Cerebral small vessel disease genomics and its implications across the lifespan

Muralidharan Sargurupremraj
University of Bordeaux

Andrew C. Heath
Washington University School of Medicine in St. Louis

Pamela A.F. Madden
Washington University School of Medicine in St. Louis

et al.

Follow this and additional works at: https://digitalcommons.wustl.edu/open_access_pubs

Please let us know how this document benefits you.

Recommended Citation
Sargurupremraj, Muralidharan; Heath, Andrew C.; Madden, Pamela A.F.; and al., et, "Cerebral small vessel disease genomics and its implications across the lifespan." Nature Communications. 11, 1. 6285 (2020). https://digitalcommons.wustl.edu/open_access_pubs/9932

This Open Access Publication is brought to you for free and open access by Digital Commons@Becker. It has been accepted for inclusion in Open Access Publications by an authorized administrator of Digital Commons@Becker. For more information, please contact vanam@wustl.edu.
White matter hyperintensities (WMH) are the most common brain-imaging feature of cerebral small vessel disease (SVD), hypertension being the main known risk factor. Here, we identify 27 genome-wide loci for WMH-volume in a cohort of 50,970 older individuals, accounting for modification/confounding by hypertension. Aggregated WMH risk variants were associated with altered white matter integrity ($p = 2.5 \times 10^{-7}$) in brain images from 1,738 young healthy adults, providing insight into the lifetime impact of SVD genetic risk. Mendelian randomization suggested causal association of increasing WMH-volume with stroke, Alzheimer-type dementia, and of increasing blood pressure (BP) with larger WMH-volume, notably also in persons without clinical hypertension. Transcriptome-wide colocalization analyses showed association of WMH-volume with expression of 39 genes, of which four encode known drug targets. Finally, we provide insight into BP-independent biological pathways underlying SVD and suggest potential for genetic stratification of high-risk individuals and for genetically-informed prioritization of drug targets for prevention trials.
As a leading cause of stroke, cognitive decline, and dementia, cerebral small vessel disease (SVD) represents a major source of morbidity and mortality in aging populations1-3. Exploring the mechanisms of SVD and their contribution to dementia risk has recently been identified as a priority research area4,5, based on its more frequent recognition with magnetic resonance imaging (MRI), its high prevalence in older community persons3,6 and the demonstration that intensive management of vascular risk factors, especially hypertension, may slow down its progression7,8. The biological underpinnings of SVD are poorly understood and no mechanism-based treatments currently are available. White matter hyperintensities of presumed vascular origin (WMH), the most common MRI-marker of SVD, can be measured quantitatively using automated software. They are highly heritable9, and confer an increased risk of stroke and dementia10, thus making them well-suited to identify potential genetic determinants of SVD and its contribution to stroke and dementia risk. WMH are most often covert, i.e., not associated with a history of clinical stroke. They are highly prevalent in the general population, and much more frequently observed than clinical stroke caused by SVD (which can be both ischemic [small vessel stroke] and hemorrhagic [deep intracerebral hemorrhage]) (Supplementary Fig. 1).

Studying the genomics of SVD also provides a powerful approach to discovery of underlying molecular mechanisms and targets in order to accelerate the development of future therapies, or identify drug repositioning opportunities10-12. Although genomic studies of WMH have been most fruitful for deciphering SVD risk variants compared with other MRI-features of SVD (lacunes, cerebral microbleeds, dilated perivascular spaces)13 or small vessel stroke14, or deep intracerebral hemorrhage15, few risk loci have been identified to date16-18. This is likely due to limited sample size of populations studied and possibly also the failure to take into account the role of hypertension (HTN), the strongest known risk factor for WMH, in confounding or modifying genetic associations. There is also mounting evidence suggesting that early-life factors play a crucial role in the occurrence of late-life vascular and neurological conditions, including SVD19, likely due to both genetic and environmental factors that may intrinsically influence the vascular substrate of SVD or modulate the brain’s resilience to SVD20-22. Identifying these early predictors could have major implications for our understanding of disease mechanisms across the lifespan and for devising effective prevention strategies.

Here, we conduct a large multiancestry meta-analysis of WMH-volume genome-wide association studies (GWAS), accounting for HTN as a potential confounder and effect modifier. We explore association of WMH risk loci with early changes in white matter microstructure on MRI using diffusion tensor imaging (DTI) in young adults in their early twenties. Last, we explore biological pathways underlying the observed genetic associations with SVD and their clinical significance through shared genetic variation and Mendelian randomization (MR) experiments with vascular risk factors and neurological traits, linking them with multiple epigenomic, transcriptomic, and drug-target databases.

Results

Genetic discovery from association analyses. Figure 1 summarizes the overall workflow of our study that included data from 50,970 participants (N = 48,454 Europeans and 2516 African-Americans) from population-based studies taking part in the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE)23 consortium and from the UK Biobank (Supplementary Data 1). The mean age of participants was 66.0 ± 7.5 years, 53% were women and 52% hypertensive (Methods, Supplementary Methods 1, and Supplementary Data 1 for cohort-specific characteristics). There was no evidence for systematic inflation of SNP-WMH association statistics at the individual cohort or meta-analysis level (Supplementary Data 3 and Supplementary Fig. 2) for the three types of analyses performed (Methods).

In the European-only SNP-main-effects analysis, 22 independent loci harbored common variants associated with WMH volume at genome-wide significance (P < 5 × 10⁻⁸, Table 1, Fig. 2), lead SNPs for each independent locus were confirmed by both LD clamping and GCTA-COJO24. Additionally, the NID2 locus reached genome-wide significance by the joint effect of multiple SNPs (P = 4.87 × 10⁻⁸, Supplementary Data 4), with P = 5.45 × 10⁻⁸ for lead SNP rs2680374, using GCTA-COJO (Methods, Supplementary Fig. 3). The African–American-only analysis identified a genome-wide significant locus at ECHDC3 (Supplementary Data 4). For loci showing heterogeneity in allelic effects across ancestry groups (PHet < 0.01), using MR-MEGA25 the ECHDC3 locus and another locus near KCNJ2 reached genome-wide significance (Table 1, Supplementary Data 5). In the HTN-adjusted model two additional loci were associated with WMH volume at P < 5 × 10⁻⁸ (PKN2 and XKRX), while three loci were no longer genome-wide significant (Table 1). The 2-df genome-wide gene-HTN interaction joint meta-analysis (JMA) did not identify any additional loci (Table 1, Supplementary Data 6). Five loci reached genome-wide significance in the small African–American-only JMA, but these were not maintained in the fixed-effects multiancestry JMA (Supplementary Data 7).

In total, 27 loci reached genome-wide significance in association with WMH volume in at least one of the aforementioned analyses, of which 18 have not previously been reported (Table 1, Fig. 2). Associations with WMH volume at these loci were similar in participants with and without HTN and when stratifying on quartiles of genetically predicted SBP and DBP levels (Methods, Supplementary Data 8-10). In aggregate, however, a weighted genetic risk score of independent WMH risk loci (WMH wGRS, Methods) showed significant 1-df interaction with HTN in association with WMH volume (βGRSxHTN = 0.15, PGRSxHTN = 0.009, Supplementary Fig. 4). One previously described risk locus for WMH burden did not reach genome-wide significance in the current analysis (near PMF1, P = 3.9 × 10⁻⁴). Of note one genome-wide significant locus (COLA2) and one suggestive locus (HTRA1, P < 5 × 10⁻⁶, Supplementary Data 11), involve genes implicated in monogenic forms of SVD26,27.

Additional, gene-based tests using MAGMA28 yielded 49 genome-wide significant associations (P < 2.8 × 10⁻⁶), of which 13 were outside GWAS loci, including the APOE gene (Methods, Supplementary Data 12). Using the Heritability Estimator from Summary Statistics (HESS)29 we found that 29 ± 2% of WMH-volume variance is explained by common and low frequency variants across the genome, the amount of heritability attributable to loci containing GWAS index SNPs being 2.4 ± 0.1%.

Implications of WMH genes across the lifespan. To examine the lifetime impact of WMH risk variants on brain structure, we explored the association of the WMH wGRS with MRI-markers of white matter microstructure in 1738 young healthy adults participating in the i-Share cohort (mean age 22.1 ± 2.3 years, 72% women). Integrity of the white matter microstructure was measured on diffusion tensor imaging (DTI) using the following metrics: fractional anisotropy (FA), mean diffusivity (MD), radial diffusivity (RD), axial diffusivity (A,D), and the recently described peak width of skeletonized mean diffusivity (PSMD)30. These MRI-markers are associated with the maturation and aging
process of white matter microstructure\textsuperscript{31,32}, and alterations in DTI metrics were shown to precede the occurrence of WMH lesions in older patients with SVD\textsuperscript{30,33}. The WMH wGRS showed significant associations with higher MD, RD, and PSMD and lower FA values in i-Share; four WMH risk loci individually showed significant associations with at least one of the DTI parameters (SH3PD2A, NMT1, KLHL24, and VCAN, Table 2). Increasing values of PSMD (but not other DTI markers) shows a trend towards association with slower information processing speed on the Stroop test in i-Share participants ($N = 1,401$, effect estimate $\pm SE$: $0.085 \pm 0.040$, $P = 0.031$), which did not survive correction for multiple testing (for three independent DTI markers). The WMH wGRS was not associated with the Stroop test in i-Share but showed a trend towards an association with poorer episodic memory performance in older community persons ($N = 24,597$, effect estimate $\pm SE$: $-0.19 \pm 0.11$, $P = 0.085$)\textsuperscript{34}.

We also examined whether genetically predicted larger WMH volume was associated with increased risk of stroke and Alzheimer-type dementia, the most common age-related neurological diseases, and with lower cognitive performance in older age, using previously reported GWAS data (Supplementary Table 2). Several genome-wide significant WMH risk loci showed significant association with ischemic stroke (three loci), all stroke and small vessel stroke (two loci each), cardioembolic stroke, deep intracerebral hemorrhage, and Alzheimer-type dementia (one locus each) (Supplementary Data 13). Using linkage disequilibrium score regression (LDSR)\textsuperscript{35}, we observed significant genetic correlation of WMH volume with all stroke, ischemic stroke, small vessel stroke, and lower general cognitive function, after Bonferroni correction for multiple testing ($P < 3.6 \times 10^{-3}$, Methods, Fig. 3, Supplementary Data 14). Using the Bayesian pairwise GWAS (GWAS-PW) approach\textsuperscript{36}, significant regional overlap (posterior probability of association for model 3, PP3 $\geq 0.90$, Methods) was observed between WMH volume and general cognitive function and between WMH volume and stroke, especially ischemic and small vessel stroke (Supplementary Data 15). This included regions previously implicated in complex and monogenic forms of stroke (FOXF2/FOXQ1\textsuperscript{37,38}, HTRA1\textsuperscript{27,37}) and cardiovascular disease (NO3)\textsuperscript{39}.

Using two-sample MR, which implements the inverse-variance weighting (IVW) method, we observed evidence for significant causal associations after Bonferroni correction for multiple testing ($P < 3.6 \times 10^{-3}$) between WMH volume and increased risk of Alzheimer-type dementia, with no statistical evidence of horizontal pleiotropy using Cochran’s Q statistic (Q-PHet $\geq 0.01$, Methods) was observed between WMH volume and stroke, especially ischemic and small vessel stroke (Supplementary Data 15). We also observed evidence for significant causal association of WMH volume with risk of any stroke, ischemic stroke, small vessel stroke, and deep intracerebral hemorrhage. There was some evidence for horizontal pleiotropy.

**Fig. 1 Study workflow and rationale.** † number of GW hits. MRI magnetic resonance imaging, CHARGE cohorts for heart and aging research in genomics epidemiology, EUR European, AFR African-American, GWAS genome-wide association study, WMH White matter hyperintensities, SNP single nucleotide polymorphism, HTN hypertension, JMA joint meta-analysis, MR-MEGA meta-regression of multi-ethnic genetic association, GWAS genome-wide, LD linkage disequilibrium, GCTA-COJO genome-wide complex trait analysis-conditional and joint analysis, MAGMA multi-marker analysis of genomic annotation, DTI diffusion tensor imaging, iSHARE internet based student health research enterprise, FA fractional anisotropy, MD mean diffusivity, RD radial diffusivity, AxD axial diffusivity, PSMD peak width of skeletonised mean diffusivity, wGRS weighted genetic risk score, GEC genetic type I error calculator, LDSR-LD score regression, GWAS-PW GWAS-pairwise analysis, HESS heritability estimator from summary statistics, EPIGWAS epigenome wide association study, TWAS transcriptome-wide association study, GTEx genotype-tissue expression, ROSMAP religious orders study and the RUSH memory and aging project, CMC common mind consortium, eQTL expression quantitative trait loci, eGene expression-associated genes, COLOC colocalisation, GREP genome for repositioning drugs.
Table 1 | Loci reaching genome-wide significance with WMH burden in one or more genetic association model.

<table>
<thead>
<tr>
<th>Region</th>
<th>Position</th>
<th>Nearest Gene</th>
<th>Analysis SNP</th>
<th>Local ( h^2 )</th>
<th>Function</th>
<th>EA/OA</th>
<th>EAF</th>
<th>SE</th>
<th>( \beta )</th>
<th>( \text{SE} )</th>
<th>( p )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Novel</td>
<td>43144218</td>
<td>NMT1§</td>
<td>rs12921170</td>
<td>0.17</td>
<td>Intronic</td>
<td>EUR/TRANS</td>
<td>0.059</td>
<td>0.047</td>
<td>0.007</td>
<td>9.82E-19</td>
<td>3.43E-18</td>
</tr>
<tr>
<td>Novel</td>
<td>31731202</td>
<td>KLHL24</td>
<td>rs5797002</td>
<td>0.12</td>
<td>Intronic</td>
<td>EUR/TRANS</td>
<td>0.050</td>
<td>0.046</td>
<td>0.006</td>
<td>9.81E-18</td>
<td>3.34E-17</td>
</tr>
<tr>
<td>Novel</td>
<td>65882074</td>
<td>CALCRL</td>
<td>rs9785772</td>
<td>0.12</td>
<td>Intronic</td>
<td>EUR/TRANS</td>
<td>0.050</td>
<td>0.046</td>
<td>0.006</td>
<td>9.81E-18</td>
<td>3.34E-17</td>
</tr>
<tr>
<td>Novel</td>
<td>16268535</td>
<td>KCNK2</td>
<td>rs6540873</td>
<td>0.03</td>
<td>Intronic</td>
<td>AFR/TRANS</td>
<td>0.027</td>
<td>0.024</td>
<td>0.006</td>
<td>7.17E-16</td>
<td>2.24E-15</td>
</tr>
<tr>
<td>Novel</td>
<td>83891379</td>
<td>KCNK2</td>
<td>rs8962910</td>
<td>0.02</td>
<td>Intronic</td>
<td>EUR/TRANS</td>
<td>0.027</td>
<td>0.024</td>
<td>0.006</td>
<td>7.17E-16</td>
<td>2.24E-15</td>
</tr>
<tr>
<td>Novel</td>
<td>29457328</td>
<td>KCNK2</td>
<td>rs4630220</td>
<td>0.02</td>
<td>Intronic</td>
<td>EUR/TRANS</td>
<td>0.027</td>
<td>0.024</td>
<td>0.006</td>
<td>7.17E-16</td>
<td>2.24E-15</td>
</tr>
<tr>
<td>Novel</td>
<td>27866281</td>
<td>KCNK2</td>
<td>rs11257311</td>
<td>0.01</td>
<td>Intronic</td>
<td>EUR/TRANS</td>
<td>0.027</td>
<td>0.024</td>
<td>0.006</td>
<td>7.17E-16</td>
<td>2.24E-15</td>
</tr>
<tr>
<td>Novel</td>
<td>18802831</td>
<td>KCNK2</td>
<td>rs21533722</td>
<td>0.01</td>
<td>Intronic</td>
<td>EUR/TRANS</td>
<td>0.027</td>
<td>0.024</td>
<td>0.006</td>
<td>7.17E-16</td>
<td>2.24E-15</td>
</tr>
<tr>
<td>Novel</td>
<td>70578698</td>
<td>KCNK2</td>
<td>rs21533722</td>
<td>0.01</td>
<td>Intronic</td>
<td>EUR/TRANS</td>
<td>0.027</td>
<td>0.024</td>
<td>0.006</td>
<td>7.17E-16</td>
<td>2.24E-15</td>
</tr>
<tr>
<td>Novel</td>
<td>70578698</td>
<td>KCNK2</td>
<td>rs21533722</td>
<td>0.01</td>
<td>Intronic</td>
<td>EUR/TRANS</td>
<td>0.027</td>
<td>0.024</td>
<td>0.006</td>
<td>7.17E-16</td>
<td>2.24E-15</td>
</tr>
<tr>
<td>Novel</td>
<td>70578698</td>
<td>KCNK2</td>
<td>rs21533722</td>
<td>0.01</td>
<td>Intronic</td>
<td>EUR/TRANS</td>
<td>0.027</td>
<td>0.024</td>
<td>0.006</td>
<td>7.17E-16</td>
<td>2.24E-15</td>
</tr>
<tr>
<td>Novel</td>
<td>70578698</td>
<td>KCNK2</td>
<td>rs21533722</td>
<td>0.01</td>
<td>Intronic</td>
<td>EUR/TRANS</td>
<td>0.027</td>
<td>0.024</td>
<td>0.006</td>
<td>7.17E-16</td>
<td>2.24E-15</td>
</tr>
<tr>
<td>Novel</td>
<td>70578698</td>
<td>KCNK2</td>
<td>rs21533722</td>
<td>0.01</td>
<td>Intronic</td>
<td>EUR/TRANS</td>
<td>0.027</td>
<td>0.024</td>
<td>0.006</td>
<td>7.17E-16</td>
<td>2.24E-15</td>
</tr>
<tr>
<td>Novel</td>
<td>70578698</td>
<td>KCNK2</td>
<td>rs21533722</td>
<td>0.01</td>
<td>Intronic</td>
<td>EUR/TRANS</td>
<td>0.027</td>
<td>0.024</td>
<td>0.006</td>
<td>7.17E-16</td>
<td>2.24E-15</td>
</tr>
<tr>
<td>Novel</td>
<td>70578698</td>
<td>KCNK2</td>
<td>rs21533722</td>
<td>0.01</td>
<td>Intronic</td>
<td>EUR/TRANS</td>
<td>0.027</td>
<td>0.024</td>
<td>0.006</td>
<td>7.17E-16</td>
<td>2.24E-15</td>
</tr>
<tr>
<td>Novel</td>
<td>70578698</td>
<td>KCNK2</td>
<td>rs21533722</td>
<td>0.01</td>
<td>Intronic</td>
<td>EUR/TRANS</td>
<td>0.027</td>
<td>0.024</td>
<td>0.006</td>
<td>7.17E-16</td>
<td>2.24E-15</td>
</tr>
<tr>
<td>Novel</td>
<td>70578698</td>
<td>KCNK2</td>
<td>rs21533722</td>
<td>0.01</td>
<td>Intronic</td>
<td>EUR/TRANS</td>
<td>0.027</td>
<td>0.024</td>
<td>0.006</td>
<td>7.17E-16</td>
<td>2.24E-15</td>
</tr>
<tr>
<td>Novel</td>
<td>70578698</td>
<td>KCNK2</td>
<td>rs21533722</td>
<td>0.01</td>
<td>Intronic</td>
<td>EUR/TRANS</td>
<td>0.027</td>
<td>0.024</td>
<td>0.006</td>
<td>7.17E-16</td>
<td>2.24E-15</td>
</tr>
<tr>
<td>Novel</td>
<td>70578698</td>
<td>KCNK2</td>
<td>rs21533722</td>
<td>0.01</td>
<td>Intronic</td>
<td>EUR/TRANS</td>
<td>0.027</td>
<td>0.024</td>
<td>0.006</td>
<td>7.17E-16</td>
<td>2.24E-15</td>
</tr>
<tr>
<td>Novel</td>
<td>70578698</td>
<td>KCNK2</td>
<td>rs21533722</td>
<td>0.01</td>
<td>Intronic</td>
<td>EUR/TRANS</td>
<td>0.027</td>
<td>0.024</td>
<td>0.006</td>
<td>7.17E-16</td>
<td>2.24E-15</td>
</tr>
<tr>
<td>Novel</td>
<td>70578698</td>
<td>KCNK2</td>
<td>rs21533722</td>
<td>0.01</td>
<td>Intronic</td>
<td>EUR/TRANS</td>
<td>0.027</td>
<td>0.024</td>
<td>0.006</td>
<td>7.17E-16</td>
<td>2.24E-15</td>
</tr>
<tr>
<td>Novel</td>
<td>70578698</td>
<td>KCNK2</td>
<td>rs21533722</td>
<td>0.01</td>
<td>Intronic</td>
<td>EUR/TRANS</td>
<td>0.027</td>
<td>0.024</td>
<td>0.006</td>
<td>7.17E-16</td>
<td>2.24E-15</td>
</tr>
<tr>
<td>Novel</td>
<td>70578698</td>
<td>KCNK2</td>
<td>rs21533722</td>
<td>0.01</td>
<td>Intronic</td>
<td>EUR/TRANS</td>
<td>0.027</td>
<td>0.024</td>
<td>0.006</td>
<td>7.17E-16</td>
<td>2.24E-15</td>
</tr>
<tr>
<td>Novel</td>
<td>70578698</td>
<td>KCNK2</td>
<td>rs21533722</td>
<td>0.01</td>
<td>Intronic</td>
<td>EUR/TRANS</td>
<td>0.027</td>
<td>0.024</td>
<td>0.006</td>
<td>7.17E-16</td>
<td>2.24E-15</td>
</tr>
<tr>
<td>Novel</td>
<td>70578698</td>
<td>KCNK2</td>
<td>rs21533722</td>
<td>0.01</td>
<td>Intronic</td>
<td>EUR/TRANS</td>
<td>0.027</td>
<td>0.024</td>
<td>0.006</td>
<td>7.17E-16</td>
<td>2.24E-15</td>
</tr>
</tbody>
</table>

For each locus, the values reaching the lowest \( P \) value in the fixed-effects transANCESTRAL meta-analysis or the fixed-effects European-only meta-analysis, respectively, is shown. EU = European, AFR = African, TRANS = transANCESTRAL meta-analysis, EUR only, TRANS only, EUR/TRANS = transANCESTRAL meta-analysis or the fixed-effects transANCESTRAL meta-analysis, and total intracranial volume, and in Model 2, which is additionally adjusted for hypertension status. Other loci reaching GW significance in the SCARGACO analysis for the main effect model.
QR values close to 1, indicating goodness of significance (pleiotropic effects) did not significantly differ from zero with Qₚ values close to 1, indicating goodness of fit of the IVW method (Methods, Supplementary Data 16)\(^{40}\).

**Shared genetic risk with vascular traits.** To assess whether genetic associations with WMH reflect known vascular mechanisms we systematically explored the shared genetic variation between WMH burden and related vascular traits, comprising established risk factors for vascular disease (systolic blood pressure [SBP], diastolic blood pressure [DBP], pulse pressure [PP], type 2 diabetes [T2D], low-density lipoprotein [LDL] cholesterol, high-density lipoprotein [HDL] cholesterol, triglycerides [TG], body mass index [BMI], glycated hemoglobin levels [HbA1c], and lifetime smoking index—a composite measure capturing smoking heaviness and duration as well as smoking initiation [SMKindex]), as well as putative risk factors for other disorders including venous thromboembolism (VTE) and migraine, using summary statistics from the most recent GWAS. Some of the latter were made available through collaborations with the relevant research consortia when the data were not publicly available (Supplementary Table 2).

First, we looked up associations of the 27 WMH risk loci with related vascular traits, including the lead WMH risk variants and nearby variants (±250 kb) in moderate to high LD (r² > 0.5). After correcting for the number of independent loci and traits tested (P < 1.3 × 10⁻⁴, Methods), 20 of the 27 WMH risk loci (74%) showed significant association with at least one other trait and/or vascular risk factors. For 13 of these, associations were found at a genome-wide significant level (Fig. 2, Supplementary Data 13). Blood pressure (BP) traits showed by far the largest number of related associations. These included systolic blood pressure (SBP), diastolic blood pressure (DBP), and pulse pressure (PP) traits showing the largest number of loci (46) associated with a given trait (upper section) or the number of loci associated with a given locus (lower section). Black arrows indicate genome-wide significant associations, and asterisks denote SNPs exhibiting unexpected directionality of associations (WMH risk allele displaying protective association with vascular or neurological traits). LOC* LOC100505841, SH3PXD2AI, COL4A2, SALL1, KLHL24, CCDC88C, TNKS, XKR6, ECHDC3, HAAO, PKN2, XKR6, AS all stroke, IS ischemic stroke, SVS small vessel stroke, CE cardioembolic stroke, ICH intracerebral hemorrhage, AD Alzheimer’s disease, BMI body mass index, LDL low-density lipoprotein, VTE venous thromboembolism, T2D type II diabetes, MIG migraine, SMKindex lifetime smoking index, WMH white matter hyperintensity.

---

**Fig. 2 Genome-wide association results with WMH burden and genetic overlap of WMH risk loci.** Circular Manhattan plot (top) displaying novel (violet) and known (dark blue) genome-wide significant WMH risk loci (dotted line: P < 5 × 10⁻⁸). Asterisks denote association signals that reach genome-wide significance only in the HTN-adjusted model (PKN2, XKR6) or in the MR-MEGA transethnic meta-analysis (KCNK2, ECHDC3). Chord diagram (center) summarizing the association of genome-wide significant risk variants for WMH burden (upper section) with vascular and neurological traits (bottom section) (P < 1.3 × 10⁻⁴, see Methods). The width of each of the stems corresponds to the number of traits associated with a given locus (upper section) or the number of loci associated with a given trait. Black arrows indicate genome-wide significant associations, and asterisks denote SNPs exhibiting unexpected directionality of associations (WMH risk allele displaying protective association with vascular or neurological traits). LOC* LOC100505841, SH3PXD2AI, COL4A2, SALL1, KLHL24, CCDC88C, TNKS, XKR6, ECHDC3, HAAO, PKN2, XKR6, AS all stroke, IS ischemic stroke, SVS small vessel stroke, CE cardioembolic stroke, ICH intracerebral hemorrhage, AD Alzheimer’s disease, BMI body mass index, LDL low-density lipoprotein, VTE venous thromboembolism, T2D type II diabetes, MIG migraine, SMKindex lifetime smoking index, WMH white matter hyperintensity.
Table 2 Association trend of WMH risk loci with white matter microstructures in young adults.

<table>
<thead>
<tr>
<th>Variables</th>
<th>WMH risk loci</th>
<th>Nearest gene</th>
<th>Region</th>
<th>Nearest gene</th>
<th>Region</th>
<th>Nearest gene</th>
<th>Region</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>SE</td>
<td>P</td>
<td>B</td>
<td>SE</td>
<td>P</td>
<td>B</td>
</tr>
<tr>
<td>PsMD</td>
<td>0.077</td>
<td>0.022</td>
<td>0.111</td>
<td>0.079</td>
<td>0.023</td>
<td>0.079</td>
<td>0.081</td>
</tr>
<tr>
<td>RD (10^-4 mm²/s)</td>
<td>-0.002</td>
<td>0.002</td>
<td>0.911</td>
<td>-0.002</td>
<td>0.002</td>
<td>0.911</td>
<td>-0.002</td>
</tr>
<tr>
<td>FA</td>
<td>-0.001</td>
<td>0.001</td>
<td>0.981</td>
<td>-0.002</td>
<td>0.002</td>
<td>0.951</td>
<td>-0.002</td>
</tr>
<tr>
<td>AD (10^-4 mm²/s)</td>
<td>0.009</td>
<td>0.002</td>
<td>0.002</td>
<td>0.010</td>
<td>0.002</td>
<td>0.002</td>
<td>0.011</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2 Association trend of WMH risk loci with white matter microstructures in young adults.

Biological interpretation of association signals. We used EPIGWAS\textsuperscript{44} to test for cell-type enrichment of WMH association signals using chromatin marks previously shown to be cell-type specific and associated with active gene-regulation (Methods). WMH risk loci were significantly enriched in enhancer (H3K4me1) and promoter sites (H3K4me3) in cell-types derived from the brain, neurosphere (developing brain), vascular tissue, digestive, epithelial and muscle tissues, as well as human embryonic stem cells after removing WMH risk loci associated with BP (Supplementary Data 19). Analysis of brain-specific single-cell expression data in mice using MAGMA.celltyping\textsuperscript{45} (Methods) revealed significant enrichment of highly cell-type-specific genes in endothelial mural cells and nominally significant enrichment for vascular leptomeningeal cells, oligodendrocytes, oligodendrocyte precursors, and ependymal astrocytes; results were substantially unchanged after removing WMH risk loci associated with BP (Supplementary Data 20).

To functionally characterize and prioritize individual WMH genomic risk loci we performed transcriptome-wide association studies (TWAS) using TWAS-Fusion\textsuperscript{46}, WMH-SNP association statistics from the main effects (EUR-only) and weights from 23 significant associations with WMH risk variants, 16 loci (59.3%) being associated with SBP, DBP, and/or PP. Further significant associations with WMH risk variants were observed for BMI (8 loci), T2D (5 loci), SMKindex (3 loci), and lipid traits (3 loci), one locus (at XKR6) being notably shared with all these risk factors. Seven loci (CI6orf95, ECHDC3, MIN1, NID2, SALL1, VCAN, KCNK2), none of which were reported previously as WMH risk loci, appear not to be associated with any of the vascular traits explored, suggesting other underlying biological pathways.

Second, we explored the genome-wide and regional genetic overlap between WMH volume and related vascular traits. Mean \(X^2\) ranged between 1.06 and 3.99 suggesting strong polygenicity for all investigated traits. The impact of possible sample overlap was estimated to be negligible using LDSR\textsuperscript{35} (Supplementary Data 14). We observed significant \((P < 3.6 \times 10^{-3})\) genetic correlation of larger WMH volume with higher SBP, DBP, SMKindex, BMI and increased risk of VTE. Using GWAS-PW\textsuperscript{36} and HESS\textsuperscript{33} (Methods), we identified 16 genomic regions harboring shared genetic risk variants with at least one other vascular trait, predominantly BP traits, but also BMI, lipid levels and SMKindex (PPA3 ≥ 0.90, Fig. 3, Supplementary Data 15).

Third, we explored the causal relations between the aforementioned vascular traits and WMH volume using two-sample MR\textsuperscript{42} (RadialMR\textsuperscript{40}, Methods), implementing the IVW method. We observed significant \((P < 3.6 \times 10^{-3})\) association of genetically predicted SBP, DBP, PP, SMKindex and T2D with larger WMH volume and of genetically predicted migraine with smaller WMH volume (Fig. 4, Supplementary Data 17). After removal of potentially pleiotropic outlier variants; for SBP, DBP, PP and SMKindex the MR-Egger intercept was nonsignificant, indicating no residual pleiotropy and suggesting causal association with WMH volume (Methods). For migraine and T2D in contrast there was evidence of residual pleiotropic effects (significant MR-Egger intercept, Supplementary Data 17) after removal of potentially pleiotropic outlier variants, and the association became only nominally significant for migraine. Importantly, associations of genetically predicted SBP and DBP with WMH volume remained significant after adjustment for HTN, and in participants with and without HTN (Supplementary Data 17), highlighting that higher levels of BP are likely causally associated with larger WMH volume even below BP thresholds typically used for the definition of hypertension (SBP ≥ 140 mmHg or DBP ≥ 90 mmHg or antihypertensive drug intake)\textsuperscript{43}.
gene-expression reference panels from blood, arterial, and brain tissues (Supplementary Methods 2). We also included non-publicly available gene-expression weights from the dorsolateral prefrontal cortex (DLPFC) of 494 older community-dwelling participants (Methods) 47,48. TWAS-Fusion identified 201 transcriptome-wide significant associations with WMH, conditionally significant on the predicted expression of a TWAS-associated gene, including 21 with splicing quantitative trait loci (sQTLs) regulating highly tissue-specific gene isoforms in DLPFC (Fig. 5, Supplementary Data 21). To rule out that observed associations reflect the random overlap between expression (eQTLs) and noncausal WMH risk variants, a colocalization analysis (COLOC)49 was performed at each significant locus, to estimate the posterior probability of a shared causal variant (PP4 ≥ 75%) between the gene expression and trait association (Methods). Colocalization was observed for 96 TWAS significant eQTLs (48%, Fig. 5): of these, 54 mapped to 8 WMH genome-wide risk loci and 16 expression-associated genes (eGenes), while 42 mapped to 12 distinct loci that were not genome-wide significant in the WMH GWAS and 23 eGenes. These additional putative WMH risk loci require confirmation in follow-up studies. Leveraging histone regulatory mark information from blood, arterial, and brain tissues (Methods, Supplementary Data 21), we observed that the majority (89%) of TWAS signals overlapped with enhancer and/or promoter elements, including eQTLs exhibiting weaker colocalization probability (PP4 < 75%). Larger WMH volume was associated with either upregulated or downregulated gene expression, the directionality being mostly consistent across broad tissue categories (Fig. 5). We found evidence for colocalization of WMH risk variants with eQTLs in brain tissues (28 eGenes), vascular tissues (15 eGenes), and blood (6 eGenes). Some eGenes (KLHL24, NMT1, DCAKD, KANSL1, AMZ2P1) showed evidence for colocalization in multiple tissues, and for some WMH risk loci (chr2q33.2, chr17q25.1, 3p24.3 (ZNF385D), 3p13 (FOXP1), 17q25.1 (TRIM47), 1q21.3 (GOLPH3L), 19q13.32 (APOE), 22q13.31 (RIBC2), 9p13.3 (PTENP1), 6p22.2 (LRRC16A), 2p16.1 (EFEMP1), 16p23.2 (CDYL2), 11q24.3 (ADAMTS8), 12q24.21 (TBX3), 10p12.31 (PLXDC2), 1p22.2 (PKN2), 17p11.2 (EPN2), 16q12.2 (RBL2), 6p25.3 (FOXF2), 13q34 (COL4A2), 2q33.2 (ALS2CR8), 10q26.3 (HTRA1), 2q32.1 (CALCRL), 10q24.33 (SH3PXD2A), 7q36.1 (NOS3), 5q23.2 (LOC100505841)).

**Fig. 3 Shared genetic architecture of WMH at genome-wide and regional level** Color coded for the direction of effect (Green: Positive genetic correlation; Red: Negative genetic correlation). The LD-score regression (LDSR) axis shows evidence for genome-wide correlations (after Bonferroni correction for multiple testing $P < 3.6 \times 10^{-3}$, Methods), with the size of the nodes corresponding to the level of significance of the association. The GWAS-pairwise (PW) axis shows evidence for regional level overlap of association signals between WMH burden and related vascular and neurological traits (PPA3 ≥ 0.90, Methods). For any given region, the nearest gene (in brackets) to the top SNP associated with WMH is shown. Bivariate heritability estimator from summary statistics ($p$-HESS) was used to infer directionality of shared association signals (Methods) and asterisks denote an unexpected directionality of association. SBP systolic blood pressure, DBP diastolic blood pressure, PP pulse pressure, AS all stroke, IS ischemic stroke, SVS small vessel stroke, CE cardioembolic stroke, BMI body mass index, HDL high-density lipoprotein, LDL low-density lipoprotein, VTE venous thromboembolism, GCF general cognitive function, SMKindex lifetime smoking index, WMH white matter hyperintensity.
chr17q21.31) colocalized variants associated with multiple eGenes with distinct tissue specificities (Fig. 5, Supplementary Data 21). WMH risk variants at the chr2q33 locus for instance showed evidence for colocalization with eQTLs for NBEAL1 in nerve and arterial tissues, for ICAIL and KRTBP15 in brain tissues, and for CARFI in right atrial appendage.

Among the 39 WMH eGenes with high colocalization probability (Fig. 5, Supplementary Data 21), 4 (MAPT, CRHR1, CALCR, NOTCH4) are registered as targets of approved drugs in the DrugBank database and the Therapeutic Target Database (Supplementary Data 22).

### Discussion

This largest genetic study to date on complex SVD leveraging genetic and brain-imaging information from 50,970 older community persons, triples the number of genetic loci associated with cerebral SVD and shows that this genetic risk results in detectable brain changes among asymptomatic young adults in their twenties. We further demonstrate the importance of higher BP as a risk factor for WMH even below clinical thresholds for HTN. MR analysis provides strong evidence for causal links of genetically determined WMH volume with risk of ischemic stroke, intracerebral haemorrhage, and Alzheimer-type dementia in later life. Importantly, we also provide insight into molecular pathways underlying SVD, highlighting relevant tissue and cell types, and suggest potential for genetic stratification of high-risk individuals and for genetically informed prioritization of drug targets for prevention trials.

Our approach focusing on the most common brain-imaging feature of SVD appears to be more powerful than GWAS of the small vessel stroke subtype to identify risk loci for SVD. Indeed, no new small vessel stroke risk locus was identified in MEGA-STROKE, the largest stroke GWAS meta-analysis to date. We show a strong association between genetically determined WMH burden and risk of stroke in the general population, notably both risk of ischemic stroke and of intracerebral hemorrhage. While corroborating epidemiological observations, this has never been demonstrated using genetic instrumental variables, providing evidence for causality. This prompts greater caution with the common empirical prescription of antplatelet therapy in persons with extensive WMH in the absence of clinical stroke, given the potential detrimental effects on intracerebral hemorrhage risk, and suggests the need for randomized clinical trials to determine the risk/benefit ratio of antplatelet therapy in this setting.

The significant association we describe between genetically determined WMH burden and Alzheimer-type dementia also has potential important implications for prevention. It strengthens recent epidemiological evidence that WMH is associated not only with an increased risk of all and vascular dementia, but also of neurodegenerative Alzheimer-type dementia, providing for the first time evidence for causality using the WMH wGRS as an instrumental variable. Because of the proven ability to treat vascular risk factors, understanding and targeting the biological mechanisms of the vascular contribution to cognitive impairment and dementia, and specifically how cerebral SVD contributes to the molecular pathology of Alzheimer disease, are areas of intense research and clinical interest, especially given the current

### Fig. 4 Mendelian randomization results of vascular risk factors with WMH burden (box A) and WMH burden with neurological traits (box B). Point estimates and confidence intervals (blue) from the inverse-variance weighted (IVW) method are shown along with the point estimates and 95% confidence interval (black) from sensitivity analyses after filtering out potentially pleiotropic outlier variants. The intercept and p-value from the MR-Egger method is displayed on the far right (an intercept term significantly differing from zero at the conservative threshold of $P < 0.05$ suggests the presence of directional pleiotropy). SBP systolic blood pressure, DBP diastolic blood pressure, PP pulse pressure, HTN hypertensive, NT normotensive, Str. stratum, AS all stroke, IS ischemic stroke, SVS small vessel stroke, ICH intracerebral hemorrhage, AD Alzheimer’s disease, T2D type II diabetes, SMKindex lifetime smoking index, WMH white matter hyperintensity, Model 1 Main effects adjusted for age, sex, principal components for population stratification, intracranial volume; Model 2 Model 1 + hypertension status.
abscence of other efficient therapies. Our results suggest that WMH should be considered a major target for preventative interventions, to mitigate not only the risk of stroke and vascular cognitive impairment but also of Alzheimer-type dementia, and support the rationale of innovative trials using WMH progression as a surrogate or intermediate endpoint for cognitive decline and dementia.

Over half of identified WMH risk loci are associated with higher BP levels. Moreover, using MR we provide evidence for a causal association between higher BP and larger WMH volume, notably even in participants without clinically defined HTN at the time of the MRI. Indeed, associations of genetically predicted SBP and DBP with WMH volume remained significant in participants without HTN, highlighting that higher levels of BP are likely causally associated with larger WMH volume even below BP

Moreover, using MR we provide evidence for a causal association between higher BP and larger WMH volume, notably even in participants without clinically defined HTN at the time of the MRI. Indeed, associations of genetically predicted SBP and DBP with WMH volume remained significant in participants without HTN, highlighting that higher levels of BP are likely causally associated with larger WMH volume even below BP thresholds typically used for the definition of HTN. Considering the recent conclusions from the SPRINT-MIND trial suggesting that more drastic lowering of BP in persons with HTN is associated with slower progression of WMH volume and a lower risk of developing the combined outcome of mild cognitive impairment or dementia, our results suggest that trials to test a similar impact of intensive BP lowering in high-risk individuals are warranted. We additionally show strong causal association between increased exposure to cigarette smoking over the lifetime (lifetime smoking index) and increased WMH burden, as has recently been described in relation with stroke risk, providing some additional evidence for the relevance of smoking cessation to prevent vascular brain injury and specifically SVD.

Importantly, a quarter of the identified WMH risk loci reflect molecular mechanisms that are not mediated by BP or other known vascular risk factors, two of these (NID2, VCAN), along with the COL4A2 and EFEMP1 locus, implicate genes encoding matrisome proteins, involved in cell membrane structure and representing core components of the extracellular matrix (ECM). Converging evidence from experimental models for monogenic SVD suggest that perturbations of the matrisome play a central role in disease pathophysiology. Our findings suggest that these could also be relevant for sporadic forms of SVD. NID2 encodes nidogen, an ECM glycoprotein and a major component of basement membranes and is recognized as having a role in post-stroke angiogenesis. Mutations in COL4A1/2, encoding collagen another basement membrane component already are known causing monogenic SVD, VCAN, which we also found to be associated with white matter microstructure in young adults, encodes versican, a proteoglycan involved in cell adhesion and ECM assembly. In CARASIL, a monogenic SVD caused by HTRA1 mutations, accumulation of versican in the thickened arterial wall was observed. Versican also can form complexes that inhibit oligodendrocyte maturation and remyelination.

EFEMP1 encodes fibulin 3, an ECM glycoprotein localized in the basement membrane, and a proteolytic target of serine protease HTRA1. Other previously unreported WMH risk loci that we have identified include KCNK2 that encodes Twik-related K+ channel (TREK1), a 2-pore-domain background ATP-sensitive potassium channel expressed throughout the central nervous system, more prominently in fetal than in adult brain. ECHDC3, near a distinct locus (r^2 < 0.01) previously implicated in Alzheimer disease, MINI, which previously has been causally related to familial meningiomas, and XKR6, which has been associated with risk of systemic lupus erythematosus.

Our results provide important insight into the lifetime impact of genetic risk for SVD. Indeed, WMH risk variants observed in older adults were already associated with changes in DTI markers of white matter integrity in young adults in their early twenties. Of these, PSMD, a DTI metric recently described to be more...
strongly correlated with cognitive performance in older persons (patients with sporadic or monogenic SVD and older community persons) than any other MRI-marker of SVD, was already showing nominal association with lower cognitive performance in young adults. This finding requires confirmation in future independent samples. The association of the WMH wGRS with subtle changes in white matter microstructure in young adults, if confirmed in independent samples, has potential important implications for the timing and paradigm of prediction and prevention of SVD progression and complications. It could reflect that biological pathways contributing to WMH at an older age already have a significant impact on brain microstructure in young adults, possibly reflecting a very early stage of SVD (typically characterized by reduced FA and increased MD and PSMD). DTI changes and WMH have been suggested to be dependent physiological processes occurring within consecutive temporal windows in older patients with SVD. Alternatively, observed associations might also reflect pleiotropy between SVD genes and genes influencing brain maturation, as the mean age of iShare participants corresponds to the peak of white matter maturation. On average FA tends to increase during childhood, adolescence, and early adulthood and then decline in middle-age, while the reverse is observed for MD and higher MD could also reflect an impaired or delayed maturation or a premature aging process. The significant association of the WMH wGRS with RD but not AxD could potentially suggest that this is predominantly reflecting an impact on myelination of fiber tracts in line with involvement of oligodendroglial dysfunction in early SVD pathology. Future follow-up studies in a longitudinal setting are warranted to better understand the impact of genetically predicted WMH burden on the progression of white matter microstructural changes observed already in young adults and on their link with SVD and its complications.

Functional characterization revealed enrichment of WMH risk variants in regulatory marks in brain and neurosphere and in single-cell gene-expression levels in endothelial mural cells (as for clinical stroke). Gene prioritization using TWAS revealed that several WMH risk loci colocalized with eQTL for multiple genes with distinct tissue specificities. This pattern could potentially partly explain why association of such loci with WMH volume remained unchanged after controlling for the presence of HTN, although they were associated at genome-wide significant level with both BP and WMH. Of the 39 eGenes identified by TWAS four encode known drug targets. MAPT is a drug target under investigation for neurodegenerative disorders: the eQTL colocalizing with the WMH risk variant is an sQTL for the MAPT isomorph in DLPFC and TWAS suggest that larger WMH volume is associated with upregulated MAPT expression. CALCRL encodes a component of the Calcitonin Gene Related Peptide receptor. TWAS suggest that lower abundance of the CALCRL transcript in arterial and nerve tissue and higher abundance in blood are associated with larger WMH volume. Monoclonal antibodies against CALCRL have recently been developed for the treatment of migraine.

We acknowledge limitations. We were underpowered for detecting additional risk variants for WMH after accounting for presence of HTN in the 2-df JMA gene-HTN interaction model. Recognizing that blood pressure is also highly variable and that a one-time blood pressure measurement may not reflect the long-term exposure of participants to high blood pressure levels, we conducted secondary analyses stratifying on quartiles of genetically predicted SBP and DBP levels, yielding similar results. In aggregate, a weighted genetic risk score of independent genome-wide significant WMH risk loci showed a significant 1-df interaction with HTN status in association with WMH volume, suggesting that effect modification of genetic associations by HTN may exist, but that to discover them at the level of individual loci likely will require much larger samples. While we were able to use gene-expression data from many tissues for TWAS, such data are lacking for certain tissues that may be relevant for WMH (e.g., small brain vessels, microglia). Finally, our study population is predominantly of European ancestry (95%) limiting our ability to extrapolate our conclusions to other ancestry groups.

In conclusion we have identified 27 genetic risk loci for WMH volume, of which two thirds are not previously reported, and provided additional insight into their association with structural brain changes in very young adults, their clinical significance and the importance of high BP as a risk factor below clinical thresholds. Our results also point to molecular pathways underlying SVD that are not mediated by vascular risk factors and suggest potential for genetic stratification of high-risk individuals and genetically informed prioritization of drug targets for prevention trials.

Methods

Study population. The study population comprised 23 population-based studies from the CHARGE consortium comprising a total of 24,182 individuals of European (N = 21,666) and African-American (N = 2516) ancestry, along with 26,788 community participants of European origin from the UK Biobank. In total, 50,970 participants were available for testing main genetic effects and 48,524 participants for the gene-hypertension interaction analysis (information on HTN status was missing in 2446 participants). Individuals with a history of stroke (or MRI-defined brain infarcts involving the cortical gray matter), or other pathologies that may influence the measurement of WMH (e.g., brain tumor, head trauma, etc.), at the time of MRI were excluded from analyses. Study participants in all participating cohorts gave written informed consent for phenotype quantification and use of genetic material (Supplementary Methods 1, Supplementary Table 1).

Phenotypes. MRI scans were obtained from scanners with field strengths ranging mostly from 1.5 to 3.0 Tesla and interpreted using a standardized protocol blinded to clinical or demographic features. In addition to T1 and T2 weighted scans along the axial plane, some cohorts included fluid-attenuated inversion recovery (FLAIR) and/or proton density (PD) sequences for better differentiation of WMH from cerebrospinal fluid. The vast majority of participating cohorts (>92% of all participants) used fully automated software to quantify WMH volume, with two cohorts using validated, visually guided semi-quantitative scales in older study subsets (Supplementary Table 1). WMH volume measures were inverse normal transformed to correct for skewness and account for differences in WMH quantification methods. Blood pressure measurements that are closest to the MRI scan were used to define HTN status. Participants with a SBP ≥ 140 mmHg, a DBP ≥ 90 mmHg, or taking antihypertensive medication were classified as having HTN.

Genotyping and imputation. Genome-wide genotyping platforms are described in Supplementary Data 2. Prior to imputation, sample-specific quality control (QC) on heterozygosity, missingness, gender mismatch, cryptic relatedness, and analysis of principal components (PC) for population stratification, as well as SNP-level QC on genotyping call rate and Hardy–Weinberg equilibrium were applied (Supplementary Data 2). Samples and SNPs passing the cohort-specific QC measures were then imputed to the 1000 genomes cosmopolitan panel phase 1 version 3 (1000G v13) for CHARGE cohorts, while for the UK Biobank the dataset version 3 was imputed to the combined UK10K and Haplotype Reference Consortium (HRC) reference panels.

Genome-wide association analyses. Each participating study conducted ancestry-specific analyses using linear regression and assuming additive genetic effects under three models: (1) marginal genetic association test of WMH volume (SNP-main effect); (2) SNP-main effect adjusted for HTN status; (3) and joint association test of both SNP-main and SNP-by-HTN interaction effects in relation with WMH volume:

\[ Y = \beta_0 + \beta_S SNP + \beta_Cov + \varepsilon \]  

\[ Y = \beta_0 + \beta_S SNP + \beta_Cov + \beta_{SNP-HTN} + \varepsilon \]  

\[ Y = \beta_0 + \beta_S SNP + \beta_Cov + \beta_{SNP-HTN} + \beta_{SNP-HTN} + \varepsilon \]  

where SNP corresponds to the dosage of a given genetic (G) variant, Cov is the vector of covariates, GEnv is the SNP-by-HTN interaction effect and \( \varepsilon \) values are the corresponding regression coefficients.
coefficients and error covariance (\(\lambda\)) of \(\beta\). The joint model (Model 3) provides effect estimates of \(G\) and \(G\times G\), their robust standard errors (SE) and robust covariance matrices, and a joint \(P\) value from a 2 degree-of-freedom (df) Wald test. Robust estimates of SEs and covariance matrices were used.

All analyses were adjusted for age, sex, PCs of population stratification and intracranial volume (ICV). Adjustment for ICV was not performed for studies using visual grading of WMH burden, as visual grades are inherently normalized for brain size34 (Supplementary Data 1, Supplementary Methods 1).

**Genome-wide association meta-analyses.** A custom harmonization script along with the R package EasyQC23 was used to perform the QC of cohort-specific GWAS results. SNPs with minor allele frequency (MAF) lower than 1%, poor imputation quality (R\(^2\) < 0.80), or a product of MAF, \(R^2\) and sample size less than 10, or 15 for the 2-df interaction analysis were excluded.

Inverse-variance weighted meta-analysis was conducted using METAL24, first within each cohort, and then followed by a meta-analysis of the ancestry-specific results. A patch implemented in the METAL25 software was used to perform a 2-df joint meta-analysis (JMA) with inverse-variance weighting. For cohort-specific GWAS results with genomic inflation factors (\(\lambda\)) exceeding 1, genomic control (GC) correction was applied to correct for any residual population stratification. After meta-analysis only SNPs represented in more than half of participating studies and/or more than half of sample size and with no evidence of between-study heterogeneity (PH\(_{\text{Het}}\) > \(1 \times 10^{-6}\)) were considered. Quantile-Quantile (QQ) plots of the P-values (observed versus expected) in the GWAS for the different models tested are presented along with the genomic inflation factor (\(\lambda\)) (Supplementary Data 1). HESS30 estimates of heterogeneity at the genetic level specific effects that is specifically due to ancestry differences is not addressed by the traditional fixed-effects meta-analysis, a multiancestry meta-regression was carried out. For each variant, SNP-main allele effects on WMH volume across GWAS were estimated in a linear regression framework weighting on the inverse of the variance of effect estimates weighted on the axes of genetic variation derived from pairwise allele frequency differences, as implemented in MR-MEGA25. It provides two heterogeneity estimates, one that is correlated with ancestry (\(PH_{\text{ANC}}\)) and accounted for in the meta-regression and the residual heterogeneity that is not due to population genetic differences (\(PH_{\text{RES}}\)).

Across all association models, the power to reject the null hypothesis of no association at the genome-wide (GW) level was set at \(P < 5 \times 10^{-8}\). Independent SNPs within genome-wide risk loci were determined by performing linkage disequilibrium (LD) based clumping implemented in PLINK using both a physical distance of \(\pm 1\) megabase (Mb) and an LD threshold of \(r^2 > 0.10\) from the index SNP of a given locus28. For constructing the LD matrix, ancestry-specific (European [EUR], African–American ancestry in South-West USA [ASW]) 1000 G p1v3 reference panels were used for ancestry-specific results and the merged (EUR +ASW) reference for multiancestry results. Stepwise conditional regression and joint analysis (cojo) implemented in GCTA25 was performed to further validate the independent effects on the main-effects GWAS in Europeans only. GCTA-2COJO additionally identifies signals with GW (\(P < 5 \times 10^{-8}\)) association level due to the LD adjusted joint effect of several neighbouring SNPs, selected based on an association prior of \(P < 1 \times 10^{-7}\). Genotypes of 6489 unrelated European individuals from the Three City (3 C) study77 were used to generate the LD matrix. Finally, gene-based association tests were conducted using MAGMA26, with \(P < 2.8 \times 10^{-6}\) as a gene-wide significance threshold. Gene regions with SNPs not reaching GW significance for WMH and/or not in LD \((r^2 < 0.10)\) with the lead WMH SNP were considered as novel.

**WMH heritability estimates.** LD-score regression (LDSR) was used to distinguish polygenicity from confounding due to population stratification or cryptic relatedness29 and to estimate the GW heritability by regressing the LD-score (measure of linked SNPs) against the chi-square association statistics of WMH volume from the European-only analysis. To address the infinitesimal-model assumption used by variance-component methods such as LDSR, we applied the heritability estimator from summary statistics (HESS)29 to estimate local SNP-level heritability. HESS does not assume any effect size distribution and by weighted summation of the ratios of beta-coefficients from the supplementary information (Supplementary Table 1), to derive DTI metrics measuring the integrity of the white matter microstructure, including fractional anisotropy (FA) and mean, radial and axial diffusivity (MD, RD, A,D), as well as peak width of skeletonized mean diffusivity (PSMD). PSMD was calculated using a fully automated method via a shell script (www.psmd-marker.com) and the correlation estimated from regions with null association evidence (PPA < 0.90). To ensure that the association of lead WMH risk variants (\(n = 27\)) with related vascular and neurological traits. For each of the related traits, association statistics of SNPs falling in a window of \(\pm 250\) kb around each of the lead WMH SNP were retrieved and SNPs satisfying the multiple testing threshold defined by \(H\) was used for the effective number of LD independent markers per locus, as implemented in Genetic Type 1 error calculator (GE20) were retained. Only SNPs showing an association with a related vascular or neurological trait at \(P < 1.3 \times 10^{-8}\) (accounting for 14 independent traits and 27 independent loci) and in moderate to high lead with the WMH SNP \((r^2 > 0.50)\) are reported. The correlation matrix estimated between the traits using individual-level DTI data from the 3 C study77 was used to estimate the number of independent traits by applying the Matrix Spectral Decomposition (matSpDlite) method (http://neurogenetics.qimrberghofer.edu.au/matSpDlite/).

**Shared genetic architecture of WMH with related traits.** We systematically explored the genetic overlap of WMH SNP-main-effects (in the European-only analysis) with (i) neurological traits (any stroke, ischemic stroke, small vessel stroke, large artery stroke, cardiomebolic stroke; any, and deep, and lobar intracerebral hemorrhage) (general cognition) and Alzheimer's (ii) vascular risk factors and traits (SBP, DBP, PP, HDL-cholesterol, LDL-cholesterol, TG, BMI, T2D, HbA1c, SMIndex, VTE, and migraine). We acquired summary statistics of European-only analyses for these traits, using the latest largest GWAS, seeking collaboration with the relevant consortia when the data were not publicly available (Supplementary Table 1). We first explored the association of lead WMH risk variants (\(n = 27\)) with related vascular and neurological traits. For each of the related traits, association statistics of SNPs falling in a window of \(\pm 250\) kb around each of the lead WMH SNP were retrieved and SNPs satisfying the multiple testing threshold defined by \(H\) was used for the effective number of LD independent markers per locus, as implemented in Genetic Type 1 error calculator (GE20) were retained. Only SNPs showing an association with the related vascular or neurological trait at \(P < 1.3 \times 10^{-8}\) (accounting for 14 independent traits and 27 independent loci) and in moderate to high lead with the WMH SNP \((r^2 > 0.50)\) are reported. The correlation matrix estimated between the traits using individual-level DTI data from the 3 C study77 was used to estimate the number of independent traits by applying the Matrix Spectral Decomposition (matSpDlite) method (http://neurogenetics.qimrberghofer.edu.au/matSpDlite/).

Using LDSR26, genetic correlation estimates between WMH and the aforementioned neurological and vascular traits were derived. For each of the related traits, association statistics of SNPs within a window of \(\pm 250\) kb around each of the lead WMH SNP were retrieved and SNPs satisfying the multiple testing threshold defined by \(H\) was used for the effective number of LD independent markers per locus, as implemented in Genetic Type 1 error calculator (GE20) were retained. Only SNPs showing an association with a related vascular or neurological trait at \(P < 1.3 \times 10^{-8}\) (accounting for 14 independent traits and 27 independent loci) and in moderate to high lead with the WMH SNP \((r^2 > 0.50)\) are reported. The correlation matrix estimated between the traits using individual-level DTI data from the 3 C study77 was used to estimate the number of independent traits by applying the Matrix Spectral Decomposition (matSpDlite) method (http://neurogenetics.qimrberghofer.edu.au/matSpDlite/).

Using LDSR, genetic correlation estimates between WMH and the aforementioned neurological and vascular traits were derived. For each of the related traits, association statistics of SNPs within a window of \(\pm 250\) kb around each of the lead WMH SNP were retrieved and SNPs satisfying the multiple testing threshold defined by \(H\) was used for the effective number of LD independent markers per locus, as implemented in Genetic Type 1 error calculator (GE20) were retained. Only SNPs showing an association with a related vascular or neurological trait at \(P < 1.3 \times 10^{-8}\) (accounting for 14 independent traits and 27 independent loci) and in moderate to high lead with the WMH SNP \((r^2 > 0.50)\) are reported. The correlation matrix estimated between the traits using individual-level DTI data from the 3 C study77 was used to estimate the number of independent traits by applying the Matrix Spectral Decomposition (matSpDlite) method (http://neurogenetics.qimrberghofer.edu.au/matSpDlite/).

**Mendelian randomization (MR).** For each vascular trait genetic variant (instrument) details were retrieved from the latest largest GWAS (Supplementary Table 2). Only independent SNPs \((r^2 < 0.10,\) based on 1000 G EU panel) reaching genome-wide significance were included as recommended32. Similarly, 25 independent WMH risk variants from the SNP-main effects were used as instruments to test the association of genetically predicted WMH volume with neurological traits. A comprehensive causal effect of an exposure on a phenotype was estimated, using the inverse-variance weighting (IVW) method, by the weighted sum of the ratios of beta-coefficients from the SNP-outcome associations for each
variant (j) over corresponding beta-coefficients from the SNP-exposure associations (β). The ratio estimates from each genetic variant were averaged after weighting on the inverse variance (Wj) of j across L uncorrelated SNPs, implemented as an R package radialMR (available through CRAN repositories)\(^\text{40}\).

\[
\text{RIVW} = \frac{\sum_{j=1}^{L} W_j \beta_j}{\sum_{j=1}^{L} W_j}
\]

Effect alleles for each risk variant were defined as the allele associated with increase in the corresponding trait values. A p value < 3.6 x 10^-3 correcting for 14 independent traits was considered significant. Cochran’s Q statistic was used to test for the presence of heterogeneity (Qcoh < 0.01) due to horizontal pleiotropy that occurs when instruments exert an effect on the outcome and exposure through independent pathways. Inflammatory outlier SNPs that have the largest contribution to the global Cochran’s Q statistic are identified by regressing the predicted causal estimate against the inverse-variance weights. After excluding the inflammatory outlier SNPs, the IVW test was repeated along with MR-Egger regression\(^\text{83}\). Relative goodness of fit of the MR-Egger over the IVW approach was quantified using the QK statistic, which is the ratio of the statistical heterogeneity around the MR-Egger fitted slope divided by the statistical heterogeneity around the IVW slope. A QK close to 1 indicates that MR-Egger is not a better fit to the data and therefore offers no benefit over IVW\(^\text{40}\). Non-significant MR-Egger intercept was used as an indicator to formally rule out horizontal pleiotropy.

Cell and tissue type enrichment analysis. Association statistics from the WMH SNP-main effects (European-only) were used to test cell/tissue-specific enrichment. First we used the EPIGWAS software\(^\text{44}\) and histone marks for promoters (H3K4me3) and enhancers (H3K4me1) from publicly available information on tissue-specific histone regulatory marks (Supplementary Methods 2). EPIGWAS calculates specificity scores for the lead WMH risk variant and its proxies (r^2 ≥ 0.80, HapMap 2 EUR) based on the distance to the strongest ChIP-seq signal, and estimates enrichment significance by comparing the relative proximity and specificity of the test set with 10,000 sets of matched background (using permutation). Bonferroni correction for the number of histone marks tested was applied (P < 2.5 x 10^-3). Second, we used MAGMA\(^\text{48}\) (gene-property analysis) and differentially expressed gene sets from the single-cell transcriptomic (scRNA) data in mouse brain from the Karolinska Institute. MAGMA generates gene-level association statistics by combining SNP p-values in a specified window (10 kb upstream and 1.5 kb downstream of each gene) accounting for LD (1000 G EUR) and under a linear regression framework performs a one-sided test between the association of genes with WMH volume and cell specificity. Using the MAGMA.celltyping R package\(^\text{45}\), scRNA expression values were obtained from five different mouse brain regions (neocortex, hippocampus, hypothalamus, striatum, and midbrain)\(^\text{84}\). A gene-expression specificity metric for each cell-type was calculated by dividing the expression level in a given cell type by the sum of the expression levels from all cell types (i.e., genes with high expression levels in two or more cell types will get a lower specificity measure than genes with high expression levels in a single-cell type), followed by binning the metric value to 40 equally sized bins. The MAGMA one-sided test was then used to test for enrichment between the top 10 percentile bins (bins with higher cell-specificity) in each cell type. Bonferroni correction for the number of cell type tests was applied (P < 2.1 x 10^-3).

Transcriptome-wide association study and colocalization. We performed transcriptome-wide association studies (TWAS) using the association statistics from the WMH SNP-main effects (European-only) and weights from 22 publicly available gene-expression reference panels (Supplementary Methods 2) from blood (Netherlands Twin Registry, NTR; Young Finns Study, YFS), arterial (Genotype-Tissue Expression, GTEx), brain (GTEx, CommonMind Consortium, CMC) and peripheral nerve tissues (GTEx). For each gene in the reference panel, precomputed SNP-expression weights in the 1-Mb window were obtained (Supplementary Methods 2), including the highly tissue-specific splicing QTL (sQTL) information from the MAGMA.celltyping R package, scRNA expression values were obtained from five different mouse brain regions (neocortex, hippocampus, hypothalamus, striatum, and midbrain)\(^\text{84}\). A gene-expression specificity metric for each cell-type was calculated by dividing the expression level in a given cell type by the sum of the expression levels from all cell types (i.e., genes with high expression levels in two or more cell types will get a lower specificity measure than genes with high expression levels in a single-cell type), followed by binning the metric value to 40 equally sized bins. The MAGMA one-sided test was then used to test for enrichment between the top 10 percentile bins (bins with higher cell-specificity) in each cell type. Bonferroni correction for the number of cell type tests was applied (P < 2.1 x 10^-3).

Transcriptome-wide association study and colocalization. We performed transcriptome-wide association studies (TWAS) using the association statistics from the WMH SNP-main effects (European-only) and weights from 22 publicly available gene-expression reference panels (Supplementary Methods 2) from blood (Netherlands Twin Registry, NTR; Young Finns Study, YFS), arterial (Genotype-Tissue Expression, GTEx), brain (GTEx, CommonMind Consortium, CMC) and peripheral nerve tissues (GTEx). For each gene in the reference panel, precomputed SNP-expression weights in the 1-Mb window were obtained (Supplementary Methods 2), including the highly tissue-specific splicing QTL (sQTL) information from the MAGMA.celltyping R package, scRNA expression values were obtained from five different mouse brain regions (neocortex, hippocampus, hypothalamus, striatum, and midbrain)\(^\text{84}\). A gene-expression specificity metric for each cell-type was calculated by dividing the expression level in a given cell type by the sum of the expression levels from all cell types (i.e., genes with high expression levels in two or more cell types will get a lower specificity measure than genes with high expression levels in a single-cell type), followed by binning the metric value to 40 equally sized bins. The MAGMA one-sided test was then used to test for enrichment between the top 10 percentile bins (bins with higher cell-specificity) in each cell type. Bonferroni correction for the number of cell type tests was applied (P < 2.1 x 10^-3).

References

74. Willer, J. C., Li, Y. & Abecasis, R. G. A. METAL: fast and efficient meta-
estimation of SNP and SNP x environment regression coefficients. Genet. 
76. Purcell, S. et al. PLINK: a tool set for whole-genome association and 
77. The 3C Study Group. Vascular factors and risk of dementia: design of the Three-City Study and baseline characteristics of the study population. 
Advantages and pitfalls in the application of mixed-model association 
79. International Consortium for Blood Pressure Genome-Wide Association, S. 
et al. Genetic variants in novel pathways influence blood pressure and 
80. Li, M.-X., Yeung, Y. J. M., Cherny, S. S. & Sham, C. P. Evaluating the effective 
numbers of independent tests and significant p-value thresholds in 
commercial genotyping arrays and public imputation reference datasets. Hum. 
81. Bulik-Sullivan, B. et al. An atlas of genetic correlations across human diseases and 
82. Svedlow, D. I. et al. Selecting instruments for Mendelian randomization in the 
wake of genomewide association studies. Int. J. Epidemiol. 45, 1600–1616 
(2016).
83. Bowden, J., Davey Smith, G. & Burgess, S. Mendelian randomization with 
invalid instruments: effect estimation and bias detection through Egger regression. 
84. Ziese, L. et al. Brain structure. Cell types in the mouse cortex and 
85. Ward, L. D. & Kells, M. haploReg: a resource for exploring chromatin states, 
conservation, and regulatory motif alterations within sets of genetically linked 
86. Sakaue, S. & Okada, Y. GREP: genome for REPositioning drugs. 

Acknowledgements

We thank all the participating cohorts from the CHARGE consortium and the UK 
biobank participants for contributing to this study. Detailed study-specific acknowledg-
ements are provided in the supplementary note.

Author contributions

analysis. M.S. performed shared genetic architecture and functional characterization 
analysis. M.S., C.B., and A.S. performed lifetime impact analysis. M.S. and S.Sakaue 
P.M.M., M.F., and S.D. wrote and edited the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41467-
020-19111-2.

Correspondence and requests for materials should be addressed to M.F. or S.D.

Peer review information Nature Communications thanks the anonymous reviewers for their 
contribution to the peer review of this work. Peer review reports are available.

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

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

Open Access

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

© The Authors (2020)
1062, Nutrition Obesity and Risk of Thrombosis, F-13385 Marseille, France. 125Aix-Marseille University, UMR_S 1062, Nutrition Obesity and Risk of Thrombosis, F-13385 Marseille, France. 126Group Health Research Institute, Group Health Cooperative, Seattle, WA 98101, USA. 127Einhoven Laboratory for Experimental Vascular Medicine, Department of Thrombosis and Hemostasis, Leiden University Medical Center, 2300 RC Leiden, The Netherlands. 128Division of Preventive Medicine, Brigham and Women's Hospital, Boston, MA 02215, USA. 129Division of Preventive Medicine, Brigham and Women's Hospital, Boston, MA 02215, USA. 130Department of Molecular Epidemiology, Leiden University Medical Center, 2300 RC Leiden, The Netherlands. 131Université Paris Descartes, Sorbonne Paris Cité, Paris, France. 132AP-HP, Hôpital Européen Georges Pompidou, Service Hématologie Biologique, Paris, France. 133INSERM, UMR_S 1140, Faculté de Pharmacie, Paris, France. 134Department of Epidemiology, University of Washington, Seattle, WA 98195, USA. 135Seattle Epidemiology Research and Information Center, VA Office of Research and Development, Seattle, WA 98108, USA. 136Los Angeles Biomedical Research Institute and Department of Pediatrics, Harbor-UCLA Medical Center, Torrence, CA 90502, USA.

International Headache Genomics Consortium (IHGC)


137Psychiatric and Neurodevelopmental Genetics Unit, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA. 138Medical and Population Genetics Program, Broad Institute of MIT and Harvard, Cambridge, MA, USA. 139Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge, MA, USA. 140Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, UK. 141Analytic and Translational Genetics Unit, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA. 142GORMI, Oslo University Hospital, P.O. 4956 Nydalen, 0424 Oslo, Norway. 143Department of Neurology, Oslo University Hospital, P.O. 4956 Nydalen, 0424 Oslo, Norway. 144Institute of Clinical Medicine, University of Oslo, P.O. 1171 Blindern, 0318 Oslo, Norway. 145Institute of Molecular Medicine Finland (FIMM), University of Helsinki, Helsinki, Finland. 146Estonian Genome Center, University of Tartu, Tartu, Estonia. 147Division of Endocrinology, Boston Children's Hospital, Boston, MA, USA. 148Statens Serum Institut, Dept of Epidemiology Research, Copenhagen, Denmark. 149Novo Nordisk Foundation Center for Basic Metabolic Research, University of Copenhagen, Copenhagen, Denmark. 150Illumina, 5200 Illumina Way, San Diego, CA, USA. 151Val d'Hebron Research Institute, Pediatric Neurology, Barcelona, Spain. 152Folkhålsans Institute of Genetics, FI-00290 Helsinki, Finland. 153Neuroscience Center, University of Helsinki, FI-00014 Helsinki, Finland. 154Research Programs Unit, Molecular Neurology, University of Helsinki, FI-00014 Helsinki, Finland. 15523andMe, Inc., 899 W. Evelyn Avenue, Mountain View, CA, USA. 156Inserm Research Center for Epidemiology and Biostatistics (U987), University of Bordeaux, 33076 Bordeaux, France. 157Division of Preventive Medicine, Brigham and Women's Hospital, Boston, MA 02215, USA. 158deCODE Genetics, 101 Reykjavik, Iceland. 159Medical Research Council (MRC) Integrative Epidemiology Unit, University of Bristol, Bristol, UK. 16023andMe University Amsterdam, Department of Biological Psychology, 1081 BT Amsterdam, The Netherlands. 161Leiden University Medical Centre, Department of Neurology, PO Box 9600, 2300 RC Leiden, The Netherlands. 162Department of Medicine, University of Washington, Seattle, WA 98195, USA. 163Seattle Epidemiology Research and Information Center, VA Office of Research and Development, Seattle, WA 98108, USA. 164Los Angeles Biomedical Research Institute and Department of Pediatrics, Harbor-UCLA Medical Center, Torrence, CA 90502, USA.
University of Helsinki, Helsinki, Finland. 163 Institute for Stroke and Dementia Research, Klinikum der Universität München, Ludwig-Maximilians-Universität München, Feodor-Lynen-Str. 17, 81377 Munich, Germany. 164 Department of Neurology and Epileptology, Hertie Institute for Clinical Brain Research, University of Tuebingen, Tübingen, Germany. 165 Karolinska Institutet, Department of Neuroscience, 171 77 Stockholm, Sweden. 166 Department of Genetics and Computational Biology, QIMR Berghofer Medical Research Institute, 300 Herston Road, Brisbane, QLD 4006, Australia. 167 Ulm University, Institute of Human Genetics, 89081 Ulm, Germany. 168 University of Oulu, Center for Life Course Epidemiology and Systems Medicine, Oulu, Finland. 169 Department of Twin Research and Genetic Epidemiology, King’s College London, London, UK. 170 Department of Epidemiology, Erasmus University Medical Center, 3015 CN Rotterdam, The Netherlands. 171 Department of Radiology, Erasmus University Medical Center, 3015 CN Rotterdam, The Netherlands. 172 Department of Clinical Chemistry, Finlab Laboratories, and School of Medicine, University of Tampere, Tampere, Finland 33520. 173 Department of Public Health, University of Helsinki, Helsinki, Finland. 174 Harvard Medical School, Boston, MA 02115, USA. 175 University Duisburg Essen, Essen, Germany. 176 Landspitali University Hospital, 101 Reykjavik, Iceland. 177 VU University Medical Center, Department of Psychiatry, 1081 HL Amsterdam, The Netherlands. 178 Department of Psychiatry, Washington University School of Medicine, 660 South Euclid, CB 8134, St. Louis, MO 63110, USA. 179 University Medical Center Groningen, University of Groningen, 9700RB Groningen, The Netherlands. 180 MRC Functional Genomics Unit, Department of Physiology, Anatomy & Genetics, Oxford University, Oxford, UK. 181 Department of Radiology, Erasmus University Medical Center, 3015 CN Rotterdam, The Netherlands. 182 Department of Pain Management and Research, Oslo University Hospital, Oslo, 0424 Oslo, Norway. 183 Medical Faculty, University of Oslo, Oslo, 0318 Oslo, Norway. 184 Division of Mental Health, Norwegian Institute of Public Health, P.O. Box 4404 Nydalen, Oslo NO-0403, Norway. 185 Kiel Pain and Headache Center, 24149 Kiel, Germany. 186 Danish Headache Center, Department of Neurology, Rigshospitalet, Glostrup Hospital, University of Copenhagen, Copenhagen, Denmark. 187 Institute of Biological Psychiatry, Mental Health Center Sct. Hans, University of Copenhagen, Roskilde, Denmark. 188 Institute Of Biological Psychiatry, MHC Sct. Hans, Mental Health Services Copenhagen, DK-2100 Copenhagen, Denmark. 189 Institute of Clinical Sciences, Faculty of Medicine and Health Sciences, University of Copenhagen, DK-2100 Copenhagen, Denmark. 190 NPSYCH—The Lundbeck Foundation’s Initiative for Integrative Psychiatric Research, DK-2100 Copenhagen, Denmark. 191 Department of Health, National Institute for Health and Welfare, Helsinki, Finland. 192 Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, Turku 20521, Finland. 193 Department of Clinical Physiology and Nuclear Medicine, Turku University Hospital, Turku 20521, Finland. 194 Department of Neurology, Erasmus University Medical Center, 3015 CN Rotterdam, The Netherlands. 195 Imperial College London, Department of Epidemiology and Biostatistics, MRC Health Protection Agency (HPE) Centre for Environment and Health, School of Public Health, London, UK W2 1PG. 196 University of Oulu, Biocenter Oulu, Oulu, Finland. 197 Box 5000, Fin-90014 University of Oulu, Oulu, Finland. 198 Oulu University Hospital, Unit of Primary Care, OuluBox 10, Fin-90029 OYSFinland. 199 University Medical Center Hamburg Eppendorf, Institute of Human Genetics, 20246 Hamburg, Germany. 200 Population Health Research Institute, St George’s, University of London, Cranmer Terrace, London SW17 0RE, UK. 201 Munich Cluster for Systems Neuroscience (SyNergy), Munich, Germany. 202 Leiden University Medical Centre, Department of Human Genetics, PO Box 9600, 2300 RC Leiden, The Netherlands. 203 Department of Medicine, University of Iceland, 101 Reykjavik, Iceland. 204 Statistical and Genomic Epidemiology Laboratory, Institute of Health and Biomedical Innovation, Queensland University of Technology, 60 Musk Ave, Kelvin Grove, QLD 4059, Australia. 205 Department of Neurology, Massachusetts General Hospital, Boston, MA, USA.