bioRxiv preprint doi: https://doi.org/10.1101/683367; this version posted June 27, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Common genetic variation indicates separate etiologies for periventricular and deep white matter hyperintensities

Nicola J Armstrong[1#], Karen A Mather[2,3#], Muralidharan Sargurupremraj[4], Maria J Knol[5], Rainer Malik[6], Claudia L Satizabal[7,8,9], Lisa R Yanek[10], Wen Wei[2], Vilmundur Gudnason[11,12], Nicole D Deuker[13], Lloyd T Elliott[14,15], Edith Hofer[16,17], Neda Jahanshad[18], Shuo Li[19], Mark A Logue[19,20,21], Michelle Luciano[22], Markus Scholz[23,24], Albert Smith[12], Stella S Trompet[25,26], Dina Vojinovic[5], Rui Xia[27], Fidel Alfaro-Almagro[15], David Ames[28,29], Najaf Amin[5], Philippe Amouyel[30], Alexa S Beiser[8,9,19], Henry Brodaty[2,31], Ian J Deary[22], Christine Fennema-Notestine[32,33], Piyush G Gampwar[34], Rebecca Gottesman[35], Ludovica Griffanti[15], Clifford R Jack[36], Mark Jenkinson[15], Jiyang Jain[2], Brian G Kral[10], John W Kwok[37,38], Leonie Lampe[39], David CM Liewald[22], Pauline Maillard[40], Jonathan Marchini[41], Mark E Bastin[42], Bernard Mazoyer[43], Lukas Pirpamer[44], José Rafael Romero[8,9], Gennady V Roshchupkin[5,45], Peter R Schofield[3], Matthias L Schroeter[46,47], David J Stott[48], Anbupalam Thalamuth[2], Julian Trollor[2,49], Christophe Tzourio[4,50], Jeroen van der Grond[51], Meike W Vernooij[5,45], Veronica A Witte[39,52], Maragret J Wright[53,54], Qiong Yang[19], Moris Zoe[54], Siggi Siggurdsson[11], Arno Villringer[46,47], Helena Schmidt[34], Asta L Haberg[55,56], Cornelia M Van Duijn[5,57], J Wouter Jukema[58,59], Martin Dichigans[6,60,61], Ralph L Sacco[62,63,64,65], Clinton B Wright[66], William S Kremen[67,68], Lewis C Becker[10], Paul M Thompson[69], Lenore Launer[70], Thomas H Mosley[71], Joanna M Wardlaw[72], M Afran Ikram[5], Hieab HH Adams[5,45,73,], Reinhold Schmidt[16], Stephen M Smith[74], Charles Decarli[75], Perminder S Sachdev[2,76], Myriam Fornage[27,77], Stephanie Debbette[4,78], Sudha Seshadri[7,8,9], Paul A Nyquist[10,79,80*]

*Indicates the corresponding author "Paul A Nyquist"

#These authors have shared first authorship "Nicola Armstrong, Karen Mather"

bioRxiv preprint doi: https://doi.org/10.1101/683367; this version posted June 27, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

[1]Mathematics and Statistics, Murdoch University, Perth, Australia. [2]Centre for Healthy Brain Ageing, School of Psychiatry, University of New South Wales, Sydney, Australia. [3]Neurosciences Research, School of Medical Sciences, University of New South Wales, Sydney Australia. [4]Univ. Bordeaux, Inserm, Bordeaux Population Health Research Center, Bordeaux, France. [5]Department of Epidemiology, Erasmus MC, University Medical Center, Rotterdam, the Netherlands. [6]Institute for Stroke and Dementia Research (ISD), University Hospital, Ludwig-Maximilians-Universität LMU Munich, Germany. [7]Glenn Biggs Institute for Alzheimer's & Neurodegenerative Diseases, UT Health San Antonio, San Antonio, TX, USA. [8]The Framingham Heart Study, Framingham, MA, USA. [9]Department of Neurology, Boston University School of Medicine, Boston, MA, USA. [10]GeneSTAR Research Program, Johns Hopkins University School of Medicine, Baltimore, MD, USA. [11]Icelandic Heart Association, Kopavogur, Iceland. [12]University of Iceland, Reykjavik, Iceland. [13]John P. Hussman Institute for Human Genomics, University of Miami, Miami, FL, USA. [14]Department of Statistics and Actuarial Science, Simon Fraser University, Burnaby, BC, Canada. [15]Wellcome Centre for Integrative Neuroimaging (WIN FMRIB), University of Oxford, Oxford, United Kingdom. [16]Clinical Division of Neurogeriatrics, Department of Neurology, Medical University of Graz, Austria. [17]Institute for Medical Informatics, Statistics and Documentation, Medical University of Graz, Austria. [18]Mark and Mary Neuroimaging and Informatics Institute, Keck School of Medicine of USC, University of Southern California, Marina del Rey, CA, USA. [19]Department of Biostatistics, Boston University School of Public Health, Boston, MA, USA. [20]National Center for PTSD: Behavioral Science Division, VA Boston Healthcare System, Boston, MA, USA. [21]Department of Psychiatry and Biomedical Genetics Section, Boston University School of Medicine, Boston, MA, USA. [22]Centre for Cognitive Ageing and Cognitive Epidemiology, Psychology, University of Edinburgh, Edinburgh, UK. [23]Institute for Medical Informatics, Statistics and Epidemiology, University of Leipzig, Leipzig, Germany. [24]LIFE Research Center for Civilization Disease, Leipzig, Germany. [25]Department of Internal Medicine, section of Gerontology and Geriatrics, Leiden University Medical

bioRxiv preprint doi: https://doi.org/10.1101/683367; this version posted June 27, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Center, The Netherlands. [26]Department of Cardiology, Leiden University Medical Center, The Netherlands. [27]Brown Foundation Institute of Molecular Medicine, McGovern Medical School, University of Texas Health Science Center at Houston, Houston, TX, USA. [28]National Ageing Research Institute, Parkville, Victoria, Australia. [29]Academic Unit for Psychiatry of Old Age, University of Melbourne, St George's Hospital, Kew, Australia. [30]Lille University, Inserm, Institut Pasteur de Lille, RID-AGE - Risk factors and molecular determinants of aging-related diseases and Labex Distalz, Lille, France. [31]Dementia Centre for Research Collaboration, University of New South Wales, Sydney, Australia. [32]Department of Psychiatry, University of California, San Diego, La Jolla, CA, USA. [33]Center for Behavior Genetics of Aging, University of California, San Diego, La Jolla, CA, US. [34]Gottfried Schatz Research Center (for Cell Signaling, Metabolism and Aging), Medical University of Graz, Austria. [35]Department of Neurology, Cerebrovascular and stroke Division, Johns Hopkins University School of Medicine, Baltimore, MD, USA. [36]Department of Radiology, Mayo Clinic, Rochester, MN, USA. [37]Brain and Mind Centre - The University of Sydney, Camperdown, NSW, Australia. [38]School of Medical Sciences, University of New South Wales, Sydney, Australia. [39]Department of Neurology, Max-Planck-Institute for Human Cognitive and Brain Sciences, Leipzig, Germany. [40]Imaging of Dementia and Aging (IDeA) Laboratory, Department of Neurology, University of California-Davis, Davis, CA, USA. [41]Statistical Genetics and Methods at Regeneron Pharmaceuticals, Inc., New York, NY, USA. [42]Centre for Clinical Brain Sciences, University of Edinburgh & Brain Research Imaging Centre, University of Edinburgh, Edinburgh, United Kingdom. [43]Univ. Bordeaux, Institut des Maladies Neurodégénératives, Bordeaux, France. [44]Clinical Division of Neurogeriatrics, Department of Neurology, Medical University of Graz, Austria. [45]Department of Radiology and Nuclear Medicine, Erasmus MC, Rotterdam, the Netherlands. [46]Department of Neurology, Max-Planck-Institute for Human Cognitive and Brain Sciences, Leipzig, Germany. [47]Day Clinic for Cognitive Neurology, University Hospital Leipzig, Leipzig, Germany. [48]Institute of Cardiovascular and Medical Sciences, College of Medical, Veterinary and Life Sciences, University of Glasgow, United Kingdom. [49]Department of Developmental Disability Neuropsychiatry, School of Psychiatry, University

bioRxiv preprint doi: https://doi.org/10.1101/683367; this version posted June 27, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

of New South Wales, Sydney, Australia. [50]CHU de Bordeaux, Public Health Department, Medical information Department, Bordeaux, France. [51]Department of Radiology, Leiden University Medical Center, The Netherlands. [52]Collaborative Research Center 1052 Obesity Mechanisms, Faculty of Medicine, University of Leipzig, Leipzig, Germany. [53]Queensland Brain Institute, The University of Queensland, St Lucia, QLD, Australia. [54]Centre for Advanced Imaging, The University of Queensland, St Lucia, QLD, Australia. [55]Department of Neuromedicine and Movement Science, Norwegian University of Science and Technology, Trondheim, Norway. [56]Department of Radiology and Nuclear Medicine, Norwegian University of Science and Technology, Trondheim, Norway. [57]Nuffield Department of Population Health, University of Oxford, Oxford, United Kingdom. [58]Department of Cardiology, Leiden University Medical Center, The Netherlands. [59]Einthoven Laboratory for Experimental Vascular Medicine, LUMC, Leiden, the Netherlands. [60]German Center for Neurodegenerative Diseases, Munich, Germany. [61]Munich Cluster for Systems Neurology (SyNergy), Munich, Germany. [62]Dr. John T. Macdonald Foundation Department of Human Genetics, University of Miami, Miami, FL, USA. [63]Department of Public Health Sciences, Miller School of Medicine, University of Miami, Miami, FL USA. [64]Department of Neurology, Miller School of Medicine, University of Miami, Miami, FL USA. [65]Evelyn F. McKnight Brain Institute, Department of Neurology, University of Miami, Miami FL USA. [66]National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, MD, USA. [67]Department of Psychiatry, University of California, San Diego, La Jolla, CA, USA. [68]Center for Behavior Genetics of Aging, University of California, San Diego, La Jolla, CA, USA. [69]Mark and Mary Neuroimaging and Informatics Institute, Keck School of Medicine of USC, University of Southern California, Marina del Rey, CA, USA. [70]Laboratory of Epidemiology and Population Sciences, National Institute on Aging, Intramural Research Program, National Institutes of Health, Bethesda, MD, USA. [71]Departemnt of Geriatric Medicine, Memory Impairment and Neurodegenerative Dementia (MIND) Center, University of Mississippi Medical Center, Jackson, MS. [72]Division of Neuroimaging Sciences, Brain Research Imaging Centre, University of Edinburgh. Edinburgh, UK. [73]Department of Clinical Genetics, Erasmus MC, Rotterdam, the Netherlands. [74]Wellcome Centre for Integrative Neuroimaging

bioRxiv preprint doi: https://doi.org/10.1101/683367; this version posted June 27, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

(WIN FMRIB), University of Oxford, Oxford United Kingdom. [75]Alzheimer's Disease Center and Imaging of Dementia and Aging (IDeA) Laboratory, Department of Neurology and Center for Neuroscience University of California at Davis, Davis, CA USA. [76] Neuropsychiatric Institute, Prince of Wales Hospital, Sydney, Australia. [77]Human Genetics Center, School of Public Health UT, Houston TX, USA. [78]CHU de Bordeaux (University Hospital), Department of Neurology, Bordeaux, France. [79]Departments of Neurology, Critical Care Medicine, Neurosurgery, Johns Hopkins University School of Medicine, Baltimore, MD, USA. [80]Critical Care Medicine Department, National Institutes of Health, Bethesda MD, USA.

bioRxiv preprint doi: https://doi.org/10.1101/683367; this version posted June 27, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Abstract

We conducted a genome-wide association meta-analysis of two ischemic white matter disease subtypes in the brain, periventricular and deep white matter hyperintensities (PVWMH and DWMH). In 26,654 participants, we found 10 independent genome-wide significant loci only associated with PVWMH, four of which have not been described previously for total WMH burden (16q24.2, 17q21.31, 10q23.1, 7q36.1). Additionally, in both PVWMH and DWMH we observed the previous association of the 17q25.1 locus with total WMH. We found that both phenotypes have shared but also distinct genetic architectures, consistent with both different underlying and related pathophysiology. PVWMH had more extensive genetic overlap with small vessel ischemic stroke, and unique associations with several loci implicated in ischemic stroke. DWMH were characterized by associations with loci previously implicated in vascular as well as astrocytic and neuronal function. Our study confirms the utility of these phenotypes and identifies new candidate genes associated only with PVWMH.

bioRxiv preprint doi: https://doi.org/10.1101/683367; this version posted June 27, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Introduction

Radiological white matter hyperintensities of presumed ischemic origin (WMH, white matter lesions, leukoaraiosis) are the most prevalent sign of cerebral small vessel disease (SVD) and represent 40% of all SVD disease burden[1]. WMH begin to be observed in middle age, are present in approximately 50% of adults by their mid to late 40's and are commonly found in community-dwelling older adults[2]. They are detected as incidental lesions on T2-weighted MRI[1]. They are one of many radiological phenotypes associated with SVD including: microhemorrhages, enlarged perivascular spaces, gliosis, and brain atrophy[1]. WMH are associated with increased risk for ischemic and hemorrhagic stroke, cognitive decline that may result in dementia, and motor gait disorders and reflect ongoing chronic subclinical brain injury[3-7]. They are also commonly related to histopathological changes caused by hypertension and other risk factors such as microvascular occlusion, lipohyalinosis, gliosis, and fibrosis of surrounding adventitia[6,8,9]. These lesion types are important clinical targets for intervention in prevention of cognitive aging, stroke, and dementia[1,10-12]. Total WMH burden has high heritability and several genome-wide association studies (GWAS) have been previously reported[13-19].

Two important and common regional classifications, based on their anatomical relationship to the lateral ventricles in the brain, are periventricular white matter hyperintensities (PVWMH) and deep white matter hyperintensities (DWMH)[6,8,20-22]. This categorization reflects proposed differences in underlying pathophysiology, histopathology, and etiology for each lesion type[6,8,20,21]. Both lesion types are hypothesized to reflect chronic microvascular injury from hypoperfusion and/or microvascular occlusion[1,6,8]. The microvasculature subserving PVWMH

bioRxiv preprint doi: https://doi.org/10.1101/683367; this version posted June 27, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

and DWMH differ based on the type of microvessels that supply these brain regions. DWMH lesions occur in the subcortex, areas primarily supplied by long microvessels, with lower estimated blood pressures thought subject to hypoperfusion[1,8,23,24]. In contrast, PVWMH lesions are related to alterations in short penetrating microvessels ending in close approximation to larger arterial blood vessels with different vascular architecture such as two leptomeningeal layers and enlarged perivascular spaces[1,21,25]. These microvessels are more directly perfused by large systemic arterial vessels and have much higher estimated mean arterial pressures than those supplying DWMH. These vessels are hypothesized to be affected more directly by hypertension and other risk factors associated with stroke[1,8,23,24].

PVWMH and DWMH are associated with different clinical phenotypes. PVWMH are associated with cognitive performance and increased systolic and arterial pressure, while DWMH are linked to BMI, mood disorders, gait impairment and arterial hypertension [12,26,27]. These sub-classifications may reflect differences in associated underlying genetic factors[13]. Twin and family studies report that both PVWMH and DWMH have high heritability[13,14]. Genetic correlations between the two phenotypes are high[13,28]. Recently, GWAS for total WMH volume identified a major genetic risk locus on chromosome 17q25.1[15,16,18,19]. Several other genome-wide significant loci were identified (e.g., 10q24, 2p21, 2q33, 6q25.1)[16,19,29]. However, the genetic determinants of regional WMH burden, specifically DWMH and PVWMH, remain elusive.

Here we combined all available participants aged 45 and above with both DWMH and PVWMH measurements from the Cohorts for Heart and Aging Research in Genomic Epidemiology

bioRxiv preprint doi: https://doi.org/10.1101/683367; this version posted June 27, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

(CHARGE) and the Enhancing Neuro-Imaging Genetics through Meta-Analysis (ENIGMA) consortia, and the UK Biobank (UKB). The combined sample size of this study was n=26,654 and is the only GWAS to date of these subclassifications of WMH. We hypothesized that separating the two WMH subclassifications would mitigate phenotype heterogeneity, allowing us to identify additional risk loci and show that DWMH and PVWMH have different genetic underpinnings. We further estimated the genetic correlations of DWMH and PVWMH with stroke, dementia, and vascular disease, hypothesizing that there would be differences between the two subclassifications.

bioRxiv preprint doi: https://doi.org/10.1101/683367; this version posted June 27, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Methods

Study Cohorts

Study participants (total N~26,654) were drawn from studies of the CHARGE and ENIGMA consortia and UK Biobank (UKB). All participants were aged 45 years and over and individuals with stroke, dementia or any neurological abnormality such as tumors were excluded. Descriptions of the individual studies are found in the Supplementary Note. The demographic details of the included participants from each study are presented in Supplementary Table 1. All participants provided written informed consent and each study received ethical approval to undertake this work.

Phenotype and covariates

The majority of studies (n=14) extracted deep and periventricular WMH volumetric data using automated methods while three studies (HUNT, LBC, AGES, n= 3843) used visual rating scales. Volumetric measures were transformed by adding 1 and taking the natural logarithm. The UKB processed their volumetric WMH data as per Elliott et al[30]. The MRI and WMH extraction methods for each study are detailed in Supplementary Table 2. For all participants from CHARGE and ENIGMA, hypertension was defined as systolic blood pressure ≥140mm Hg and diastolic blood pressure ≥90mm Hg or on current antihypertensive treatment and coded as yes (1) versus no (0). For volumetric measures, intracranial volume (ICV) was included as a covariate unless ICV had already been adjusted for when the WMH measures were calculated.

Genotyping and Imputation

bioRxiv preprint doi: https://doi.org/10.1101/683367; this version posted June 27, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Genotyping platforms, imputation and quality control procedures used for each participating study are described in Supplementary Table 3. For the majority of studies, genotype data were imputed using the 1000 Genomes Project reference panel (phase 1, version 3) using standard software and methods. One cohort, VETSA, used the phase 3, version 5 panel. The UKB details and procedures are detailed in Elliott et al[30].

Statistics

All statistical tests were two-sided and, unless stated otherwise, the conventional *p*-value of 0.05 was used for significance. When necessary, Bonferroni correction for multiple testing was applied.

Genome-wide association analysis

Each study fitted linear regression models to test the association of DWMH and PVWMH with individual SNPs. Additive genetic effects were assumed and the models were adjusted for age (years), sex and ICV (where applicable). In addition, principal components for population stratification and other covariates, such as familial structure, were included if necessary. Models were also fitted with hypertension as an additional covariate. Details of the covariates included for each study are listed in Supplementary Table 3.

Prior to meta-analysis, quality control of summary statistics for each cohort was performed using EasyQC[31]. Marker names, alleles and chromosome base pair positions were harmonized to the appropriate 1000 Genome reference panel. Meta-analyses were restricted to biallelic SNPs. At the individual study level, SNPs were removed due to duplication, missing

bioRxiv preprint doi: https://doi.org/10.1101/683367; this version posted June 27, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

information, allelic mismatch to the reference panel or allele frequency outliers (where the study allele frequency differed from the reference frequency by more than 0.2). The effective allele count was defined as twice the product of the MAF, imputation accuracy (r^2, info score or oevar_imp), and the study sample size. Variants with an effective allele count < 10 were excluded. QQ plots, allele-frequency plots and P-Z plots were constructed for each study and visually checked pre and post QC.

Fixed-effects, inverse-variance-weighted meta-analysis was carried out in METAL[32], with correction for genomic control. Two meta-analyses were carried out: all cohorts excluding UKB (discovery, phase I) and all cohorts (phase II). Post meta-analysis QC, filtering SNPs by their heterozygosity index (I^2>0.8) and sample size (phase I: N> 11,000; phase II: N>14,000) was also performed.

SNPs were denoted as genome-wide significant if p<5e-8, and suggestive if p<1e-5. Replication of the phase I genome-wide significant SNPs was performed in the UKB. The *p*-value for replication was set to the nominal significance level (0.05) divided by total number of genome-wide significant SNPs (DWMH *p*< 3.60e-4; PVWMH, *p*<2.75e-4). Manhattan and Q-Q plots for the meta-analysis were created in FUMA[33]. Regional association plots were generated with LocusZoom[34]. Annotation of genome-wide significant SNPs was performed using FUMA[35]. Further *in silico* functional annotation was undertaken for the top SNPs using publicly available databases (SNiPA (http://snipa.helmholtz-muenchen.de/snipa/), SNPnexus (https://www.snp-nexus.org/), and GTEx (http:// www.gtexportal.org/home/))[35-37].

bioRxiv preprint doi: https://doi.org/10.1101/683367; this version posted June 27, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Gene-based and pathway analysis

Gene-based tests and gene set (pathway) analyses were performed using MAGMA as implemented in FUMA[33,38]. For gene-based tests, significance was set to 2.667e-6 (0.05 divided by the number of protein coding genes (18745)). For biological pathways, 10655 gene sets (4738 curated gene sets and 5917 GO terms from MsigDB v6.1) were tested, resulting in a significance level of 4.69e-6.

Heritability

In the Rotterdam study, SNP-based heritability was computed with GCTA[39] adjusting for age and sex. Additionally, univariate genome-wide heritability estimates were calculated using the phase II GWAS summary statistics, for both DWMH and PVWMH, using LD score regression (LDSC)[40] and the European LD-score files calculated from the 1000G reference panel provided by the developers.

Genetic correlations with stroke and dementia

Cross-trait genetic correlation between the two sub-classifications of WMH (deep and periventricular), stroke (overall and ischemic subtypes) and dementia were estimated using LDSC [40] on the GWAS summary statistics from phase II, MEGASTROKE (European ancestry-only)[41 42,43]. LD scores were based on the HapMap3 European reference panel. The significance threshold was set to 0.0026 (0.05/19 tests).

Regional level correlation was investigated using the GWAS-pairwise (GWAS-PW) method [44]. GWAS-PW is a Bayesian approach that estimates the probability of a shared locus harboring genetic variant that influences only the

bioRxiv preprint doi: https://doi.org/10.1101/683367; this version posted June 27, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

first trait (model 1), only the second trait (model 2) or both traits (model 3). The posterior probability of association (PPA) for model 3 greater than 0.90 was considered significant. Additionally, fgwas v.0.3.6[45] with the default settings was run to determine the PPA for each of the traits separately and the overall correlation in Z-scores estimated from genomic regions with PPA less than 0.20 was used to correct for possible sample overlap between the traits.

bioRxiv preprint doi: https://doi.org/10.1101/683367; this version posted June 27, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Results

Demographic characteristics

Study descriptions are provided in the Supplementary Information and Supplementary Table 1. The discovery cohort was comprised of dementia and stroke free older adults (≥45 years, n ~18,234) from 16 international studies. The majority of the sample had DWMH and PVWMH data extracted by automated or semi-automated volumetric methods (n = 14,383), while the remainder used visual ratings (Supplementary Table 2). The discovery cohort was primarily Caucasian but included a smaller subsample of African Americans (n=736) and Hispanics (n= 658). The replication cohort consisted of predominantly Caucasian participants from the UK Biobank (n=8,428), all of whom had DWMH and PVWMH extracted using automated methods. Genotyping, imputation and analysis details for each study are provided in Supplementary Table 3.

Discovery and Replication

The discovery association meta-analysis (Phase I) found genome-wide significant associations in the 17q25.1 region for both DWMH (n=167 SNPs) and PVWMH (n=199 SNPs; Supplementary Tables 4 and 5). This region has previously been associated with total WMH burden[15,16,19]. In contrast, only the PVWMH analysis found additional genome-wide significant associations on chromosomes 2 (3 *NBEAL* intronic SNPs) and 10 (2 regions, [i] a previously undescribed locus at 10q23.1 - 14 *TSPAN14/FAM231A* intronic & intergenic SNPs; [ii] locus at 10q24.33 - one *SH3PXD2A* intronic SNP). Two of these regions have also previously been described for total WMH burden (chr 2, *NBEAL*[16,19], 10q24.33,

bioRxiv preprint doi: https://doi.org/10.1101/683367; this version posted June 27, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

SH3PXD2A[19]). We repeated our analysis adjusting for hypertension, a known risk factor for WMH and cerebral SVD; this adjustment made almost no difference to our findings (Supplementary Tables 8 and 9). Replication of the majority of genome-wide significant results using participants from the UKB for both DWMH and PVWMH was observed after Bonferroni adjustment for multiple testing (DWMH p <3.6e-4, PVWMH p < 2.76e-4, Supplementary Tables 6 and 7).

Overall Meta-analysis

Given the relatively large size of the replication cohort, a combined meta-analysis (Phase II) was undertaken using all of the available samples from the CHARGE, ENIGMA and UK Biobank cohorts (n~26,654). Removing either the small subsample of non-Caucasians, or the cohorts with visual ratings, did not substantially change the findings for either phenotype (beta value $r^2 > 0.93$).

Results of the Phase II GWAS meta-analysis for DWMH and PVWMH are shown in Table 1, Figure 1 and Supplementary Tables 10 and 11. We identified 513 genome-wide significant SNPs for PVWMH, and 236 for DWMH. In total, 209 SNPs were found to be significant for both phenotypes, all of which are located in 17q25.1 (Figure 2a).

New genome-wide significant PVWMH loci

Ten chromosomal regions containing 290 genome-wide significant SNPs for PVWMH only were identified on chromosomes 2 (3 regions), 6, 7, 10 (2 regions), 13, 16 and 17q23.1 (Supplementary Table 11; Supplementary Figure 1). Four loci had not been previously

bioRxiv preprint doi: https://doi.org/10.1101/683367; this version posted June 27, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

reported for associations with total WMH at the genome-wide significant level. These new loci included a novel region (7.2kb) containing 2 genome-wide significant exonic SNPs in the *NOS3* gene on 7q36.1, with the top SNP (rs1799983) resulting in an amino-acid change, which is predicted to be tolerated. A new region on chr 10 was identified (10q23.1, 50.5kb, 4 intronic SNPs in *TSPAN14* & *FAM231A*) as well as one on 16q24.2 (1.2kb) which contained two genome-wide significant intergenic SNPs. An additional novel region to the 17q25.1 locus was also observed for PVWMH only. This 27.2kb region at 17q21.31 contained 8 genome-wide significant SNPs, most of which were intronic and found in the *NMT1* gene. Many of these are eQTLs or participate in chromatin interactions (Figure 2d).

Regions previously linked to total WMH burden were significant only for PVWMH in our analysis, which included three separate genomic loci on chromosome 2. The first region (2p16.1) spanned 27.3 kb and contained 6 genome-wide significant SNPs, with the top SNP (rs7596872, intronic, *EFEMP1*) having been reported previously[29]. The second region (2q33.2) spanned 442.3kb containing 250 significant SNPs, with the top SNP being rs72934583, located in an intron of *NBEAL1*. This SNP is an eQTL for several genes (*CARF*, *ICA1L*, *NBEAL1*, *FAM117B*) in various tissues. The third region (2p21) contained 7 SNPs, spanning 7.2 kb, with the top SNP, rs57242328, being intergenic and previously reported as an eQTL for *OXER1* (tibial nerve). Many SNPs in this region were suggestive ($p<1e-5$, n=215, Supplementary Table 11). Of the genome-wide SNPs, twelve were located in two regions (2p16.1, 2q33.2) previously associated with total WMH burden (rs78857879, rs7596872 – both *EFEMP1* intronic; rs72934505- *NBEAL* intronic), coronary heart disease, myocardial infarction, migraine and cholesterol (Supplementary Table 13). The third locus at

bioRxiv preprint doi: https://doi.org/10.1101/683367; this version posted June 27, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

2p21 has also previously been significantly associated with total WMH burden (rs11679640), although this SNP was in low LD (R^2=0.232) with our top SNP in this region (rs57242328)[19].

The region on 6q25.1 contained six genome-wide significant SNPs spanning a region of 7.7 kb. The top SNP, rs275350, is an intronic SNP in *PLEKHG1*, which has been reported as an eQTL for *PLEKHG1* (testis) and recently associated with total WMH[29]. A region on 10q24.33 (4.3kb, 2 intronic SNPs in *SH3PXD2A*) was previously reported by Verhaaren et al[19], who observed 4 genome-wide significant intronic SNPs in the *NEURL*, *PDCD11* and *SH3PXD2A* genes. Additionally, two genome-wide SNPs were located in the *COL4A2* gene (intronic) on 13q34 (spanning 2.6kb); all of these SNPs are in high LD (R^2>0.8). In prior GWAS these SNPs have been associated with coronary artery disease (rs11838776) and myocardial infarction (rs55940034) (Supplementary Table 13).

Genome-wide significant locus for DWMH and PVWMH

As expected, the association of the 17q25.1 locus with both phenotypes was confirmed. The size of this region including genome-wide significant SNPs only was similar for both DWMH (236 SNPs, BP 73757836-74025656, Figure 1b) and PVWMH (223 SNPs, BP 73757836-74024711, Supplementary Figure 1a). The top genome-wide significant results in this locus were rs3744020 for DWMH, (*p*=7.06e-35, intronic SNP, *TRIM47*) and rs35392904 for PVWMH (*p*=3.989e-28, intronic SNP, *TRIM65*), which are in high linkage disequilibrium (LD, R^2=0.902) (Table 1). The top SNP for DWMH, rs3744020 was also reported genome-wide significant in the most recent total WMH GWAS[29].Both SNPs (rs3744020, rs35392904) have been reported as eQTLs for several genes, including *TRIM47* and *TRIM65* in a range of

bioRxiv preprint doi: https://doi.org/10.1101/683367; this version posted June 27, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

tissues, including the cerebellum and cortex (GTEx[36]). Of the genome-wide significant SNPs, eight have previously been identified in GWAS for total WMH burden, psoriasis, systolic blood pressure or eosinophil and basophil cell counts (Supplementary Tables 12 and 13). Many of these SNPs are eQTLs or have chromatin interactions (Figure 2c-d).

Gene-based association analysis

Using FUMA[33], 13 genes reached genome-wide significance (p<2.66e-6) for association with both DWMH and PVWMH (Table 2, Supplementary Tables 14-15, Figure 2b). All of these genes are located at 17q25.1. For PVWMH, an additional 19 genes (Table 2, Supplementary Table 15) reached genome-wide significance, covering the majority of regions/loci found in the SNP-based analysis (apart from 2p16.1 and 2p21). Four genes were located in previously unidentified regions: *CALCRL* (2q32.1), *KLHL24* (3q27.1), *VCAN* (5q27.1) and *POLR2F* (22q13.1).

Pathway Analysis

We tested enrichment of MAGMA-derived gene-based *p*-values from the phase II analysis, in gene sets from MsigDBv6.1, as implemented in FUMA[33]. No significant associations with PVWMH were found, and only one gene set (*caffarel_response_to_thc_8hr_5_dn*) remained significant after multiple testing correction for DWMH.

Heritability

Applying LD score regression[40] to the phase II GWAS summary statistics, heritability was estimated to be 0.1 for DWMH and 0.18 for PVWMH. For the Rotterdam study, the largest

bioRxiv preprint doi: https://doi.org/10.1101/683367; this version posted June 27, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

individual cohort in Phase I, we used GCTA[39] on the individual level data, adjusting for age and sex. We found heritability estimates of 0.29 and 0.47 for DWMH and PVWMH respectively.

Genetic correlation with other traits

The genetic correlation between DWMH and PVWMH, estimated using the phase II GWAS results, was high (r_g= 0.927), indicating a shared genetic architecture (Supplementary Table 16). Since higher WMH burden is associated with the presence of stroke and dementia, we also examined their genetic correlation with DWMH and PVWMH (Figure 3). Positive genetic correlations with both phenotypes were found for 'all stroke', ischemic stroke and SVD. Intracerebral haemorrhage (ICH, all types) was correlated with DWMH only. Deep ICH was correlated with both DWMH and PVWMH, but the relationship was much stronger for DWMH. No significant correlation was found with dementia.

Using GWAS-PW[46] we investigated if there were any genomic regions with shared genetic influence between PVWMH and DWMH. We identified several regions, including the major locus on chromosome 17 as well as chromosome 2 (two regions), 6 and 16 (Supplementary Table 17). These regions of shared influence overlap several genome-wide significant loci that were identified for PVWMH. However, using an alternative method[47] to calculate local shared influence, only one region was significant, and was located within the 17q25.1 region (p=4.6e-6) (Supplementary Table 17).

bioRxiv preprint doi: https://doi.org/10.1101/683367; this version posted June 27, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Finally, we investigated local regions of shared influence between the WMH subtypes and stroke (Supplementary Table 17). A region on chromosome 7 (BP 149861853-150709571, *GIMAP4/KCNH2*) exhibited shared genetic influence of 'all stroke' with both DWMH and PVWMH. This region encompasses the PVWMH genome-wide significant *NOS3* exonic SNP. Other regions of shared influence with all stroke were observed for PVWMH only. For the sub-types of stroke (i.e. ischemic, cardioembolic and SVD), significant regions were identified for DWMH and PVWMH, but none were found for both phenotypes except the chromosome 7 region for ischemic stroke (also identified for all stroke). Additional regions of shared genetic influence were observed for PVWMH with ischemic stroke, cardioembolic and SVD. In contrast, DWMH shared a region of genetic influence with SVD stroke only (chr 6: BP 202452-1452004, *FOXF2/FOXQ1*). Using HESS[47], no regions were found to have significant correlation between DWMH and any of the stroke subtypes. However, PVWMH had suggestive regions of genetic correlation with cardioembolic stroke (chr7, unadjusted p=0.03) and SVD (chr13, unadjusted p=0.02, *COL4A1/COL4A2*; chr2, unadjusted p=0.02, *NBEAL1*).

bioRxiv preprint doi: https://doi.org/10.1101/683367; this version posted June 27, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Discussion

In our analysis using all available individuals (n=26,654, Phase II analysis), we found common and different results for DWMH and PVWMH burden. Newly discovered significant genes for PVWMH include *CALCRL*, *VCAN* and *NOS3*. The only significant locus (17q25.1) for DWMH was also found for PVWMH, which was the previously reported major locus for total WMH burden[15,16,19]. We found significant associations with PVWMH for all previously reported loci associated for total WMH. Our meta-analysis discovered a high number of significant SNPs, genes and loci that are only associated with PVWMH, more than any single total WMH GWAS to date. Most genes and loci previously reported as significant in total WMH[29,30,32,37,39] were found to be associated with PVWMH alone, including *PLEKHG1*[29], *SH3PXD2A*[17,41,48] and *COL4A2*[48]. Most of the genes associated with PVWMH have direct effects on vascular function, or with vascular diseases such as ischemic stroke, intracranial hemorrhage or coronary artery disease. Other genes discussed as potential candidates in prior total WMH studies we now find to be significantly associated only with PVWMH including *DYDC2* and *NEURL1* as well as *NMT1*, *GALK1*, *H3F3B*, *UNK*, *UNC13D*, *EVPL*, *ICAL1*, *WDR12/CARF*, *NBEAL1*, and *EFEMP1*.

Genetic correlation analyses between DWMH and PVWMH and other phenotypes indicate independent and shared genetic architecture. The genetic correlation between the two phenotypes was high, as previously reported in a twin analysis[13]. Regional overlap analyses demonstrated shared genomic regions between PVWMH and DWMH at 2p16.1 (*EFEMP1*), 2q33.2 (*CARF* & *NBEAL*), 6q25.1 (*PLEKHGI*[29]), 16q24.2 (*C16orf95*), and 17q25.1 (*TRIM47*, *TRIM65*). Comparison of these phenotypes with ischemic and hemorrhagic stroke phenotypes from the

bioRxiv preprint doi: https://doi.org/10.1101/683367; this version posted June 27, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

MEGASTROKE Consortium[41] for shared genetic architecture yielded interesting results. In general, DWMH and PVWMH both had shared and unique associations with different genomic regions and stroke phenotypes, including "any stroke" and "ischemic stroke". For the first time, we demonstrated shared genetic architecture between any stroke, SVD stroke and PVWMH, at loci on chr 2 (*NBEAL1*), 3 (*PLSCR1*) and 10 (*SORBS1*). Significantly we discovered overlap for PVWMH only with deep intracranial hemorrhage (ICH) at genomic regions containing the genes *COL4A1/COL4A2*[49] (chr13) and *CHD6* (chr20). There was also overlap for PVWMH with SVD stroke at the chr 2 region, which includes the *NBEAL1* gene. Many of these genes are implicated in cerebral small vessel disease, such as *COL4A1/COL4A2* and *NBEAL1*. Deep ICH is a regional classification of ICH occurring in deep subcortical regions approximating the same brain region as PVWMH. Together, these results are suggestive of an underlying vascular pathophysiology influenced by variation in genes with influences on microvessels in the periventricular region that are associated with deep ICH, periventricular SVD, and periventricular ischemic stroke.

We identified 11 independent loci for PVWMH, and one locus for DWMH. The only significant locus observed for DWMH was the previously reported total WMH 17q25.1 locus[15,16,19,29], which was also found for PVWMH. Several of the identified loci overlap with prior total WMH GWAS results, specifically 17q25.1, 2p16.1, 2p21, 2q33.2, 6q25.1, 10q24.33 and 13q34[15,17,19,29]. Many of the previously reported candidates (17q25.1) achieved genome-wide significance for the first time in both PVWMH and DWMH including *TRIM47*, *TRIM65*, *MRPL38*, *WBP2*, *FBF1*, and *ACOX1*[15-17,19] indicating shared biology. For both phenotypes, the majority of the top SNPs from each locus were intronic. As discussed, several of these SNPs are reported cis eQTLs in various tissues. There were many genes located within 10kb of our significant SNPs (Supplementary Tables 9 and 10). Fourteen of these genes overlapped with the two phenotypes

bioRxiv preprint doi: https://doi.org/10.1101/683367; this version posted June 27, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

and are located in the gene-rich locus, 17q25.1, including *EVPL*, *UNK*, and *CDK3*. Interestingly, in a post-mortem study comparing gene expression in DWMH versus non-lesion control white matter, *EVPL* was downregulated in DWMH[50]. Pathway analyses did not find any significant results for PVWMH and only one pathway for DWMH was identified, which was not informative with respect to WMH.

PVWMH had significant associations with loci containing genes implicated in large and small vessel disease, as well as ischemic and hemorrhagic stroke. The association of these vascular genes with PVWMH and not DWMH suggests a potential underlying regional difference in the microvasculature affecting the risk of vascular injury in the periventricular region. The different physiology of the short penetrating microvessels emerging from the midline arteries of the brain unique to the periventricular region may provide explanations for our observed differences when comparing the two phenotype results [1,7,21,23]. Some of these genes such as *SH3PXD2A* have been previously associated with total WMH and ischemic stroke[16,41]. The *NOS3* gene is associated with coronary artery disease, migraine, vascular dysfunction, SVD, and ischemic stroke[29,43,51,52]. *PLEKHG1* is associated with dementia and ischemic stroke[17,53]. The most notable associated vascular gene is *COL4A2* that encodes for a subunit of type IV collagen, which has been associated with SVD, ischemic stroke, intracranial hemorrhage, and coronary artery disease[17,49,54-56]. It is a proposed therapeutic target for the prevention of intracranial hemorrhage[57,58]. Verhaaren et al.[19] identified a suggestive intronic *COL4A2* SNP for total WMH, rs9515201, which was genome-wide significant in Traylor et al.[29]. Our genome-wide significant top SNP (intronic) in the *COL4A2* gene, rs11838776, is in high LD with rs9515201 (R^2=0.83). New vascular genes not previously reported associated with total WMH, include the *MAT1A*

bioRxiv preprint doi: https://doi.org/10.1101/683367; this version posted June 27, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

gene, which catalyzes adenosine transfer from ATP to methionine and is associated with stroke and hypertension[59]. *FAM213A* encodes for the peroxiredoxin like 2A protein, which is an antioxidant associated with endothelial activation[60]. The *CALCRL* gene encodes the calcitonin receptor like receptor and modulates the calcitonin signaling system, implicated in ischemic stroke [61].

Other newly significant non-vascular PVWMH genes identified by gene-based analysis include *TSPAN14* (10q23.1), which encodes one of the tetraspanins, which organize a network of interactions referred to as the tetraspanin web. This includes ADAM10 - a metalloprotease that cleaves the precursor of cell surface proteins, including the amyloid precursor protein and Notch[62]. Another newly identified gene, *VCAN* (5q14.2-q14.3), encodes a large chondroitin sulfate proteoglycan that is found in the extracellular matrix. In a recent meta-analysis, *VCAN* was associated with white matter microstructural integrity in UK Biobank participants with WMH, stroke, and dementia[63]. *KLHL24* (3q27.1), encodes a ubiquitin ligase substrate receptor that is regulated by auto-ubiquitination and is associated with deficits in skin keratization[64]. These candidate genes for PVWMH may influence the immediate tissues surrounding microvessels and may contribute to SVD-associated biological changes.

DWMH and PVWMH share a previously reported locus on 17q25.1 for total WMH, with these SNP associations having the largest effect sizes for both phenotypes. Thus, this locus appears to influence both WMH subtypes and is the most predominant genetic association with the largest potential clinical effect for both phenotypes. The top genome-wide significant hits for DWMH and PVWMH (17q25.1) were either identical with the SNP recently reported by Traylor et al[29]

bioRxiv preprint doi: https://doi.org/10.1101/683367; this version posted June 27, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

for total WMH (PVWMH SNP rs3744020) or in high LD (R^2>0.9) with the previously identified top ranked SNPs from prior total WMH GWAS in the same locus (rs3744028, Fornage et al.[15], rs7214628, Verhaaren et al.[19]). These SNPs were only in moderate LD (R^2≤0.396) with the top SNP (rs3760128) identified in a recent exome association analysis[16]. All of these SNPs fall within or between the previously reported *TRIM47* and *TRIM65* genes[15,17,19,29]. This gene-rich locus contains genes that influence glial cell proliferation and have been hypothesized to influence gliosis, which is a histological and MRI marker of microvascular injury[1]. It includes previously identified total WMH genes, such as *TRIM47/TRIM65* (glial proliferation, astrocytoma's)[15,19], the *ACOX1* gene (cell replication, hepatic cancer)[15,16,19,65] and the *MRPL38* gene (protein synthesis)[16]. Genes associated with neuronal injury and/or neurodegenerative disorders are also found in the 17q25.1 locus, including *CDK3* (neuronal cell death in stroke)[66], *H3F3B* (schizophrenia pathogenesis)[67] and the *GALK1* gene (galactosemia, a neurodegenerative metabolic disease)[68]. Interestingly, two genome-wide significant intronic *UNC13D* SNPs identified in this study and reported previously for total WMH burden[19], rs9894244 and rs7216615, have been reported as eQTLs for *GALK1* and *H3F3B* respectively[69].

The PVWMH specific loci also contained genes that potentially influence astrocytic function and gliosis. There were several previously reported total WMH candidate genes such as *NBEAL1* (2q33.2, astrocytoma)[16,19], *WDR12* (2q33.2, cell cycle, apoptosis, myocardial infarction, coronary artery disease, carotid intima-media thickness)[16], *NEURL1* (10q24.33, notch signaling, cell proliferation in medulloblastoma, atrial fibrillation, total WMH[15,16,19]), *CARF* (2q33.2, modulates BDNF expression, tumor suppression, astrocyte function[70]) and *EFEMP1* (2p16.1, malignant gliomas[54]). Newly identified genes include the *NMT1* gene (17q21.31), which

bioRxiv preprint doi: https://doi.org/10.1101/683367; this version posted June 27, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

catalyzes the transfer of myristate from CoA to proteins[71]. The *ICA1L*, islet cell autoantigen 1-like protein gene, approximating CARF, is associated with amyotrophic lateral sclerosis and may influence neuronal death in that disease[72]. *POLR2F* (22q13.1) is associated with colon cancer[73] while *OBFC1* (10q24.33) initiates DNA replication[74] and *DYDC2* (10q23.1) is associated with myofibrillar myopathy and ventricular cardiomegaly[75].

While our GWAS were only slightly larger than the previous biggest GWAS on total WMH burden with 21,079 participants[19], our detection rate of significant SNPs was substantially higher. This improved detection may be the result of reduced heterogeneity by separately analyzing the DWMH and PVWMH phenotypes. Additionally, sensitivity also may have been increased by including more modern samples (e.g., from cohorts such as UK Biobank), with most data acquired on the more powerful 1.5-3.0 Tesla MRI scanners using sequences such as FLAIR and incorporating segmentation programs with better lesion detection and volumetric quantification. The results obtained did not change significantly when cohorts with visual rather than volumetric measures were excluded from the analysis.

Shortcomings of this study include the potential presence of phenotype heterogeneity due to the different WMH extraction algorithms used, with a minority of samples using visual ratings. However, this is a common problem encountered in this type of study[15,16,19]. The majority of the participants in this study were Caucasian and hence these results may not apply to other ethnicities. Sex differences have been previously reported but were also not examined[76]. For the Phase II meta-analysis, we did not have an independent replication cohort. Even though our results suggest improved power and reduction in potential bias through the discrimination of PVWMH from DWMH, the Euclidean methodology used by the majority of studies, which

bioRxiv preprint doi: https://doi.org/10.1101/683367; this version posted June 27, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

prescribed PVWMH as all WMH within one centimeter of the ventricle walls, undoubtedly missed PVWMH lesions outside this boundary. It is also possible that a number of DWMH lesions were accidentally included although review of the data makes this seem less likely. There are unique "cuff regions" or "border zone regions", which fall within the PVWMH phenotype that represent a lesion type that is specific to the regions abutting the anterior and posterior horns of the lateral ventricles. These lesions have been attributed with an as yet not fully classified pathophysiology that could be quite different from PVWMH and DWMH. This could represent a hitherto unclassified source of phenotype heterogeneity unaccounted for in the present analysis. Further, we did not investigate other WMH sub-classifications or alternative measures. For example, other regional sub-classifications may prove fruitful (e.g. lobar).

Conclusion

When using the two anatomically based sub-classifications (PVWMH and DWMH), we identified new genetic associations with WMH. Novel genetic associations unique to PVWMH alone were found and include genes such as *CALCRL*, *KLHL24*, *VCAN*, *MAT1A*, *OBFC1*, *DCAKD*, and *POLRF2*. We observed that the PVWMH phenotype was associated with genes previously associated with vascular phenotypes such as SVD, ICH and ischemic stroke. We found shared genetic architecture between PVWMH and deep ICH that includes loci containing the genes *COL4A1/COL4A2* and *CHD6*, supporting a strong etiologic association between these clinical phenotypes and the periventricular region. Our findings identify underlying cellular processes that are involved in the response to chronic brain injury. These include genes implicated in astrocytic/glial and vascular responses. Finally, we discovered a new set of putative PVWMH genes including *TSPAN14* and *VCAN*, which suggest new mechanisms that may

bioRxiv preprint doi: https://doi.org/10.1101/683367; this version posted June 27, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

influence the connective tissue outside the microvasculature. In conclusion, our findings suggest that the two WMH subclassifications have shared (17q25.1) but also unique and divergent genetic underpinnings. This work provides stimulus for further study to assess the roles of the identified genetic variants and genes in diseases of the aging brain.

bioRxiv preprint doi: https://doi.org/10.1101/683367; this version posted June 27, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

References

1. Wardlaw, J.M. *et al.* Neuroimaging standards for research into small vessel disease and its contribution to ageing and neurodegeneration. *Lancet Neurol* **12**, 822-38 (2013).
2. Wen, W., Sachdev, P.S., Li, J.J., Chen, X. & Anstey, K.J. White matter hyperintensities in the forties: their prevalence and topography in an epidemiological sample aged 44-48. *Hum Brain Mapp* **30**, 1155-67 (2009).
3. Vermeer, S.E. *et al.* Silent brain infarcts and white matter lesions increase stroke risk in the general population: the Rotterdam Scan Study. *Stroke* **34**, 1126-9 (2003).
4. Gouw, A.A. *et al.* Heterogeneity of small vessel disease: a systematic review of MRI and histopathology correlations. *J Neurol Neurosurg Psychiatry* **82**, 126-35 (2011).
5. Wardlaw, J.M., Smith, C. & Dichgans, M. Mechanisms of sporadic cerebral small vessel disease: insights from neuroimaging. *Lancet Neurol* **12**, 483-97 (2013).
6. Fazekas, F. *et al.* Pathologic correlates of incidental MRI white matter signal hyperintensities. *Neurology* **43**, 1683-9 (1993).
7. Vermeer, S.E. *et al.* Silent brain infarcts and the risk of dementia and cognitive decline. *N Engl J Med* **348**, 1215-22 (2003).
8. Fernando, M.S. *et al.* White matter lesions in an unselected cohort of the elderly: molecular pathology suggests origin from chronic hypoperfusion injury. *Stroke* **37**, 1391-8 (2006).
9. Iadecola, C. The pathobiology of vascular dementia. *Neuron* **80**, 844-66 (2013).
10. Mok, V.C. *et al.* Effects of statins on the progression of cerebral white matter lesion: Post hoc analysis of the ROCAS (Regression of Cerebral Artery Stenosis) study. *J Neurol* **256**, 750-7 (2009).
11. Godin, O., Tzourio, C., Maillard, P., Mazoyer, B. & Dufouil, C. Antihypertensive treatment and change in blood pressure are associated with the progression of white matter lesion volumes: the Three-City (3C)-Dijon Magnetic Resonance Imaging Study. *Circulation* **123**, 266-73 (2011).
12. Gottesman, R.F. *et al.* Blood pressure and white-matter disease progression in a biethnic cohort: Atherosclerosis Risk in Communities (ARIC) study. *Stroke* **41**, 3-8 (2010).
13. Sachdev, P.S. *et al.* White Matter Hyperintensities Are Under Strong Genetic Influence. *Stroke* **47**, 1422-8 (2016).
14. Fennema-Notestine, C. *et al.* White matter disease in midlife is heritable, related to hypertension, and shares some genetic influence with systolic blood pressure. *Neuroimage Clin* **12**, 737-745 (2016).
15. Fornage, M. *et al.* Genome-wide association studies of cerebral white matter lesion burden: the CHARGE consortium. *Ann Neurol* **69**, 928-39 (2011).
16. Jian, X. *et al.* Exome Chip Analysis Identifies Low-Frequency and Rare Variants in MRPL38 for White Matter Hyperintensities on Brain Magnetic Resonance Imaging. *Stroke* **49**, 1812-1819 (2018).
17. Traylor, M. *et al.* Genome-wide meta-analysis of cerebral white matter hyperintensities in patients with stroke. *Neurology* **86**, 146-53 (2016).
18. Verhaaren, B.F. *et al.* Replication study of chr17q25 with cerebral white matter lesion volume. *Stroke* **42**, 3297-9 (2011).

bioRxiv preprint doi: https://doi.org/10.1101/683367; this version posted June 27, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

19. Verhaaren, B.F. *et al.* Multiethnic genome-wide association study of cerebral white matter hyperintensities on MRI. *Circ Cardiovasc Genet* **8**, 398-409 (2015).
20. Griffanti, L. *et al.* Classification and characterization of periventricular and deep white matter hyperintensities on MRI: A study in older adults. *Neuroimage* **170**, 174-181 (2018).
21. Doubal, F.N., MacLullich, A.M., Ferguson, K.J., Dennis, M.S. & Wardlaw, J.M. Enlarged perivascular spaces on MRI are a feature of cerebral small vessel disease. *Stroke* **41**, 450-4 (2010).
22. Albrecht, M. *et al.* Remote ischemic preconditioning regulates HIF-1alpha levels, apoptosis and inflammation in heart tissue of cardiosurgical patients: a pilot experimental study. *Basic Res Cardiol* **108**, 314 (2013).
23. Blanco, P.J., Muller, L.O. & Spence, J.D. Blood pressure gradients in cerebral arteries: a clue to pathogenesis of cerebral small vessel disease. *Stroke Vasc Neurol* **2**, 108-117 (2017).
24. De Reuck, J. The human periventricular arterial blood supply and the anatomy of cerebral infarctions. *Eur Neurol* **5**, 321-34 (1971).
25. Potter, G.M. *et al.* Enlarged perivascular spaces and cerebral small vessel disease. *Int J Stroke* **10**, 376-81 (2015).
26. Krishnan, M.S. *et al.* Relationship between periventricular and deep white matter lesions and depressive symptoms in older people. The LADIS Study. *Int J Geriatr Psychiatry* **21**, 983-9 (2006).
27. Kreisel, S.H. *et al.* Deterioration of gait and balance over time: the effects of age-related white matter change--the LADIS study. *Cerebrovasc Dis* **35**, 544-53 (2013).
28. Nyquist, P.A. *et al.* Age differences in periventricular and deep white matter lesions. *Neurobiol Aging* **36**, 1653-8 (2015).
29. Traylor, M. *et al.* Genetic variation in PLEKHG1 is associated with white matter hyperintensities (n = 11,226). *Neurology* **92**, e749-e757 (2019).
30. Elliott, L.T. *et al.* Genome-wide association studies of brain imaging phenotypes in UK Biobank. *Nature* **562**, 210-216 (2018).
31. Winkler, T.W. *et al.* Quality control and conduct of genome-wide association meta-analyses. *Nat Protoc* **9**, 1192-212 (2014).
32. Willer, C.J., Li, Y. & Abecasis, G.R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190-1 (2010).
33. Watanabe, K., Taskesen, E., van Bochoven, A. & Posthuma, D. Functional mapping and annotation of genetic associations with FUMA. *Nat Commun* **8**, 1826 (2017).
34. Pruim, R.J. *et al.* LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics* **26**, 2336-7 (2010).
35. Dayem Ullah, A.Z. *et al.* SNPnexus: assessing the functional relevance of genetic variation to facilitate the promise of precision medicine. *Nucleic Acids Res* **46**, W109-W113 (2018).
36. Carithers, L.J. *et al.* A Novel Approach to High-Quality Postmortem Tissue Procurement: The GTEx Project. *Biopreserv Biobank* **13**, 311-9 (2015).
37. Arnold, M., Raffler, J., Pfeufer, A., Suhre, K. & Kastenmuller, G. SNiPA: an interactive, genetic variant-centered annotation browser. *Bioinformatics* **31**, 1334-6 (2015).
38. de Leeuw, C.A., Mooij, J.M., Heskes, T. & Posthuma, D. MAGMA: generalized gene-set analysis of GWAS data. *PLoS Comput Biol* **11**, e1004219 (2015).

bioRxiv preprint doi: https://doi.org/10.1101/683367; this version posted June 27, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

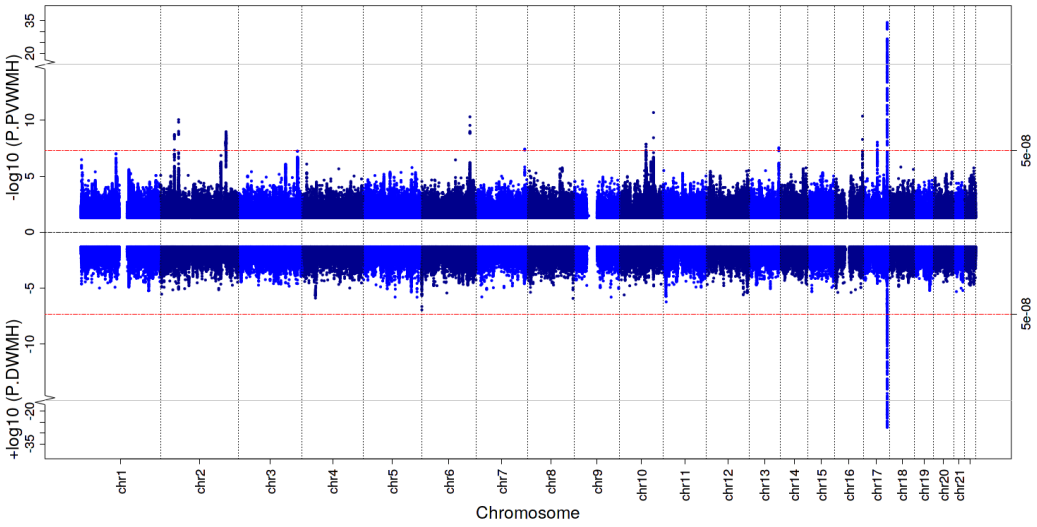
39. Yang, J., Lee, S.H., Goddard, M.E. & Visscher, P.M. GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet* **88**, 76-82 (2011).
40. Bulik-Sullivan, B.K. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet* **47**, 291-5 (2015).
41. Malik, R. *et al.* Multiancestry genome-wide association study of 520,000 subjects identifies 32 loci associated with stroke and stroke subtypes. *Nat Genet* **50**, 524-537 (2018).
42. Lambert, J.C. *et al.* Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat Genet* **45**, 1452-8 (2013).
43. Malik, R. *et al.* Genome-wide meta-analysis identifies 3 novel loci associated with stroke. *Ann Neurol* **84**, 934-939 (2018).
44. Pickrell, J.K. *et al.* Detection and interpretation of shared genetic influences on 42 human traits. *Nat Genet* **48**, 709-17 (2016).
45. Pickrell, J.K. Joint analysis of functional genomic data and genome-wide association studies of 18 human traits. *Am J Hum Genet* **94**, 559-73 (2014).
46. Pickrell, J.K. *et al.* Detection and interpretation of shared genetic influences on 42 human traits (vol 48, pg 709, 2016). *Nature Genetics* **48**(2016).
47. Shi, H., Mancuso, N., Spendlove, S. & Pasaniuc, B. Local Genetic Correlation Gives Insights into the Shared Genetic Architecture of Complex Traits. *Am J Hum Genet* **101**, 737-751 (2017).
48. Traylor, M. *et al.* Genetic variation at 16q24.2 is associated with small vessel stroke. *Ann Neurol* **81**, 383-394 (2017).
49. Rannikmae, K. *et al.* Common variation in COL4A1/COL4A2 is associated with sporadic cerebral small vessel disease. *Neurology* **84**, 918-26 (2015).
50. Simpson, J.E. *et al.* Microarray RNA expression analysis of cerebral white matter lesions reveals changes in multiple functional pathways. *Stroke* **40**, 369-75 (2009).
51. Chen, X.J. *et al.* The association between an endothelial nitric oxide synthase gene polymorphism and coronary heart disease in young people and the underlying mechanism. *Mol Med Rep* **17**, 3928-3934 (2018).
52. Bastian, C. *et al.* NOS3 Inhibition Confers Post-Ischemic Protection to Young and Aging White Matter Integrity by Conserving Mitochondrial Dynamics and Miro-2 Levels. *J Neurosci* **38**, 6247-6266 (2018).
53. Traylor, M. & Lewis, C.M. Genetic discovery in multi-ethnic populations. *Eur J Hum Genet* **24**, 1097-8 (2016).
54. Yang, L. *et al.* Impact of interaction between the G870A and EFEMP1 gene polymorphism on glioma risk in Chinese Han population. *Oncotarget* **8**, 37561-37567 (2017).
55. Verbeek, E. *et al.* COL4A2 mutation associated with familial porencephaly and small-vessel disease. *Eur J Hum Genet* **20**, 844-51 (2012).
56. Rannikmae, K. *et al.* COL4A2 is associated with lacunar ischemic stroke and deep ICH: Meta-analyses among 21,500 cases and 40,600 controls. *Neurology* **89**, 1829-1839 (2017).
57. Kuo, D.S., Labelle-Dumais, C. & Gould, D.B. COL4A1 and COL4A2 mutations and disease: insights into pathogenic mechanisms and potential therapeutic targets. *Hum Mol Genet* **21**, R97-110 (2012).

bioRxiv preprint doi: https://doi.org/10.1101/683367; this version posted June 27, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

58. Murray, L.S. *et al.* Chemical chaperone treatment reduces intracellular accumulation of mutant collagen IV and ameliorates the cellular phenotype of a COL4A2 mutation that causes haemorrhagic stroke. *Hum Mol Genet* **23**, 283-92 (2014).
59. Lai, C.Q. *et al.* MAT1A variants are associated with hypertension, stroke, and markers of DNA damage and are modulated by plasma vitamin B-6 and folate. *Am J Clin Nutr* **91**, 1377-86 (2010).
60. He, H. *et al.* Adiporedoxin suppresses endothelial activation via inhibiting MAPK and NF-kappaB signaling. *Sci Rep* **6**, 38975 (2016).
61. Koyama, T. *et al.* Genetic Variants of RAMP2 and CLR are Associated with Stroke. *J Atheroscler Thromb* **24**, 1267-1281 (2017).
62. Saint-Pol, J. *et al.* Regulation of the trafficking and the function of the metalloprotease ADAM10 by tetraspanins. *Biochem Soc Trans* **45**, 937-44 (2017).
63. Rutten-Jacobs, L.C.A. *et al.* Genetic Study of White Matter Integrity in UK Biobank (N=8448) and the Overlap With Stroke, Depression, and Dementia. *Stroke* **49**, 1340-1347 (2018).
64. Has, C. The "Kelch" Surprise: KLHL24, a New Player in the Pathogenesis of Skin Fragility. *J Invest Dermatol* **137**, 1211-1212 (2017).
65. Zeng, J. *et al.* Specific Inhibition of Acyl-CoA Oxidase-1 by an Acetylenic Acid Improves Hepatic Lipid and Reactive Oxygen Species (ROS) Metabolism in Rats Fed a High Fat Diet. *J Biol Chem* **292**, 3800-3809 (2017).
66. Osuga, H. *et al.* Cyclin-dependent kinases as a therapeutic target for stroke. *Proc Natl Acad Sci U S A* **97**, 10254-9 (2000).
67. Manley, W. *et al.* Validation of a microRNA target site polymorphism in H3F3B that is potentially associated with a broad schizophrenia phenotype. *PLoS One* **13**, e0194233 (2018).
68. Timmers, I. *et al.* White matter microstructure pathology in classic galactosemia revealed by neurite orientation dispersion and density imaging. *J Inherit Metab Dis* **38**, 295-304 (2015).
69. Wild, P.S. *et al.* Large-scale genome-wide analysis identifies genetic variants associated with cardiac structure and function. *J Clin Invest* **127**, 1798-1812 (2017).
70. Tao, X., West, A.E., Chen, W.G., Corfas, G. & Greenberg, M.E. A calcium-responsive transcription factor, CaRF, that regulates neuronal activity-dependent expression of BDNF. *Neuron* **33**, 383-95 (2002).
71. Lu, Y. *et al.* Expression of N-myristoyltransferase in human brain tumors. *Neurochem Res* **30**, 9-13 (2005).
72. Hadano, S. *et al.* A gene encoding a putative GTPase regulator is mutated in familial amyotrophic lateral sclerosis 2. *Nat Genet* **29**, 166-73 (2001).
73. Antonacopoulou, A.G. *et al.* POLR2F, ATP6V0A1 and PRNP expression in colorectal cancer: new molecules with prognostic significance? *Anticancer Res* **28**, 1221-7 (2008).
74. Casteel, D.E. *et al.* A DNA polymerase-{alpha}{middle dot}primase cofactor with homology to replication protein A-32 regulates DNA replication in mammalian cells. *J Biol Chem* **284**, 5807-18 (2009).
75. Kuhl, A. *et al.* Myofibrillar myopathy with arrhythmogenic right ventricular cardiomyopathy 7: corroboration and narrowing of the critical region on 10q22.3. *Eur J Hum Genet* **16**, 367-73 (2008).

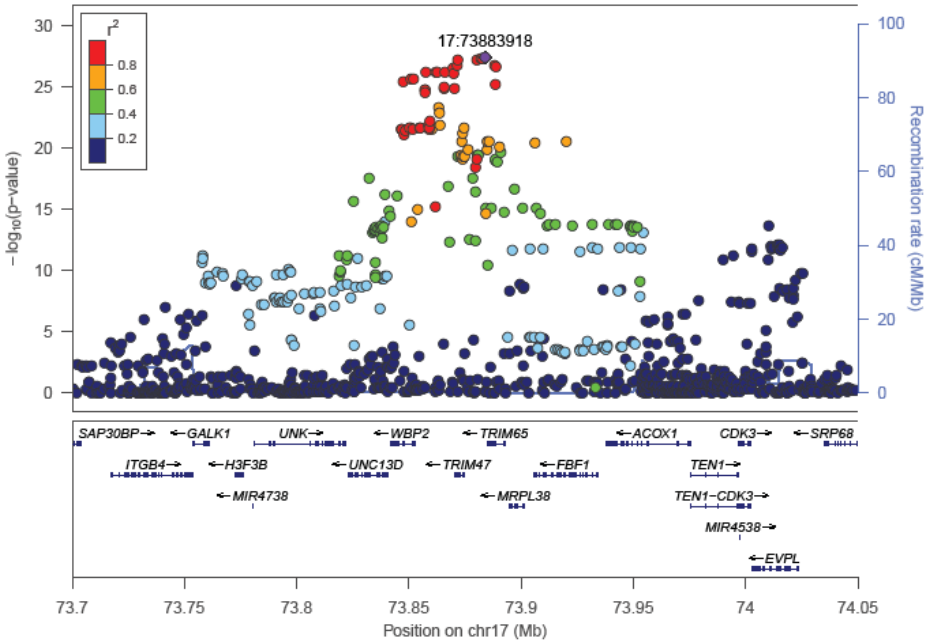
bioRxiv preprint doi: https://doi.org/10.1101/683367; this version posted June 27, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

76. Sachdev, P., Chen, X. & Wen, W. White matter hyperintensities in mid-adult life. *Curr Opin Psychiatry* **21**, 268-74 (2008).



bioRxiv preprint doi: https://doi.org/10.1101/683367; this version posted June 27, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

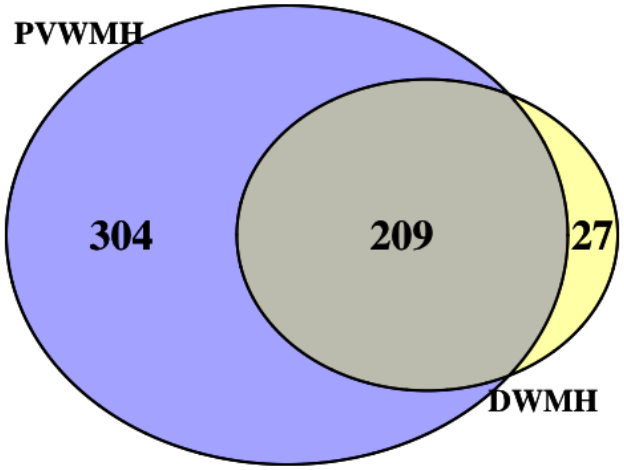
(a)



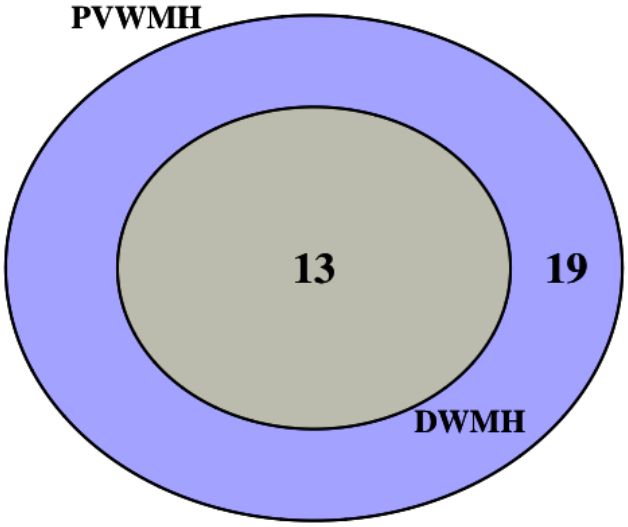
(b)

Figure 1. (a) Phase II GWAS meta-analysis results (n=26,654). Miami plot for periventricular WMH (PVWMH, upper panel) and deep WMH (DWMH, lower panel). Red dashed line shows genome-wide significance threshold (*p*<5e-8). (1b) Chr17 regional plot of genome-wide significant SNPs for DWMH. Colors of the SNPs indicate the level of LD with the top SNP (purple), rs35392904.

bioRxiv preprint doi: https://doi.org/10.1101/683367; this version posted June 27, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

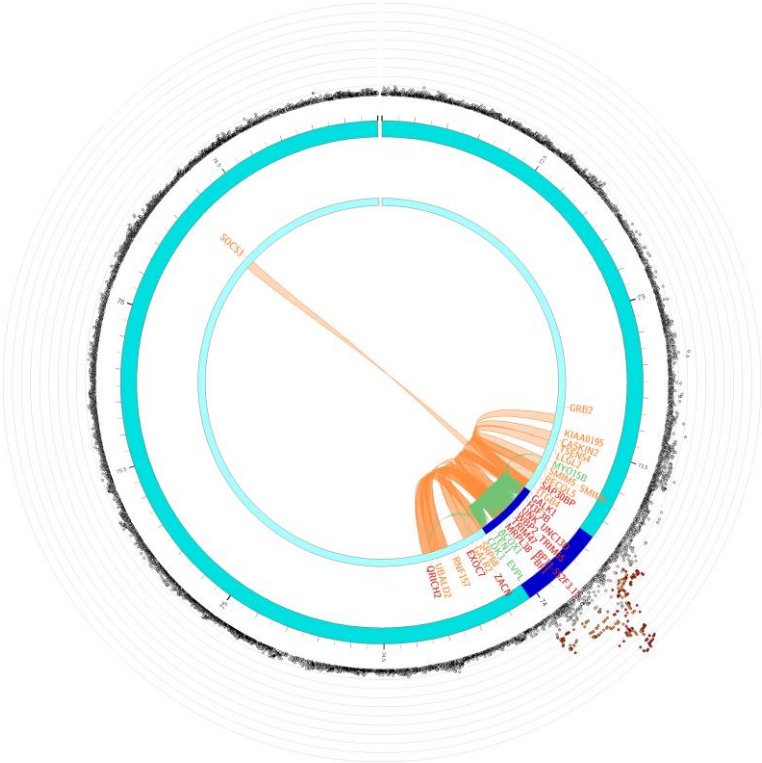


(a)

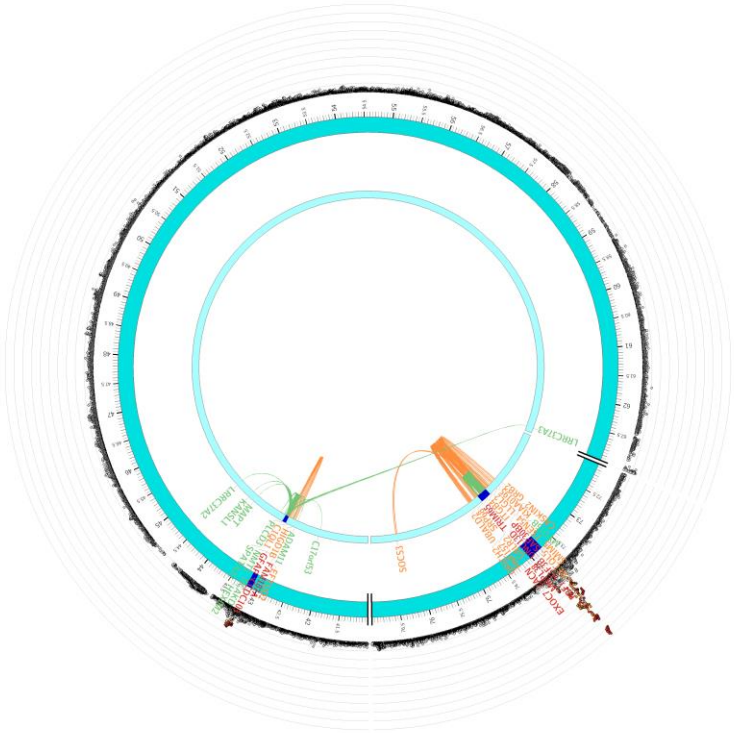


(b)

bioRxiv preprint doi: https://doi.org/10.1101/683367; this version posted June 27, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



(c) Chr17 DWMH



(d) Chr17 PVWMH

bioRxiv preprint doi: https://doi.org/10.1101/683367; this version posted June 27, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure 2. (a) Overlap between genome-wide significant SNPs (*p*<5e-8) for DWMH and PVWMH. (b) Overlap between significant genes identified by MAGMA for both phenotypes. (c - d) Circos plots for chr17 for both phenotypes, showing one identified region for DWMH but two for PVWMH. Outer ring shows SNPs <.05 with the most significant SNPs located towards the outermost ring. SNPs in high LD with the independent significant SNPs in each locus are colored in red (r^2>.8)-blue (r^2>.2); no LD (grey). Genomic risk loci are colored in dark blue (2nd layer). Genes are mapped by chromatin interaction (orange), eQTL (green) or both (red).

bioRxiv preprint doi: https://doi.org/10.1101/683367; this version posted June 27, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

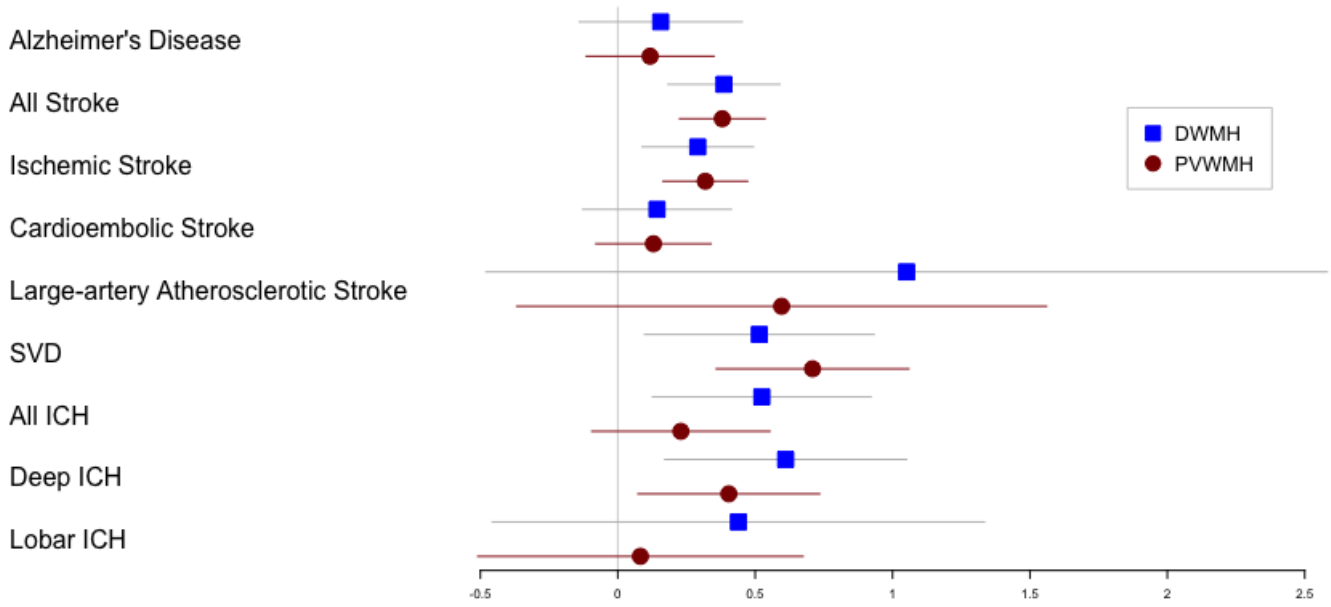


Figure 3. Genetic correlations (r_g) between deep WMH (DWMH), periventricular WMH (PVWMH), Alzheimer's disease (AD) and stroke phenotypes. Horizontal bars represent standard errors and the size of the square corresponds precision. SVD = small vessel disease stroke, All ICH = All intracranial hemorrhage, Deep ICH = deep intracranial hemorrhage, Lobar ICH = lobar intracranial hemorrhage

Table 1. Top genome-wide significant SNP results from each genomic locus identified from the Phase II GWAS meta-analysis for deep and periventricular (PV) WMH.

WMH phenotype	rsID	CHR	POS	Nearest gene	Function/ Position	Allele 1	Allele2	Freq (A1)	Beta (A1)	StdErr	*N*	Di
PV	rs3744020	17q25.1	73871773	*TRIM47*	Intronic	A	G	0.1897	0.0899	0.0073	26438	++
Deep	rs35392904	17q25.1	73883918	*TRIM65*	Intronic	T	C	0.7981	-0.0765	0.0070	26642	
PV	rs3758575	10q24.33	105454881	*SH3PXD2A*	Intronic	A	G	0.4904	0.0388	0.0058	26654	++
PV	**rs12928520**	**16q24.2**	**87237568**	***C16orf95***	**Intergenic**	**T**	**C**	**0.4252**	**0.0431**	**0.0065**	**26327**	**++**
PV	rs275350	6q25.1	151016058	*PLEKHG1*	Intronic	C	G	0.4202	0.0374	0.0057	26654	+-
PV	rs7596872	2p16.1	56128091	*EFEMP1*	Intronic	A	C	0.0975	0.0642	0.0099	25730	-+
PV	rs72934583	2q33.2	204009057	*NBEAL1*	Intronic	T	G	0.8740	0.0529	0.0087	25730	-+
PV	rs57242328	2p21	43073247	*AC098824.6*	Intergenic	A	G	0.3317	-0.0368	0.0061	25730	---
PV	**rs7213273**	**17q21.31**	**43155914**	***NMT1***	**Intronic**	**A**	**G**	**0.6668**	**0.0341**	**0.0059**	**26111**	**++**
PV	**rs1993484**	**10q23.1**	**82222698**	***TSPAN14***	**Intronic**	**T**	**C**	**0.2388**	**0.0378**	**0.0067**	**26654**	**++**
PV	rs11838776	13q34	111040681	*COL4A2*	Intronic	A	G	0.2793	0.0350	0.0063	26654	-+
PV	**rs1799983**	**7q36.1**	**150696111**	***NOS3***	**Exonic**	**T**	**G**	**0.3201**	**0.0373**	**0.0068**	**26654**	**++**

Notes: Effect allele is allele 1. Those loci bolded have not been previously associated with total WMH.

bioRxiv preprint doi: https://doi.org/10.1101/683367; this version posted June 27, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Table 2. Thirty-two significant genes were identified for PVWMH using gene-based tests (p<2.66e-6). Thirteen of these genes (chr17) were also significant for DWMH (*).

GENE	CHR	START	STOP	N SNPS	N	*p* PVWMH	*p* DWMH
WBP2	17	73841780	73852588	28	24682	3.19E-26	1.16E-21*
TRIM65	17	73876416	73893084	52	24555	7.73E-24	9.12E-19*
TRIM47	17	73870242	73874656	13	24185	1.70E-23	9.04E-19*
RP11-552F3.12	17	73894726	73926210	53	24351	2.15E-20	1.76E-15*
FBF1	17	73905655	73937221	55	24338	3.98E-17	1.23E-13*
GALK1	17	73747675	73761792	36	24307	6.34E-16	3.23E-14*
MRPL38	17	73894724	73905899	21	24481	7.62E-15	1.18E-13*
UNC13D	17	73823306	73840798	73	23788	3.10E-14	1.22E-13*
UNK	17	73780681	73821886	120	22768	3.28E-13	4.85E-10*
H3F3B	17	73772515	73781974	23	24009	4.43E-12	1.41E-10*
SH3PXD2A	10	105348285	105615301	788	24847	8.43E-12	0.21731
ACOX1	17	73937588	73975515	151	24198	7.72E-11	1.1E-09*
EVPL	17	74000583	74023533	67	24582	1.26E-10	2.82E-14*
PLEKHG1	6	150920999	151164799	1022	24922	1.59E-10	0.011765
WDR12	2	203739505	203879521	322	23753	2.53E-10	0.00104
ICA1L	2	203640690	203736708	224	23843	8.44E-10	0.001301
CARF	2	203776937	203851786	157	24076	2.41E-09	0.001763
NMT1	17	43128978	43186384	221	24766	7.18E-08	0.00034
CDK3	17	73996987	74002080	12	24433	8.54E-08	1.82E-08*
OBFC1	10	105642300	105677963	99	25461	1.41E-07	0.054127
NOS3	7	150688083	150711676	58	24608	1.73E-07	0.000371
DCAKD	17	43100708	43138473	111	25229	2.60E-07	0.000363
DYDC2	10	82104501	82127829	91	25050	2.88E-07	0.003460
NBEAL1	2	203879602	204091101	367	23413	3.83E-07	0.040539
NEURL1	10	105253736	105352309	296	25038	4.84E-07	0.098303
MAT1A	10	82031576	82049440	66	25295	4.90E-07	0.002421
TSPAN14	10	82213922	82292879	213	24731	6.73E-07	0.006605
CALCRL	2	188207856	188313187	278	24309	7.87E-07	0.000574
KLHL24	3	183353356	183402265	207	24356	1.29E-06	0.002571
POLR2F	22	38348614	38437922	105	23525	1.94E-06	0.252540
VCAN	5	82767284	82878122	316	24248	2.52E-06	0.065044
COL4A2	13	110958159	111165374	1140	24876	2.61E-06	0.365300

Notes: Those loci bolded have not been previously associated with total WMH.