Genes From a Translational Analysis Support a Multifactorial Nature of White Matter Hyperintensities

Lorna M. Lopez, PhD; W. David Hill, MSc; Sarah E. Harris, PhD; Maria Valdes Hernandez, PhD; Susana Munoz Maniega, PhD; Mark E. Bastin, PhD; Emma Bailey, PhD; Colin Smith, MD; Martin McBride, PhD; John McClure, PhD; Delyth Graham, PhD; Anna Dominiczak, MD; Qiong Yang, PhD; Myriam Fornage, PhD; M. Arfan Ikram, MD, PhD; Stephanie Debette, MD, PhD; Lenore Launer, PhD; Joshua C. Bis, PhD; Reinhold Schmidt, MD; Sudha Seshadri, MD; David J. Porteous, PhD; John Starr, MD; Ian J. Deary, PhD; Joanna M. Wardlaw, MD

Background and Purpose—White matter hyperintensities (WMH) of presumed vascular origin increase the risk of stroke and dementia. Despite strong WMH heritability, few gene associations have been identified. Relevant experimental models may be informative.

Methods—We tested the associations between genes that were differentially expressed in brains of young spontaneously hypertensive stroke–prone rats and human WMH (using volume and visual score) in 621 subjects from the Lothian Birth Cohort 1936 (LBC1936). We then attempted replication in 9361 subjects from the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE). We also tested the subjects from LBC1936 for previous genome-wide WMH associations found in subjects from CHARGE.

Results—Of 126 spontaneously hypertensive stroke–prone rat genes, 10 were nominally associated with WMH volume or score in subjects from LBC1936, of which 5 (AFP, ALB, GNAI1, RBM8a, and MRPL18) were associated with both WMH volume and score (P<0.05); 2 of the 10 (XPNPEP1, P=6.7×10⁻⁴; FAR1, P=0.024) plus another spontaneously hypertensive stroke–prone rat gene (USMG5, P=0.00014), on chromosomes 10, 13, and 10 respectively, were associated with WMH in subjects from CHARGE. Gene set enrichment showed significant associations for downregulated spontaneously hypertensive stroke–prone rat genes with WMH in humans. In subjects from LBC1936, we replicated CHARGE’s genome-wide WMH associations on chromosomes 17 (TRIM65 and TRIM47) and, for the first time, 1 (PFM1).

Conclusions—Despite not passing multiple testing thresholds individually, these genes collectively are relevant to known WMH associations, proposed WMH mechanisms, or dementia: associations with Alzheimer’s disease, late-life depression, ATP production, osmotic regulation, neurodevelopmental abnormalities, and cognitive impairment. If replicated further, they suggest a multifactorial nature for WMH and argue for more consideration of vascular contributions to dementia. (Stroke. 2015;46:341-347. DOI: 10.1161/STROKEAHA.114.007649.)

Key Words: genetics ■ humans ■ leukoencephalopathies ■ magnetic resonance imaging
White matter hyperintensities (WMH) of presumed vascular origin, a major component of cerebral small vessel disease (SVD), double the risk of stroke and dementia. Despite considerable societal effect, the causes of WMH and SVD are poorly understood. Conventional vascular risk factors explain little of the WMH variance. Several rare monogenic SVD disorders, and epidemiology suggest that genetic predisposition is important.

Identification of genetic factors for SVD has been challenging. Several replicable single-nucleotide polymorphisms (SNPs) associated with WMH have been identified in 1 locus on chromosome 17q25,2,3 although the exact gene(s) and biological pathways to WMH are unclear. Few other replicable genes have been found in genome-wide association studies (GWAS),4,5 and little is known of their functional significance.

Experimental SVD models provide insight into human SVD. The spontaneously hypertensive stroke-prone rat (SHRSP) is a relevant model of spontaneous SVD.4 It was selectively crossbred (1974) from Wistar-Kyoto (WKY) rats via the spontaneously hypertensive rat (SHR, 1963).2 Hypertension, established in SHRSP rats by 10 weeks of age, is considered to be the main cause of their brain disease. However, differences in protein and gene expression in SHRSP rats versus WKY rats at 5 weeks of age (before measurable blood pressure rises) suggest underlying susceptibilities to SVD.6 Compared with WKY controls, 5-week-old SHRSP rats have reduced claudin 5 (tight junction) and myelin basic protein and increased microglia (IBA1) and glial activation (GFAP)7; at 16 and 21 weeks, increase in smooth muscle actin was seen, thought to reflect arteriolar smooth muscle hyperplasia secondary to hypertension. SHRSP gene expression differences at 5 weeks of age were more numerous than at 16 or 21 weeks of age and included downregulation of Mmp14, Mbp, Gfap, Avp, Alb, and Igf2, upregulation of Gucy1a3, Rps9, Fos, and JunB, early-growth response, cell-signaling genes, and overexpression of genes involved in neurological diseases (stroke, depression, and blood–brain barrier leakage),8 rather than just hypertension. Recent gene sequencing of SHRSP rats (and 26 other rat models of common human diseases)9 revealed that genes that were either shared between or uniquely mutated in these rat models were significantly over-represented in human GWAS hits for hypertension or metabolism-related phenotypes, suggesting coevolution of these genes and their role in common diseases in models and humans.10

In a hypothesis-driven collaborative approach, we tested for associations between genes that were differentially expressed in the brains of 5-week-old SHRSP rats11 and WMH in humans. We used data from 5-week-old rats because gene expression differences were more frequent at that age than at 16 or 21 weeks, and we wanted to minimize the confounding of tissue changes by secondary effects of hypertension and to optimize the chances of detecting genes related to WMH susceptibility. We focused on WMH as the most frequent feature of SVD with the most data available in replication cohorts. We first tested the subjects from Lothian Birth Cohort 1936 (LBC1936)12,13 and then attempted replication in subjects from the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) consortium.7 To provide confidence in the relevance of subjects from LBC1936, we also sought CHARGE’s previously reported WMH-gene associations in the subjects from LBC1936.

Methods

Subjects
The subjects from LBC1936 are community-dwelling individuals living in South East Scotland who underwent detailed cognitive, biomedical, genetic assessments, and detailed brain MRI at ≈73 years of age (n=866).14,15 The MRI acquisition, methods for assessing WMH burden qualitatively and quantitatively,16,17 and proportions with WMH by either method have been reported. This study was approved by the Lothian (REC/07/MRE00/58) and Scottish Multicentre

Table 1. Genes Associated With Cerebral Small Vessel Disease in Rats That Are Associated With WMH in Older Humans: 126 Differentially Expressed Genes Between Spontaneously Hypertensive Stroke Prone and Wild-Type Rats Were Tested for Association With WMH in Subjects From LBC1936 and 10 Genes Were Significantly Associated (P<0.05) With Either WMH Volume or Fazekas Score

| Chromosome | Gene | Start Position | Stop Position | nSNPs | Discovery: LBC1936 | Replication: CHARGE |
|------------|------|----------------|---------------|-------|--------------------|---------------------|
|            |      | WMH Volume     | Fazekas Score |       | nSNPs P Value      | nSNPs P Value       |
|            |      |                |               |       |                    |                     |
| 4          | AFP  | 74520796       | 74540356      | 13    | 0.0021             | 0.00009             |
| 4          | ALB  | 74488869       | 74505834      | 11    | 0.0026             | 0.0017              |
| 7          | GNA1 | 79602075       | 79868661      | 42    | 0.034              | 0.033               |
| 1          | RBM6A| 144218994      | 144222801     | 13    | 0.038              | 0.057               |
| 2          | INPP5D| 233633279      | 233781288     | 69    | 0.041              | 0.78                |
| 10         | XPNPEP1| 111614513      | 111673192     | 18    | 0.042              | 0.14                |
| 9          | NIRA3| 101623957      | 101668994     | 13    | 0.045              | 0.16                |
| 13         | FARP1| 9795343        | 97900024      | 154   | 0.049              | 0.25                |
| 6          | MRPL18| 160131481      | 160139451     | 24    | 0.059              | 0.039               |
| 1          | SIPA1L2| 230600334      | 230717866     | 80    | 0.087              | 0.0093              |

nSNPs is the number of SNPs considered in the gene test. CHARGE indicates Cohorts for Heart and Aging Research in Genomic Epidemiology; LBC1936, Lothian Birth Cohort 1936; SNP, single-nucleotide polymorphism; and WMH, white matter hyperintensities.
The subjects from LBC1936 had genome-wide SNP data on 542,050 SNPs,21 imputed to 2.5 million SNPs with HapMap2.22 There were 621 participants (392 men) from LBC1936 with both MRI and genetic data (mean age, 72.67 years; SD=0.73 years; Table I and Methods in the online-only Data Supplement). We excluded 48 subjects from LBC1936 with a history of stroke or dementia.

**Gene Analysis**

In the 5-week-old SHRSP rats, 162 genes were differentially expressed compared with 5-week-old WKY rats in frontal and midcortical brain sections (Table II in the online-only Data Supplement).14 We used the following databases to match the SHRSP Illumina IDs to human genes (Materials and Table II in the online-only Data Supplement): Ensembl—http://www.ensembl.org, GeneCards—http://www.genecards.org, Illumina ID search—http://www.ncbi.nlm.com, NCBI—http://www.ncbi.nlm.nih.gov, and Rat Genome Database—http://www.rgd.mcw.edu. Of the 162 SHRSP genes, 132 had an equivalent human gene, 8 transcripts were mapped to the same gene, 20 were uncharacterized in humans, and 2 had no human homologue. Of the 132 genes, 126 were available for association testing using the Versatile Gene-based Association Study (VEGAS) test.23 We first performed a genome-wide association analysis on subjects from LBC1936 using PLINK software24 to test the genetic association between 542,050 genotyped SNPs and 2 WMH measurements using a linear regression analysis: (1) log transformed WMH volume (mL), with age, sex, intracranial volume, and first 4 multiple dimension scaling components for population stratification as covariates; and (2) summed Fazekas score of periventricular and deep WMH, with age, sex, and the first 4 multiple dimension scaling population stratification components as covariates. We used both WMH volume and Fazekas score20 to increase the reliability of the results. We did not stratify by vascular risk factors because hypertension (although it was the strongest vascular risk factor) explained <2% of WMH variance in subjects from LBC1936.20 The VEGAS software summarized evidence for association with WMH in subjects from LBC1936 per gene by considering the P values of all 543,050 SNPs that were located within 17681 unique autosomal genes (including SNPs±50 kb outside of genes to include regulatory regions). For a more direct comparison with CHARGE (which used imputed data), we also performed a gene-based test on LBC1936's 2447226 HapMap2 derived P values (after removing SNPs with a minor allele frequency of <0.01 and imputation quality of <0.3) with VEGAS software as above.

**Replication in Subjects From CHARGE**

We then tested whether any of the 126 SHRSP genes were also associated with WMH in subjects from CHARGE by using data from CHARGE's published genome-wide meta-analysis of WMH in 9361 stroke-free individuals from 7 community-based cohorts.25 We performed a gene-based test using VEGAS software, which summarized the evidence for association with WMH burden on a per gene basis, as above, by considering the associated P values of all HapMap2 SNPs located within 17787 autosomal genes (including SNPs±50 kb outside of genes to include regulatory regions).

**Gene Set Enrichment**

We performed a gene set enrichment analysis26 to investigate the enrichment of the 126 SHRSP genes in the LBC1936 and CHARGE data associated with WMH, accounting for whether these were upregulated or downregulated (online-only Data Supplement),26 corrected for multiple testing using a false discovery rate (FDR) method.27

**Replication of Previous CHARGE Findings in Subjects From LBC1936**

To demonstrate our ability to detect WMH-gene associations in subjects from LBC1936, we attempted replication of CHARGE's genome-wide associations with WMH28 in the subjects from the LBC1936 Cohort in a genome-wide association analysis using the 2534887 SNPs imputed to HapMap2, with WMH (volume and Fazekas score) in Mach2QTL software.29

We applied Bonferroni correction for multiple testing (P=0.05/126 genes=0.0004). We did not include the 2 WMH phenotypes in the Bonferroni correction as they are highly correlated (r=0.77). Because of the overconservative nature of Bonferroni correction for multiple testing,29 a nominal significance threshold of P value of <0.05 was required for replication efforts.

**Results**

**SHRSP Genes in Subjects From LBC1936**

Of the 126 candidate SHRSP-derived genes, 10 were nominally associated with WMH in subjects from LBC1936 (P<0.05; Table 1). Using imputed or genotyped data, 5 genes were associated with WMH volume (AFP, ALB, GNAI1 [RBMSA and INPP5D, both borderline]); 3 of these (AFP, ALB, and GNAI1) and 2 others (MRPL18 and SIPAI12) were associated with WMH Fazekas scores. Three other genes were associated with WMH volume using genotyped data only (XNPPPEP1, NNR4A3, and FARP1). None of these genes individually passed Bonferroni correction in subjects from LBC1936 (all were P>0.0004), in part, reflecting the LBC1936 sample size.

**SHRSP Genes in Subjects From CHARGE**

Two of these 10 genes were also associated with WMH in subjects from CHARGE (XNPPPEP1, P=6.7×10−5; and FARP1, P=0.024; Table 1). Full details of all 126 SHRSP to LBC1936 to CHARGE gene associations are given in Table III in the online-only Data Supplement. Several other of the 126 SHRSP genes (outside the 10/126 described above) showed significance at P<0.05 in subjects from CHARGE (eg, USMG5, MED17, ZNF461, C20orf7, EGR1, ARC, NUDT14, and MMP14) of which 1 (USMG5, P<0.000142) passed Bonferroni correction (P<0.0004).

**Gene Set Enrichment**

Using gene set enrichment analysis, all 126 SHRSP candidate genes were not enriched in subjects from LBC1936 for association with WMH in the 17681 genes tested here (WMH volume, P=0.34; Fazekas score, P=0.81), but this would not preclude the possibility that in either upregulated or downregulated gene sets, there was an abundance of genes showing an enriched association. We tested the upregulated (n=76) and downregulated (n=50) SHRSP genes separately and found significant enrichment for Fazekas scores in SHRSP downregulated genes (P=0.035; FDR, 0.046) but not SHRSP upregulated genes (P=0.921; FDR, 0.899). WMH volume showed significant enrichment in downregulated (P=0.018; FDR, 0.025) but not upregulated (P=0.802; FDR, 0.780) genes. In the CHARGE consortium, there was no significant enrichment for either the total set of 126 genes (P=0.0514), the upregulated (P=0.109; FDR, 0.266) or the downregulated genes (P=0.173; FDR, 0.149).

**Replication of CHARGE’s Previous Genome-Wide Association in Subjects From LBC1936**

We sought CHARGE’s previous genome-wide association results for WMH in subjects from LBC1936. Of CHARGE’s
15 SNPs ($P<1\times10^{-5}$) associated with WMH (Table 2). Of the 2 SHRSP genes found in LBC1936 and CHARGE, *XPNPEP1* is X-prolyl aminopeptidase (aminopeptidase P) 1, a soluble, associated with biliary atresia, and located in a region on chromosome 10 that is associated with Alzheimer's disease. *FARP1* is Pleckstrin domain protein 1, associated with brain volume differences, and important in synapse development. The SHRSP-CHARGE–associated gene *USMG5* is upregulated during skeletal muscle growth 5 homolog (also known as diabetes mellitus–associated protein in insulin sensitive tissues, or *DAPIT*), sits on chromosome 10, and maintains ATP synthase populations in mitochondria. All 5 SHRSP genes associated with both WMH volume and Fazekas score in subjects from LBC1936 (*AFP*, *ALB*, *GNA11*, *RBMSA*, and *MRPL18*) are associated with white matter–relevant diseases in humans. Despite not surviving correction for multiple testing, there was a notable consistency in their association with 2 separate WMH measures. *AFP* encodes α-fetoprotein, a major plasma protein produced in the yolk sac and liver during fetal life. Abnormally, high amounts of α-fetoprotein are found in ataxia telangiectasia, also associated with abnormal white matter. *ALB* encodes albumin, a soluble monomeric protein important for maintaining plasma oncotic pressure found in cerebral WMH, and cerebrospinal fluid as blood–brain barrier function deteriorates with ageing and dementia. *GNA11* encodes guanine nucleotide–binding protein (G protein), alpha-inhibiting activity polypeptide 1, implicated with Alzheimer's disease. *RBMSA* is an RNA binding protein that has differential expression in Alzheimer's disease, associations with a range of intellectual disabilities in humans and anxiety-related behavior in mice, with schizophrenia, several neurodevelopmental intellectual disabilities, anxiety behavior and may target neuronal genes to regulate behaviors. WMH in old age are known associates of late-onset depression, and they are also associated with lower age 11 IQ. *MRPL18* is the mitochondrial ribosomal protein L18, previously associated with development. The SHRSP-CHARGE–associated gene *USMG5* is upregulated during skeletal muscle growth 5 homolog (also known as diabetes mellitus–associated protein in insulin sensitive tissues, or *DAPIT*), sits on chromosome 10, and maintains ATP synthase populations in mitochondria. All 5 SHRSP genes associated with both WMH volume and Fazekas score in subjects from LBC1936 (*AFP*, *ALB*, *GNA11*, *RBMSA*, and *MRPL18*) are associated with white matter–relevant diseases in humans. Despite not surviving correction for multiple testing, there was a notable consistency in their association with 2 separate WMH measures. *AFP* encodes α-fetoprotein, a major plasma protein produced in the yolk sac and liver during fetal life. Abnormally, high amounts of α-fetoprotein are found in ataxia telangiectasia, also associated with abnormal white matter. *ALB* encodes albumin, a soluble monomeric protein important for maintaining plasma oncotic pressure found in cerebral WMH, and cerebrospinal fluid as blood–brain barrier function deteriorates with ageing and dementia. *GNA11* encodes guanine nucleotide–binding protein (G protein), alpha-inhibiting activity polypeptide 1, implicated with Alzheimer's disease. *RBMSA* is an RNA binding protein that has differential expression in Alzheimer's disease, associations with a range of intellectual disabilities in humans and anxiety-related behavior in mice, with schizophrenia, several neurodevelopmental intellectual disabilities, anxiety behavior and may target neuronal genes to regulate behaviors. WMH in old age are known associates of late-onset depression, and they are also associated with lower age 11 IQ. *MRPL18* is the mitochondrial ribosomal protein L18, previously associated with development.
with multiple sclerosis. These 7 SHRSP-derived genes are related to pathologies (ataxia telangiectasia, blood-brain barrier impairment, Alzheimer's disease, multiple sclerosis, depression, developmental intellectual disabilities, and brain size) that display white matter abnormalities or affect intellectual function. Impaired ATP production because of defects in USMG5, the gene that replicated from SHRSP to CHARGE, could increase susceptibility to WMH via ischemia.

The genes that were downregulated in the SHRSP were significantly enriched in subjects from LBC1936 for WMH. This may be because, in a complex disease such as SVD/WMH, several individually modest genetic defects in different components of key pathways, when present in combination, increase disease risk. This interpretation is consistent with differential protein expression seen in SHRSP and the absence, so far, of individual major human gene defects explaining either sporadic WMH or lacunar stroke.9

The lack of consistent replication from SHRSP to LBC1936 to CHARGE requires caution. The power and required significance threshold of the LBC1936 was modest for GWAS, hence our hypothesis-driven approach. Genes associated with WMH in subjects from LBC1936 but not CHARGE could be false positives; other factors include greater heterogeneity of WMH assessment and greater age range in subjects from CHARGE. The narrow age range of subjects from LBC1936 minimizes the effect of age, possibly helping to expose relevant genes. CHARGE-contributing studies used several methods of quantifying WMH, different MR scanner field strengths, and generations of technology and sequences. However, WMH volume and visual scores are highly correlated, and our replication of 3 findings from CHARGE in subjects from LBC1936 suggests that our approach has some validity. The CHARGE cohorts may have used different imputation platforms or more SNPS may have failed quality assurance in subjects from LBC1936, contributing to differences between the imputation results. There are several limitations to gene-based analysis, including the omission of nonautosomal genes, the effect of noncausal SNPs to dilute association (in particular, in the presence of a strong genetic association with a single locus within or in the regulatory region of a given gene, thus missing important associations), the lack of knowledge on (and overlap of) gene boundaries, the possibility that an SNP variant may influence a gene distal to its site, thus not corresponding to a gene that it is located next to it, and the potential of the genetic data not to tag causative genetic variants. Power may have been limited (despite CHARGE's large sample size) to detect associations with some genes. We did not stratify the human cohorts by risk factors as these explained <2% of WMH variance in subjects from LBC1936, and risk-stratified genetic data were unavailable for CHARGE. We did not test gene associations with other SVD features in addition to WMH because a total SVD burden score was not available for CHARGE. Although it is a relevant model of spontaneous SVD and of human hypertension and metabolic disorders, like any model, the SHRSP has translational limitations, arguing for additional studies at different ages and brain regions, with or without environmental stressors.

This work has the following strengths: accurate LBC1936 WMH phenotyping and genetic information in this relatively large narrow age-range older population. The Glasgow SHRSP colony is long established, with carefully controlled environments. The mRNA data were obtained from the same rats that provided protein expression data. Replication in other SHRSP colonies and examination of related strains (eg, SHR’s) may be informative. The genomes of SHRSP and 26 other complex disease phenotype models were recently sequenced, showing associations between genes in rat models of hypertension and human GWAS hits for hypertension phenotypes. This provides support for our reverse-translational discovery approach, suggesting that genes in disease models have coevolved and may contribute to disease-related phenotypes in humans.

Our findings require validation. The selection of candidate genes for investigation could be widened by examining more genes from the 5-week-old SHRSP rats (Table II in the online-only Data Supplement), other models, and in larger samples of well-phenotyped humans, such as from METASTROKE and the Wellcome Trust Case-Control Consortium. This translational analysis of experimental models and human disease suggests some aspects of the genetic architecture underlying SVD, stroke, and dementia and argues for greater awareness of vascular contributions to neurodegeneration.

Figure I and Tables IV and V in the online-only Data Supplement provide the top SNP (P<1x10^-5) and gene (P<0.001) associations with WMH variables in subjects from LBC1936 for further reference.

Acknowledgments
We thank the Lothian Birth Cohort 1936 participants and research team members, Wellcome Trust Clinical Research Facility (http://www.wtcrf.ed.ac.uk, subject testing and genotyping), and Brain Imaging Centre (http://www.bric.ed.ac.uk, brain imaging and analyses). Cohorts for Heart and Aging Research in Genomic Epidemiology thanks the staff and participants of the Aging Gene-Environment Susceptibility-Reykjavik Study, Atherosclerosis Risk in Community Study (ARIC), Austrian Stroke Prevention Study (ASPS), Cardiovascular Health Study, and Framingham Heart and Rotterdam Studies for their important contributions. ASPS thanks Birgit Reinhart for her long-term administrative commitment and Ing Johann Semmler for the technical assistance at creating the DNA bank. Drs Wardlaw, Bailey, McBride, Graham, Dominiczak, Deary, Starr, Seshadri, Fornage, Ikram, Debelette, Launer, Bis, and Schmidt contributed to data collection. Dr Lopez, Harris, Hill, Yang, Bailey, McClure, McBride, Smith, Hernandez, Maniega, Bastin, and Wardlaw contributed to data analysis. Drs Wardlaw, Deary, and Seshadri contributed to study design, co-ordination, and funding. Lopez, Wardlaw, Seshadri, and Deary contributed to article preparation. Lopez, Harris, Hill, Porteous, Smith, Deary, Starr, Seshadri, Yang, Fornage, Ikram, Debelette Launer, Bis, Schmidt, Bailey, McBride, Graham, McClure, Dominiczak, Hernandez, Maniega, Bastin, and Wardlaw contributed to article review. Wardlaw was the guarantor and provided the overall concept.

Sources of Funding
Lothian Birth Cohort 1936 was funded by Age UK’s Disconnected Mind programme (http://www.disconnectedmind.ed.ac.uk) and by Research Into Ageing (references 251 and 285). Whole-genome association was funded by Biotechnology and Biological Sciences Research Council (reference BB/F019394/1), brain image analysis was funded by Medical Research Council (G1001401 and 8200), and imaging was funded by Brain Research Imaging Centre (http://www.bric.ed.ac.uk), The University of Edinburgh Centre for Cognitive Ageing and Cognitive Epidemiology (http://www.
References

1. Debette S, Markus HS. The clinical importance of white matter hyper-intensities on brain magnetic resonance imaging: systematic review and meta-analysis. BMJ. 2010;341:c3666. doi: 10.1136/bmj.c3666.

2. Wardlaw JM, Smith C, Dichgans M. Mechanisms of sporadic cerebral small vessel disease: insights from neuroimaging. Lancet Neurol. 2013;12:483–497. doi: 10.1016/s1474-4422(13)70060-7.

3. Wardlaw JM, Allerhand M, Doublé FN, Valdés Hernández M, Morris Z, Gow AJ, et al. Vascular risk factors, large artery atheroma, and brain white matter hyperintensities. Neurology. 2014;82:1331–1338. doi: 10.1212/WNL.0000000000003132.

4. Kochunov P, Glahn D, Winkler A, Duggirala R, Olvera RL, Cole S, et al. Analysis of genetic variability and whole genome linkage of whole-brain, subcortical, and ependymal hypertensive white matter volume. Stroke. 2009;40:3685–3690. doi: 10.1161/STROKEAHA.109.565390.

5. Yamamoto Y, Crabbage B, Mannion M, Kalimo H, Kalaria RN. Review: molecular genetics and pathology of hereditary small vessel diseases of the brain. Neuropathol Appl Neurobiol. 2011;37:94–113. doi: 10.1111/j.1365-2990.2010.01147.x.

6. Jackson CA, Hutchison A, Dennis MS, Wardlaw JM, Lindgren A, Norrving B, et al. Differing risk factor profiles of ischemic stroke subtypes: evidence for a distinct lacunar arteriopathy? Stroke. 2010;41:642–629. doi: 10.1161/STROKEAHA.109.558809.

7. Fornage M, Debette S, Bis JC, Schmidt H, Ikram MA, Dufouil C, et al. Genome-wide association studies of cerebrovascular white matter lesion burden: the CHARGE consortium. Ann Neurol. 2011;69:926–939. doi: 10.1002/ana.22493.

8. Verhaaren BF, de Boer R, Vroomoo MW, Rivadeneira F, Uitterlinden AG, Hofman A, et al. Replication study of chr17q25 with cerebral white matter lesion volume. Stroke. 2011;42:3297–3299. doi: 10.1161/STROKEAHA.111.623090.

9. Traylor M, Farrall M, Holliday EG, Sudlow C, Hopewell JC, Cheng YC, et al. Australian Stroke Genetics Collaborative; Wellcome Trust Case Control Consortium 2 (WTCCC2); International Stroke Genetics Consortium. Genetic risk factors for ischaemic stroke and its subtypes (the METASTROKE collaboration): a meta-analysis of genome-wide association studies. Lancet Neurol. 2011;10:951–962. doi: 10.1016/S1474-4422(12)70234-X.

10. Adh-Samii P, Rost N, Traylor M, Devan W, Bifft A, Lanfranconsi S, et al. Australian Stroke Genetics Collaborative; Wellcome Trust Case Control Consortium-2 (WTCCC2); METASTROKE; International Stroke Genetics Consortium. 17q25 Locus is associated with white matter hyper-intensity volume in ischemic stroke, but not with lacunar stroke status. Stroke. 2013;44:1609–1615. doi: 10.1161/STROKEAHA.113.679396.

11. Bailey EL, McCulloch J, Sudlow C, Wardlaw JM. Potential animal models of lacunar stroke: a systematic review. Stroke. 2009;40:e451–e458. doi: 10.1161/STROKEAHA.108.528430.

12. Bailey EL, Smith C, Sudlow CM, Wardlaw JM. Is the spontaneously hypertensive stroke prone rat a pertinent model of subcortical ischemic stroke? A systematic review. Int J Stroke. 2011;6:434–444. doi: 10.1111/j.1747-4949.2011.00659.x.

13. Bailey EL, Wardlaw JM, Graham D, Dominiczak AF, Sudlow CL, Smith C. Cerebral small vessel endothelial structural changes predate hyperten-sion in stroke-prone spontaneously hypertensive rats: a blinded, controlled immunohistochemical study of 5- to 21-week-old rats. Neuropathol Appl Neurobiol. 2011;37:711–726. doi: 10.1111/j.1365-2990.2011.01170.x.

14. Bailey EL, McBride MW, Beattie W, McClure JD, Graham D, Dominiczak AF, et al. Differential gene expression in multiple neurological, inflammatory and connective tissue pathways in a spontaneous model of human small vessel stroke. Neuropathol Appl Neurobiol. 2014;40:855–872. doi: 10.1111/nan.12116.

15. Atanur SS, Diaz AG, Maratou K, Sarkis A, Rotival M, Game L, et al. Genome sequencing reveals loci under artificial selection that underlie disease phenotypes in the laboratory rat. Cell. 2013;154:691–703. doi: 10.1016/j.cell.2013.06.040.

16. Deary IJ, Gow AJ, Taylor MD, Corley J, Brett C, Wilson V, et al. The Lothian Birth Cohort 1936: a study to examine influences on cognitive ageing from age 11 to age 70 and beyond. BMC Geriatr. 2007;7:28. doi: 10.1186/1471-2318-7-28.

17. Wardlaw JM, Bastin ME, Valdés Hernández MC, Manienga SM, Royle NA, Morris Z, et al. Brain aging, cognition in youth and old age and vascular disease in the Lothian Birth Cohort 1936: rationale, design and methodology of the imaging protocol. Int J Stroke. 2011;6:547–559. doi: 10.1111/j.1747-4949.2011.00683.x.

18. Fickis P, Chwialk JB, Alavi A, Hurtig HI, Zimmerman RA. MR signal abnormalities at 1.5 T in Alzheimer’s dementia and normal aging. AJR Am J Roentgenol. 1987;149:351–356. doi: 10.2214/ajr.149.2.3531.

19. Hernández Mdel C, Ferguson KJ, Chappell FM, Wardlaw JM. New multispectral MRI data fusion technique for white matter lesion segmentation: method and comparison with thresholding in FLAIR images. Eur Radiol. 2010;20:1684–1691. doi: 10.1007/s00330-010-1718-6.
20. Valdés Hernández Mdcl C, Morris Z, Dickie DA, Royle NA, Muñoz Maniega S, Ariasala BS, et al. Close correlation between quantitative and qualitative assessments of white matter lesions. Neuroepidemiology. 2013;40:13–22. doi: 10.1159/000341859.

21. Houlihan LM, Davies G, Tenesa A, Harris SE, Luciano M, Gow AJ, et al. Common variants of large effect in F12, KNG1, and HRG are associated with activated partial thromboplastin time. Am J Hum Genet. 2010;86:626–631. doi: 10.1016/j.ajhg.2010.02.016.

22. Wain LV, Verwoert GC, O’Reilly PF, Shi G, Johnson T, Johnson AD, et al; LifeLines Cohort Study; EchoGen consortium; AortaGen Consortium; CHARGE Consortium Heart Failure Working Group; KidneyGen consortium; CKDGen consortium; CardioGen consortium; CardioGram. Genome-wide association study identifies six new loci influencing pulse pressure and mean arterial pressure. Nat Genet. 2011;43:1005–1011. doi: 10.1038/ng.922.

23. Corley C, Martinsson H, S appeals to the public and policy makers to support systemic approaches in Alzheimer’s disease research. Sci Transl Med. 2012;4:e341. doi: 10.1038/tp.2013.114.

24. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007;81:559–575. doi: 10.1086/519795.

25. Hill WD, Davies G, van de Lagemaat LN, Christoforou A, Marioni RE, Fernandes CP, et al. Human cognitive ability is influenced by genetic variation in components of postsynaptic signalling complexes assembled by NMDA receptors and MAGUK proteins. Transl Psychiatry. 2014;4:e341. doi: 10.1038/tp.2013.114.

26. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. 2005;102:15545–15550. doi: 10.1073/pnas.0506501102.

27. Wang K, Li M, Bucan M. Pathway-based approaches for analysis of genomewide association studies. Am J Hum Genet. 2007;81:1278–1283. doi: 10.1086/523274.

28. Li Y, Willer CJ, Ding J, Scheet P, Abecasis GR. MaCH: using sequence and genotype data to estimate haplotypes and unobserved genotypes. Genet Epidemiol. 2010;34:816–834. doi: 10.1002/gepi.20533.

29. Perneger TV. What’s wrong with Bonferroni adjustments. BMJ. 1998;316:1236–1238. doi: 10.1136/ bmj.316.7139.1236.

30. Grube A, Li Y, Rowland C, Nowotny P, Hinrichs AL, Smemo S, et al. A scan of chromosome 10 identifies a novel locus showing strong association with late-onset Alzheimer disease. Am J Hum Genet. 2006;78:78–88. doi: 10.1086/498851.

31. Stein JL, Hua X, Lee S, Ho AJ, Leow AD, Toga AW, et al. Alzheimer’s Disease Neuroimaging Initiative. Voxelwise genome-wide association study (vGWAS). Neuroimage. 2010;53:1174–1174. doi: 10.1016/j.neuroimage.2010.02.032.

32. Cheadle L, Biederer T. The novel synaptogenic protein Farp1 links postsynaptic cytoskeletal dynamics and transsynaptic organization. J Cell Biol. 2012;199:985–1001. doi: 10.1083/jcb.201205041.

33. Ohsaka S, Fujikawa M, Hisabori T, Yoshida M. Knockdown of DAPIT (diabetes-associated protein in insulin-sensitive tissue) results in loss of ATP synthase in mitochondria. J Biol Chem. 2011;286:20292–20296. doi: 10.1074/jbc.M110.198523.

34. Waldmann TA, McIntire KR. Serum-alpha-fetoprotein levels in patients with ataxia-telangiectasia. Lancet. 1972;2:1112–1115. doi: 10.1186/s1440-6736(72)92717-1.

35. Cierniak JJ, Horowitz AL. Abnormal white matter signal in ataxia telangiectasia. AJNR Am J Neuroradiol. 2000;21:1483–1485.

36. Grinberg LT, Thal DR. Vascular pathology in the aged human brain. Acta Neuropathol. 2010;119:277–290. doi: 10.1007/s00401-010-0652-7.

37. Farrall AJ, Wardlaw JM. Blood-brain barrier: ageing and microvascular disease—systematic review and meta-analysis. Neurobiol Aging. 2009;30:337–352. doi: 10.1016/j.neurobiolaging.2007.07.015.

38. Silver M, Janousova E, Hua X, Thompson PM, Montana G; Alzheimer’s Disease Neuroimaging Initiative. Identification of gene pathways implicated in Alzheimer’s disease using longitudinal imaging phenotypes with sparse regression. Neuroimage. 2012;63:1681–1694. doi: 10.1016/j.neuroimage.2012.08.002.

39. Wong J. Altered expression of RNA splicing proteins in Alzheimer’s disease patients: evidence from two microarray studies. Dement Geriatr Cogn Dis Extra. 2013;3:74–85. doi: 10.1159/000348406.

40. Alachkar A, Jiang D, Harrison M, Zhou Y, Chen G, Mao Y; An EJC factor RBM18a regulates anxiety behaviors. Curr Mol Med. 2013;13:887–899. doi: 10.2174/156652401313990019.

41. Herrmann LL, Le Masurier M, Ebnemer KP. White matter hyperintensities in late life depression: a systematic review. J Neurol Neurosurg Psychiatry. 2008;79:619–624. doi: 10.1136/jnnp.2007.124651.

42. Valdés Hernández Mdcl C, Booth T, Murray C, Gow AJ, Penke L, Morris Z, et al. Brain white matter damage in aging and cognitive ability in youth and older age. Neurobiol Aging. 2013;34:2740–2747. doi: 10.1016/j.neurobiolaging.2013.05.032.

43. Fischer MT, Sharma R, Lim JL, Haider L, Frischer JM, Drekhage J, et al. NADPH oxidase expression in active multiple sclerosis lesions in relation to oxidative tissue damage and mitochondrial injury. Brain. 2012;135(pt 3):886–899. doi: 10.1093/brain/awt012.
Genes From a Translational Analysis Support a Multifactorial Nature of White Matter Hyperintensities

Lorna M. Lopez, W. David Hill, Sarah E. Harris, Maria Valdes Hernandez, Susana Munoz Maniega, Mark E. Bastin, Emma Bailey, Colin Smith, Martin McBride, John McClure, Delyth Graham, Anna Dominiczak, Qiong Yang, Myriam Fornage, M. Arfan Ikram, Stephanie Debette, Lenore Launer, Joshua C. Bis, Reinhold Schmidt, Sudha Seshadri, David J. Porteous, John Starr, Ian J. Deary and Joanna M. Wardlaw

Stroke. 2015;46:341-347; originally published online January 13, 2015;
doi: 10.1161/STROKEAHA.114.007649

Stroke is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2015 American Heart Association, Inc. All rights reserved.
Print ISSN: 0039-2499. Online ISSN: 1524-4628

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://stroke.ahajournals.org/content/46/2/341
Free via Open Access

Data Supplement (unedited) at:
http://stroke.ahajournals.org/content/suppl/2015/01/30/STROKEAHA.114.007649.DC1

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Stroke can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to Stroke is online at:
http://stroke.ahajournals.org/subscriptions/
SUPPLEMENTAL MATERIAL

Genes from a translational analysis support a multifactorial nature of white matter hyperintensities
Lorna M. Lopez, PhD1, W. David Hill, MSc1, Sarah E. Harris, PhD2, Maria Valdez Hernandez, PhD1,3 Susana Munoz Maniega, PhD1,3 Mark E Bastin, PhD1,3 Emma Bailey, PhD4 Colin Smith, MD5, Martin McBride, PhD6, John McClure, PhD6, Delyth Graham, PhD6, Anna Dominiczak, MD6, Qiong Yang, PhD7,8, Myriam Fornage, PhD9, M. Arfan Ikram, MD, PhD10-12, Stephanie Debette, MD, PhD13-16 Lenore Launer, PhD17, Joshua C. Bis, PhD18 Reinhold Schmidt, MD19, Sudha Seshadri, MD8,15, David J. Porteous, PhD2, John Starr, MD1, Ian J. Deary, PhD1, Joanna M. Wardlaw, MD3*.

Index
Supplementary Methods
Supplementary tables and figures
Table I: Description of LBC1936 white matter hyperintensity (WMH) variables.
Table II: Candidate genes from SHRSP rat model. This is a list of the 162 transcripts differentially expressed between SHRSP and WKY at 5 weeks in two relevant brain regions, and the corresponding human genes where available.
Table III: Candidate gene-based association results with white matter hyperintensity variables (WMH) in the discovery cohort, LBC1936, and the replication cohort, CHARGE.
Table IV: 5 Top hits from genome wide association study with WMH variables in LBC1936 (P < 1x10-5).
Table V: Top gene-based results from LBC1936 for the WMH variables run in Vegas.
Figure I: Genome-wide association study results of WMH (a) and Fazekas score (b) on 542,050 SNPs. QQ and Manhattan plots are shown.
Supplementary Methods

Genotyping
A detailed description of the genotyping method is described elsewhere. Briefly, genotyping was performed using Illumina Human 610-Quadv1 arrays on blood-extracted DNA at the WTCRF Genetics Core. All individuals were checked for disagreement between genetic and reported gender. Relatedness between subjects was investigated and, for any related pair of individuals, one was removed. Samples with a call rate \( \leq 0.95 \), and those showing evidence of non-Caucasian ascent by multidimensional scaling (MDS), were also removed. SNPs were included in the analyses if they met the following conditions: call rate \( \geq 0.98 \), minor allele frequency \( \geq 0.01 \), and Hardy-Weinberg Equilibrium test with \( P \geq 0.001 \). The final number of genotyped SNPs included in the study was 542,050 in 1,005 individuals.

Genetic imputation
~2.5M common SNPs included in HapMap, using the HapMap phase II CEU data as the reference sample were imputed. NCBI build 36 (UCSC hg18) was used and genotype data were imputed using MACH software. Prior to imputation SNPs were removed that diverged from HWE with a significance \( p < 1 \times 10^{-3} \) and SNPs with a minor allele frequency < 0.01.

Gene mapping.
Listed below are the databases used to match the Illumina IDs from the SHRSP study to human genes as shown in Supplementary Table II.
Ensembl – [www.ensembl.org](http://www.ensembl.org), GeneCards – [www.genecards.org](http://www.genecards.org), Illumina ID search - [www.genscript.com](http://www.genscript.com), NCBI - [http://www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov), Rat Genome Database - [rgd.mcw.edu](http://rgd.mcw.edu).

Gene set enrichment analysis
A gene set enrichment analysis was performed to investigate the enrichment of 126 SHRSP genes in WMH gene associations. First, the gene based statistics from VEGAS were rank ordered before being \( -\log(10) \) transformed. Gene set enrichment analysis (GSEA) uses a set of candidate gene identifiers and a genome wide set of genes, ranked based on their association with a phenotype. Next, a weighted Kolmogorov-Smirnov type statistic, walks down the genome wide ranked set of genes and increases the test statistic each time it finds a gene that matches one from the candidate gene set and decreases it when it does not.\(^2\), \(^3\) The magnitude of the increase is proportional to its p value, allowing for information regarding rank and distance between ranks to be used in the calculation of enrichment. The maximum deviation from zero is assigned to the candidate gene set (this is the enrichment score or ES). The gene set is then permuted before the ES being re-calculated. The p-value describes the proportion of the 5,000 permuted enrichment scores that the observed enrichment score was greater than.
Table I: Description of LBC1936 WMH variables. Fazekas scale for periventricular lesions, Fazekas scale for deep lesions and the sum of these Fazekas scores are described. WMH as percentage of true WMH alone in intracranial volume (ICV) are listed.

| Trait          | Median | Mean   | Standard deviation | Minimum | Maximum |
|----------------|--------|--------|--------------------|---------|---------|
| Age in years   | 72.72  | 72.67  | 0.73               | 71.04   | 74.22   |
| Fazekas Peri   | 1      | 1.34   | 0.635              | 0       | 3       |
| Fazekas Deep   | 1      | 1.08   | 0.655              | 0       | 3       |
| Fazekas Sum    | 2      | 2.42   | 1.12               | 0       | 6       |
| ICV mm$^3$     | 1,448,490 | 1,452,235.57 | 141,870.37     | 1,059,966 | 1,876,420 |
| WMH volume mm$^3$ | 7,554  | 11,885.97 | 12,826.30        | 0       | 98,378  |
| WMH transformed| 2.1464 | 2.1135 | 0.9920             | 0       | 4.60    |
| WMH in ICV     | 0.5245 | 0.8203 | 0.8999             | 0       | 7.47    |
Table II: Candidate genes from SHRSP rat model. This is a list of the 162 transcripts differentially expressed between SHRSP and WKY at 5 weeks in two relevant brain regions from Bailey et al. and the corresponding human genes.

**SHRSP versus WKY differential gene expression at five weeks of age.**

| PROBE_ID      | ILMN_GENE            | CHROMOSOME | WKY vs SHRSP frontal (FDR) | WKY vs SHRSP mid-coronal (FDR) |
|---------------|----------------------|------------|---------------------------|---------------------------------|
| ILMN_1365113  | RGD1564649_PREDICTED |            | 0                         | 0                               |
| ILMN_2038795  | RP59                 |            | 1                         | 0                               |
| ILMN_1371357  | LOC497757            |            |                           | 0                               |
| ILMN_2038796  | RP59                 |            | 1                         | 0                               |
| ILMN_1359040  | RGD1561110_PREDICTED |            | 2                         | 0                               |
| ILMN_1368735  | RGD1311103_PREDICTED |            |                           | 0                               |
| ILMN_2039673  | ARC                  | 7          | 0                         | 0                               |
| ILMN_1361865  | ZFSP597              | 10         | 0                         | 0                               |
| ILMN_1372230  | RNF149               | 9          | 4.13E-05                  | 0                               |
| ILMN_1350784  | JUNB                 | 19         | 0.00010101                | 0.00030303                      |
| ILMN_1351340  | LOC500950            |            | 7.27E-05                  | 0.000187166                     |
| ILMN_1368356  | FOS                  | 6          | 0.000117302               | 0.00030303                      |
| ILMN_1367486  | DUSP1                | 10         | 0.00010101                | 0                               |
| ILMN_1371004  | RP516                |            | 0.000117302               | 0                               |
| ILMN_1367530  | LOC497727            |            | 0.000160428               | 0.000146628                     |
| ILMN_1372711  | LOC502316            |            | 0.000160428               | 0.000146628                     |
| ILMN_1367467  | PER2                 | 9          | 0.000383838               | 0.000465632                     |
| ILMN_1367162  | GPM6A                | 16         | 0.000317125               | 0.000383634                     |
| ILMN_1352667  | NAB1                 |            | 0.000317125               | 0.000890538                     |
| ILMN_1366713  | ZNF575_PREDICTED     |            | 0.000227273               | 0.001603306                     |
| ILMN_1353766  | RGD1566136_PREDICTED | X          | 0.000317125               | 8.26E-05                        |
| ILMN_1353839  | PER1                 |            | 0.000317125               | 0.00030303                      |
| ILMN_1360786  | FKBP8                | 16         | 0.000317125               | 0.000125392                     |
| ILMN_1363342  | SLC1A3               | 2          | 0.000618182               | 0.001018182                     |

*** Human equivalent of SHRSP differentially expressed genes

| Rat Gene Symbol | Human Gene | Human Gene Symbol, Human Gene Description | Notes |
|-----------------|------------|-------------------------------------------|-------|
| Rps9            | RP59       | Homo sapiens ribosomal protein 59 (RP59), mRNA. |      |
| Rps9            | RP59       | As above (GUCY1A3, transcript variant 1, mRNA). |      |
| Gucy1a3         | GUCY1A3    | Homo sapiens guanylate cyclase 1, soluble, alpha 3 (GUCY1A3), transcript variant 1, mRNA. |      |
| Rps9            | RP59       | As above (FAM151B, member B [FAM151B], mRNA). |      |
| Fam151b         | FAM151B    | Homo sapiens family with sequence similarity 151, member B. |      |
| Arc             | ARC        | Homo sapiens activity-regulated cytoskeleton-associated protein (ARC), mRNA. |      |
| Zfp597          | ZNF597     | Homo sapiens zinc finger protein 597 (ZNF597), mRNA. |      |
| Rnf149          | RNF149     | Homo sapiens ring finger protein 149 (RNF149), mRNA. |      |
| JUNB            | Junb       | Homo sapiens jun B proto-oncogene (JUNB), mRNA. |      |
| Zfp317          | ZNF317     | Homo sapiens zinc finger protein 317 (ZNF317), transcript variant 1, mRNA. |      |
| FOS             | FOs        | Homo sapiens FB1 murine osteosarcoma viral oncogene homolog (FOs), mRNA. |      |
| Dusp1           | DUSP1      | Homo sapiens dual specificity phosphatase 1 (DUSP1), mRNA. |      |
| Rps16           | RPS16      | Homo sapiens ribosomal protein 516 (RPS16), mRNA. |      |
| Sipa112         | Sipa112    | Homo sapiens signal-induced proliferation-associated 1 like 2 (Sipa112), mRNA. |      |
| Zfp566          | ZNF566     | Homo sapiens zinc finger protein 566 (ZNF566), transcript variant 1, mRNA. |      |
| Per2            | PER2       | Homo sapiens period homolog 2 (Drosophila) (PER2), mRNA. |      |
| Gpm6a           | GPM6A      | Homo sapiens glycoprotein M6A (GPM6A), transcript variant 1, mRNA. |      |
| Nab1            | NAB1       | Homo sapiens NGFI-A binding protein 1 (EGR1 binding protein 1) (NAB1), mRNA. |      |
| Zfp575          | ZNF575     | Homo sapiens zinc finger protein 575 (ZNF575), mRNA. |      |
| Rps9            | RP59       | Homo sapiens period homolog 1 (Drosophila) (PER1), mRNA. |      |
| Fkbp8           | FKBP8      | Homo sapiens FK506 binding protein 8, 38kDa (FKBP8), mRNA. |      |
| SCl1A3          | SLC1A3     | Homo sapiens solute carrier family 1 (glial high affinity glutamate transporter), member 3 (SLC1A3), transcript variant 1, mRNA. |      |

**Notes:**
- OOE is an ortholog.
- Similar to zinc finger protein 75.
- Hypothetical protein.
- Similar to 40S ribosomal protein 59.
| GenBank ID | LOC | FDR P | q-Value | Description                                                                 |
|------------|-----|-------|---------|-----------------------------------------------------------------------------|
| ILMN_1359704 | LOC307332 | 0.000371901 | 0.00161442 | **SHRSP versus WKY differential gene expression at five weeks of age.** |
| ILMN_1364120 | POL | 1 | 0.000714286 | 0.000847107 | POL Homo sapiens polymerase (DNA directed), lambda (POL), transcript variant 1, mRNA. |
| ILMN_1362834 | DUSP6 | 0.01124807 | 0.01020475 | **Human equivalent of SHRSP differentially expressed genes** |
| ILMN_1369573 | LOC688712 | 0.01792208 | 0.006753247 | Rnf149 RNF149 As above similar to goliath-related E3 ubiquitin ligase 4 |
| ILMN_1375922 | NR4A3 | 0.000482375 | 0.007217069 | **Human equivalent of SHRSP differentially expressed genes** |
| ILMN_1349793 | LOC684139 | 0.000651801 | 0.002306649 | **Human equivalent of SHRSP differentially expressed genes** |
| ILMN_1358205 | RGDP1560975_PREDICTED | 0.008090404 | 0.028941878 | **Human equivalent of SHRSP differentially expressed genes** |
| ILMN_1360210 | LOC499068 | 0.001305738 | 0.01180303 | **Human equivalent of SHRSP differentially expressed genes** |
| ILMN_1376530 | RT1-A3 | 0.000861244 | 0.0119697 | **Human equivalent of SHRSP differentially expressed genes** |
| ILMN_1368305 | LOC499613 | 0.00124807 | 0.00277778 | **Human equivalent of SHRSP differentially expressed genes** |
| ILMN_1373217 | ADPGK | 0.001633729 | 0.000847107 | **Human equivalent of SHRSP differentially expressed genes** |
| ILMN_1355124 | GALNT2_PREDICTED | 0.000714286 | 0.008354978 | **Human equivalent of SHRSP differentially expressed genes** |
| ILMN_1368809 | BTG2 | 0.01057352 | 0.01379231 | **Human equivalent of SHRSP differentially expressed genes** |
| ILMN_1364113 | CTGF | 0.002249417 | 0.004427391 | **Human equivalent of SHRSP differentially expressed genes** |
| ILMN_1367428 | ZFP189_PREDICTED | 0.004554637 | 0.013664773 | **Human equivalent of SHRSP differentially expressed genes** |
| ILMN_1370045 | TRAPPC2 | 0.0002683983 | 0.008354978 | **Human equivalent of SHRSP differentially expressed genes** |
| ILMN_1364821 | LOC500720 | 0.009307057 | 0.028941878 | **Human equivalent of SHRSP differentially expressed genes** |
| ILMN_1350533 | RGDP1563551_PREDICTED | 0.001792118 | 0.00016442 | **Human equivalent of SHRSP differentially expressed genes** |
| ILMN_1367827 | LOC298998 | 0.003996212 | 0.005117845 | **Human equivalent of SHRSP differentially expressed genes** |
| ILMN_1352722 | LOC316550 | 0.003839548 | 0.005268474 | **Human equivalent of SHRSP differentially expressed genes** |
| ILMN_1349422 | PTGS2 | 0.004818182 | 0.016441558 | **Human equivalent of SHRSP differentially expressed genes** |
| ILMN_1359441 | PLA2G2A | 0.000950187 | 0.024345238 | **Human equivalent of SHRSP differentially expressed genes** |
| ILMN_1357461 | ZFP61 | 1 | 0.000237998 | 0.002540107 | **Human equivalent of SHRSP differentially expressed genes** |
### SHRS versus WKY differential gene expression at five weeks of age.

| Gene ID         | Accession | Fold Change | |  |
|-----------------|-----------|-------------|-----------------|-----------------|
| ILMN_1368780    | CLIC2     | 20          | 0.004115584     | 0.003713188     |
| ILMN_1359630    | LOC679663 | 16/NW_047479.1 | 0.008343109     | 0.004530577     |
| ILMN_1372466    | ZCCHC9    | 2           | 0.008783107     | 0.001607143     |
| ILMN_1353304    | RGD1561287_Predicted | 20 | 0.009307057 | 0.004159402 |
| ILMN_1374612    | TM9SF4_Predicted | 5 | 0.012253444 | 0.021175449 |
| ILMN_1368506    | LOC497770 | 2           | 0.023521336     | 0.054842609     |
| ILMN_1362561    | LOC498378 | 2           | 0.009918495     | 0.02418476      |
| ILMN_1359795    | LOC499555 | 16           | 0.01039312      | 0.046209617     |
| ILMN_1356628    | NFKBIA    | 6           | 0.006464646     | 0.013769231     |
| ILMN_1374199    | GIOT1     | 7           | 0.009307057     | 0.004530577     |
| ILMN_1353935    | FARP1_Predicted | 5 | 0.012253444 | 0.021175449 |
| ILMN_1362409    | BAI2_Predicted | 2           | 0.009307057     | 0.001607143     |
| ILMN_1370369    | EGR2      | 20          | 0.009918495     | 0.02418476      |
| ILMN_1372919    | CYR61     | 2           | 0.009918495     | 0.02418476      |
| ILMN_1355401    | RGD1563543_Predicted | 10 | 0.011252914 | 0.030909091 |
| ILMN_1362451    | RGS2      | 13          | 0.007420147     | 0.013769231     |
| ILMN_1365095    | DHX40     | 10          | 0.008343109     | 0.03276328      |
| ILMN_1376434    | PGRMC1    | X           | 0.009918495     | 0.02418476      |
| ILMN_1349269    | SGK       | 1           | 0.011252914     | 0.030909091     |
| ILMN_1361423    | RKHD2_Predicted | 6 | 0.01487781 | 0.042803325 |
| ILMN_1649981    | STRN3     | 6           | 0.015932282     | 0.033140909     |
| ILMN_1361339    | TMPRSS8   | 10          | 0.03056229      | 0.030909091     |
| ILMN_1357368    | LOC497841 | 20          | 0.01375383      | 0.06753247      |
| ILMN_1356949    | COL6A1_Predicted | 20 | 0.017802335 | 0.015773059 |
| ILMN_1368116    | LOC367398 | 2           | 0.004115584     | 0.003713188     |

### Human equivalent of SHRS differentially expressed genes

**Clic2** Homo sapiens chloride intracellular channel 2 (CLIC2), mRNA.  
**Tceb1** Homo sapiens transcription elongation factor B (SIII), polypeptide 1 (SIII, elongin C) (TCEB1), transcript variant 1, mRNA.  
**Zcch9** Homo sapiens zinc finger, CCHC domain containing 9 (ZCCHC9), transcript variant 1, mRNA.  
**Fam32a** Homo sapiens family with sequence similarity 32, member A (FAM32A), mRNA.  
**Tm9sf4** Homo sapiens transmembrane 9 superfamily protein member 4 (TM9SF4), mRNA.  
**Scn3a** Homo sapiens sodium channel, voltage-gated, type III, alpha subunit (SCN3A), transcript variant 1, mRNA.  
**Nfkbia** Homo sapiens nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, alpha (NFKBIA), mRNA.  
**Farp1** Homo sapiens FERM, RhoGEF (ARHGEF) and pleckstrin domain protein 1 (chondrocyte-derived) (FARP1), transcript variant 1, mRNA.  
**Bai2** Homo sapiens brain-specific angiogenesis inhibitor 2 (BAI2), mRNA.  
**Egr2** Homo sapiens early growth response 2 (EGR2), transcript variant 1, mRNA.  
**Cyr61** Homo sapiens cysteine-rich, angiogenic inducer, 61 (CYP61), mRNA.  
**Rpl31** Homo sapiens regulator of G-protein signaling 2, 24kDa (RGS2), mRNA.  
**Dhx40** Homo sapiens DEAH (Asp-Glu-Ala-His) box polypeptide 40 (DHX40), transcript variant 1, mRNA.  
**Pgrmc1** Homo sapiens progesterone receptor membrane component 1 (PGRMC1), mRNA.  
**Sgk1** Homo sapiens serum/glucocorticoid regulated kinase 1 (SGK1), transcript variant 1, mRNA.  
**Me3c** Homo sapiens mex-3 homolog C (C. elegans) (ME3C), mRNA.  
**Strn3** Homo sapiens striatin, calmodulin binding protein 3 (STRN3), transcript variant 1, mRNA.  
**Prss30p** Homo sapiens protease, serine, 30 homolog (mouse), pseudogene (PRSS30P), non-coding RNA.  
**Col6a1** Homo sapiens collagen, type VI, alpha 1 (COL6A1), mRNA.  
**Rpl17** Homo sapiens ribosomal protein L17 (RPL17), transcript variant 1, mRNA.
**SHRSP versus WKY differential gene expression at five weeks of age.**

| Accession | Gene Symbol | Gene Name | Human Gene Name | Human Gene Description |
|-----------|-------------|-----------|-----------------|------------------------|
| ILMN_1359375 | LOC499418 | Spatc1l | C21orf56 | Homo sapiens chromosome 21 open reading frame 56 (C21orf56), transcript variant 1, mRNA. |
| ILMN_1352529 | IER2 | Ier2 | IER2 | Homo sapiens immediate early response 2 (IER2), mRNA. |
| ILMN_1354120 | LOC691762 | C21orf56 | NECAB2 | Homo sapiens N-terminal EF-hand calcium binding protein 2 (NECAB2), mRNA. |
| ILMN_1373383 | TIPARP_PREDICTED | TIPARP | NECAB2 | Homo sapiens ring finger protein 40 (RNF40), transcript variant 1, mRNA. |
| ILMN_1364683 | NECAB2 | NECAB2 | NECAB2 | Homo sapiens N-terminal EF-hand calcium binding protein 2 (NECAB2), mRNA. |
| ILMN_1360868 | RNF40 | RNF40 | RNF40 | Homo sapiens ring finger protein 40 (RNF40), transcript variant 1, mRNA. |
| ILMN_1373434 | Rab28 | Rab28 | Rab28 | Homo sapiens ring finger protein 40 (RNF40), transcript variant 1, mRNA. |
| ILMN_1354535 | ZNF386 | ZFP368 | C1GALT1C1 | Homo sapiens C1GALT1-specific chaperone 1 (C1GALT1C1), transcript variant 1, mRNA. |
| ILMN_1362029 | LOC502490 | C1GALT1C1 | C1GALT1C1 | Homo sapiens C1GALT1-specific chaperone 1 (C1GALT1C1), transcript variant 1, mRNA. |
| ILMN_1351127 | PLCL1 | Plcl1 | PLCL1 | Homo sapiens phospholipase C-like 1 (PLCL1), mRNA. |
| ILMN_1361932 | MLL5 | Mll5 | MLL5 | Homo sapiens myeloid/lymphoid or mixed-lineage leukemia 5 (MLL5), transcript variant 1, mRNA. |
| ILMN_1369005 | EGR1 | Egr1 | EGR1 | Homo sapiens early growth response 1 (EGR1), mRNA. |
| ILMN_1360758 | RGD1308626 | Slain1 | SLAIN1 | Homo sapiens SLAIN motif family, member 1 (SLAIN1), transcript variant 1, mRNA. |
| ILMN_1359043 | EGR4 | Egr4 | EGR4 | Homo sapiens early growth response 4 (EGR4), mRNA. |
| ILMN_1368493 | RGD1562629_PREDICTED | Ftdc | FTCD | Homo sapiens formiminotransferase cyclodeaminase (FTCD), transcript variant A, mRNA. |
| ILMN_1349772 | FTCD | Ftdc | FTCD | Homo sapiens formiminotransferase cyclodeaminase (FTCD), transcript variant A, mRNA. |
| ILMN_1361370 | RYBP_PREDICTED | RYBP | RYBP | Homo sapiens RING1 and YY1 binding protein (RYBP), mRNA. |
| ILMN_1373798 | PLCB1 | Plcb1 | PLCB1 | Homo sapiens phospholipase C, beta 1 (phosphoinositide-specific) (PLCB1), transcript variant 1, mRNA. |
| ILMN_1372236 | ZFP36L1 | Zfp36l1 | ZFP36L1 | Homo sapiens zinc finger protein 36, C3H type-like 1 (ZFP36L1), transcript variant 1, mRNA. |
| ILMN_1349546 | MDGA2 | Mdga2 | MDGA2 | Homo sapiens MAM domain containing glycosylphosphatidylinositol anchor 2 (MDGA2), transcript variant 1, mRNA. |
| ILMN_1353576 | GPR149 | Gpr149 | GPR149 | Homo sapiens G protein-coupled receptor 149 (GPR149), mRNA. |

*** Human equivalent of SHRSP differentially expressed genes

| Gene Symbol | Gene Name | Gene Description |
|-------------|-----------|-----------------|
| Spatc1l | C21orf56 | Homo sapiens chromosome 21 open reading frame 56 (C21orf56), transcript variant 1, mRNA. |
| Ier2 | IER2 | Homo sapiens immediate early response 2 (IER2), mRNA. |
| C21orf56 | C21orf56 | Homo sapiens chromosome 21 open reading frame 56 (C21orf56), transcript variant 1, mRNA. |
| NECAB2 | NECAB2 | Homo sapiens N-terminal EF-hand calcium binding protein 2 (NECAB2), mRNA. |
| TIPARP | TIPARP | Homo sapiens TCD-inducible poly(ADP-ribose) polymerase (TIPARP), transcript variant 1, mRNA. |
| NECAB2 | NECAB2 | Homo sapiens N-terminal EF-hand calcium binding protein 2 (NECAB2), mRNA. |
| RNF40 | RNF40 | Homo sapiens ring finger protein 40 (RNF40), transcript variant 1, mRNA. |
| Rab28 | Rab28 | Homo sapiens ring finger protein 40 (RNF40), transcript variant 1, mRNA. |
| ZNF519 | ZNF519 | Homo sapiens zinc finger protein 519 (ZNF519), transcript variant 1, mRNA. |
| C1GALT1C1 | C1GALT1C1 | Homo sapiens C1GALT1-specific chaperone 1 (C1GALT1C1), transcript variant 1, mRNA. |
| PLCL1 | PLCL1 | Homo sapiens phospholipase C-like 1 (PLCL1), mRNA. |
| MLL5 | MLL5 | Homo sapiens myeloid/lymphoid or mixed-lineage leukemia 5 (MLL5), transcript variant 1, mRNA. |
| EGR1 | EGR1 | Homo sapiens early growth response 1 (EGR1), mRNA. |
| SLAIN1 | SLAIN1 | Homo sapiens SLAIN motif family, member 1 (SLAIN1), transcript variant 1, mRNA. |
| EGR4 | EGR4 | Homo sapiens early growth response 4 (EGR4), mRNA. |
| NBEA | NBEA | Homo sapiens neurobeachin (NBEA), transcript variant 1, mRNA. |
| FTCD | FTCD | Homo sapiens formiminotransferase cyclodeaminase (FTCD), transcript variant A, mRNA. |
| RYBP | RYBP | Homo sapiens RING1 and YY1 binding protein (RYBP), mRNA. |
| PLCB1 | PLCB1 | Homo sapiens phospholipase C, beta 1 (phosphoinositide-specific) (PLCB1), transcript variant 1, mRNA. |
| ZFP36L1 | ZFP36L1 | Homo sapiens zinc finger protein 36, C3H type-like 1 (ZFP36L1), transcript variant 1, mRNA. |
| MDGA2 | MDGA2 | Homo sapiens MAM domain containing glycosylphosphatidylinositol anchor 2 (MDGA2), transcript variant 1, mRNA. |
| GPR149 | GPR149 | Homo sapiens G protein-coupled receptor 149 (GPR149), mRNA. |

Uncharacterised genes are marked with **Uncharacterised**.
**SHRSP versus WKY differential gene expression at five weeks of age.**

| GenBank ID   | Gene Symbol | Fold Change | Uncharacterised | Human Equivalent |
|--------------|-------------|-------------|-----------------|------------------|
| ILMN_1370101| LOC499058   | 0.046880878| 0.02287929      | CEACAM1 Homo sapiens carcinoembryonic antigen-related cell adhesion molecule 1 (biliary glycoprotein) (CEACAM1), transcript variant 4, mRNA. |
| ILMN_1352135| CEACAM10    | 1           | 0.03833884      | CEACAM10 not in RefSeq but overlaps with CEACAM1 |
| ILMN_1371662| PPP2R1A     | 1           | 0.04218835      | Ppp2r1a Homo sapiens protein phosphatase 2, regulatory subunit A, alpha (PPP2R1A), transcript variant 1, mRNA. |
| ILMN_1363414| LOC365025   | 1           | 0.05140693      | Uncharacterised |
| ILMN_1370033| LOC498604   | 0.037413057| 0.04984         | Uncharacterised |

***Human equivalent of SHRSP differentially expressed genes***

| GenBank ID   | Gene Symbol | Fold Change | Uncharacterised | Human Equivalent |
|--------------|-------------|-------------|-----------------|------------------|
| ILMN_1359027| RGD1563482 | 0.047780599| 0.01925846      | Ceacam1 Homo sapiens family with sequence similarity 129, member B (FAM129B), transcript variant 1, mRNA. |
| ILMN_1360094| DPRA        | 0.019282511| 0.011280632     | Sdpr Homo sapiens serum deprivation response (SDPR), mRNA. |
| ILMN_2039346| HLA-DMA     | 0.01780235 | 0.0082139       | HLA-DMA Homo sapiens major histocompatibility complex, class II, DM alpha (HLA-DMA), mRNA. |
| ILMN_1349530| CRSP6       | 0.01569697 | 0.004775604     | ATP11B Homo sapiens ATPase, class VI, type 11B (ATP11B), mRNA. |
| ILMN_1369447| RGD1565673 | 0.015577014| 0.043683386     | TTF1 Homo sapiens transcription termination factor, RNA polymerase I (TTF1), transcript variant 1, mRNA. |
| ILMN_1360785| LOC361929   | 0.01569697 | 0.004775604     | XPNPEP1 Homo sapiens X-prolyl aminopeptidase (aminopeptidase P) 1, soluble (XPNPEP1), transcript variant 1, mRNA. |
| ILMN_1365118| ANKRD15     | 0.013436679| 0.029966683     | Kank1 Homo sapiens KN motif and ankyrin repeat domains 1 (KANK1), transcript variant 1, mRNA. |
| ILMN_1359741| COL3A1      | 0.008265852| 0.007226814     | COL3A1 Homo sapiens collagen, type III, alpha 1 (COL3A1), mRNA. |
| ILMN_1349043| EAF1        | 0.011225914| 0.012742299     | USM5 Homo sapiens guanine nucleotide binding protein (G protein), alpha inhibiting activity polypeptide 1 (GNA11), transcript variant 1, mRNA. |
| ILMN_1351156| GNA1        | 0.038758971| 0.006236786     | USM5 Homo sapiens up-regulated during skeletal muscle growth 5 homolog (mouse) (USM5G5), transcript variant 1, mRNA. |
| ILMN_1352524| USMG5       | 0.012053872| 0.004427391     | Uncharacterised |
| ILMN_1374825| LOC294789   | 0.00608658 | 0.029223587     | Uncharacterised |
| ILMN_1369541| TCF4        | 0.001125291| 0.000871212     | Uncharacterised |
| ILMN_1357413| LOC360443   | 0.007355372| 0.004427391     | Uncharacterised |
| ILMN_1650955| IFI27L      | 0.028212577| 0.01020475      | Uncharacterised |
| ILMN_1375028| HTATIP2     | 0.023521336| 0.036142857     | HTATIP2 Homo sapiens HIV-1 Tat interactive protein 2, 30KDa (HTATIP2), transcript variant 1, mRNA. |
| ILMN_1348843| SLC17A6     | 0.009307057| 0.016441558     | HTATIP2 Homo sapiens solute carrier family 17 (sodium-dependent}
### **SHRSP versus WKY differential gene expression at five weeks of age.**

| Gene ID           | Predicted Gene | SHRSP Value | WKY Value | Fold Change |
|-------------------|----------------|-------------|-----------|-------------|
| ILMN_1361722      | RBM8_PREDICTED | 0.01111111  | 0.001194296 | 0.01111111  |
| ILMN_1359627      | LOC360919      | 0.002249417 | 0.01020475  | 0.002249417 |
| ILMN_1358978      | RGD1306126     | 0.00258885  | 0.007217069 | 0.00258885  |
| ILMN_1350985      | INPP5D         | 0.002191781 | 0.004427391 | 0.002191781 |
| ILMN_1367329      | PPP1R16A_PREDICTED | 0.002684492 | 0.00848199  | 0.002684492 |
| ILMN_1361915      | PDE10A         | 0.001931818 | 0.008296558 | 0.001931818 |
| ILMN_1365885      | MMP14          | 0.001564171 | 0.003168831 | 0.001564171 |
| ILMN_1371064      | RPS18          | 0.001505682 | 0.007204301 | 0.001505682 |
| ILMN_1372988      | FHL2           | 0.0015427   | 0.001276224 | 0.0015427   |
| ILMN_1651148      | POLR2I_PREDICTED | 0.001505682 | 0.016747759 | 0.001505682 |
| ILMN_1350743      | PRMT5_PREDICTED | 0.001505682 | 0.001276224 | 0.001505682 |
| ILMN_1350481      | LOC499790      | 0.00482375  | 0.014687924 | 0.00482375  |
| ILMN_1352269      | RT1-149        | 0.0015427   | 0.016747759 | 0.0015427   |
| ILMN_1362726      | HMGN3          | 0.001564171 | 0.011285266 | 0.001564171 |
| ILMN_1357163      | SYMPK          | 0.001251863 | 0.008868687 | 0.001251863 |
| ILMN_1357432      | CYP11B1        | 0.00641711  | 0.03168831  | 0.00641711  |
| ILMN_1359619      | LYZL4_PREDICTED | 0.001564171 | 0.0046942205 | 0.001564171 |
| ILMN_1360418      | RGD1302996     | 0.00714286  | 0.001320756 | 0.00714286  |
| ILMN_1649821      | HAGHL          | 0.0040619   | 0.001603306 | 0.0040619   |
| ILMN_1349624      | KIF5C_PREDICTED | 0.00223587  | 0.002306649 | 0.00223587  |
| ILMN_1358541      | IGFBP6         | 0.00117302  | 0.002201705 | 0.00117302  |
| ILMN_1363262      | CNK2A1         | 0.00395257  | 0.000817866 | 0.00395257  |
| ILMN_1359301      | IGF2           | 0.000181818 | 0.011375291 | 0.000181818 |
| ILMN_1358234      | RGD1562351_PREDICTED | 0.000142045 | 0.000233766 | 0.000142045 |

### ***Human equivalent of SHRSP differentially expressed genes***

- **Rbm8a (RBM8A)**: Homo sapiens RNA binding motif protein 8A (RBM8A), mRNA.
- **Afp (AFP)**: Homo sapiens alpha-fetoprotein (AFP), mRNA.
- **Fam173a (FAM173A)**: Homo sapiens family with sequence similarity 173, member A (FAM173A), mRNA.
- **Inpp5d (INPP5D)**: Homo sapiens inositol polyphosphate-5-phosphatase, 145kDa (INPP5D), transcript variant 1, mRNA.
- **Ppp1r16a (PPP1R16A)**: Homo sapiens phosphodiesterase 10A (PDE10A), transcript variant 1, mRNA.
- **Mmp14 (MMP14)**: Homo sapiens matrix metalloproteinase 14 (membrane-inserted) (MMP14), mRNA.
- **Rps18 (RPS18)**: Homo sapiens ribosomal protein S18 (RPS18), mRNA.
- **Fhl2 (FHL2)**: Homo sapiens RNA binding motif protein 8A (RBM8A), mRNA.
- **Polr2i (POLR2I)**: Homo sapiens polymerase (RNA II) (DNA directed) polypeptide 1, 14.5kDa (POLR2I), mRNA.
- **Prmt5 (PRMT5)**: Homo sapiens arginine methyltransferase 5 (PRMT5), transcript variant 1, mRNA.
- **Hmgn3 (HMGN3)**: Homo sapiens high mobility group nucleosomal binding domain 3 (HMGN3), transcript variant 1, mRNA.
- **Sympk (SYMPK)**: Homo sapiens symplekin (SYMPK), mRNA.
- **Cyp11b1 (CYP11B1)**: Homo sapiens cytochrome P450, family 11, subfamily B, polypeptide 1 (CYP11B1), nuclear gene encoding mitochondrial protein, transcript variant 1, mRNA.
- **Lyzl4 (LYZL4)**: Homo sapiens lysozyme-like 4 (LYZL4), mRNA.
- **C6orf136**: Homo sapiens chromosome 6 open reading frame 136 (C6orf136), transcript variant 3, mRNA.
- **Haghl (HAGHL)**: Homo sapiens hydroxyacylglycerol-3-acylglycerol-1-hydrolase-like (HAGHL), transcript variant 2, mRNA.
- **Kif5c (KIF5C)**: Homo sapiens kinesin family member 5C (KIF5C), mRNA.
- **Igfbp6 (IGFBP6)**: Homo sapiens insulin-like growth factor binding protein 6 (IGFBP6), mRNA.
- **Csnk2a1 (CSNK2A1)**: Homo sapiens casein kinase 2, alpha 1 polypeptide (CSNK2A1), transcript variant 1, mRNA.
- **Igf2 (IGF2)**: Homo sapiens insulin-like growth factor 2 (somatomedin A) (IGF2), transcript variant 1, mRNA.

**Uncharacterised**

- **C6orf136**: Human homologue.
**SHRSP versus WKY differential gene expression at five weeks of age.**

| Gene ID     | Gene Symbol | Uncharacterised | Uncharacterised | Uncharacterised |
|-------------|-------------|-----------------|-----------------|-----------------|
| ILMN_1364214 | LOC497864   | 0.000117302     | 0.000170455     | NDUFAF5 Homo sapiens NADH dehydrogenase (ubiquinone) complex I, assembly factor 5 (NDUFAF5), nuclear gene encoding mitochondrial protein, transcript variant 1, mRNA. |
| ILMN_1351068 | RGD1309829_PREDICTED | 7.27E-05       | 9.74E-05       | Uncharacterised |
| ILMN_1368752 | LOC499378   | 7.27E-05        | 9.74E-05       | Uncharacterised |
| ILMN_1351487 | RT1-A1      | 20              | 0              | 9.74E-05       |
| ILMN_1353260 | PXMP4       | 3               | 0              | 4.33E-05       |
| ILMN_1353156 | COLQ        | 0               | 9.74E-05       | Uncharacterised |
| ILMN_1651096 | RGD1560364_PREDICTED | 0            | 9.74E-05       | Uncharacterised |
| ILMN_2038792 | ALB         | 14              | 0              | 0.000871212    |
| ILMN_1650062 | RGD1563903_PREDICTED | 0          | 0              | Uncharacterised |
| ILMN_1370031 | LOC362068   | 0               | 0              | Uncharacterised |
| ILMN_1371684 | LOC499103   | 0               | 0              | Uncharacterised |
| ILMN_1358480 | LOC365566   | 0               | 0              | Uncharacterised |
| ILMN_1376663 | LOC287167   | 10              | 0              | 0.000871212    |
| ILMN_1349205 | RGD1562905_PREDICTED | 1           | 0              | Uncharacterised |
| ILMN_1359180 | MRPL18_PREDICTED | 1             | 0              | 0              |

*** Human equivalent of SHRSP differentially expressed genes

| Gene ID     | Gene Symbol | Uncharacterised | Uncharacterised |
|-------------|-------------|-----------------|-----------------|
| ILMN_1364214 | LOC497864   | 0.000117302     | 0.000170455     |
| ILMN_1351068 | RGD1309829_PREDICTED | 7.27E-05       | 9.74E-05       |
| ILMN_1368752 | LOC499378   | 7.27E-05        | 9.74E-05       |
| ILMN_1351487 | RT1-A1      | 20              | 0              |
| ILMN_1353260 | PXMP4       | 3               | 0              |
| ILMN_1353156 | COLQ        | 0               | 9.74E-05       |
| ILMN_1651096 | RGD1560364_PREDICTED | 0            | 9.74E-05       |
| ILMN_2038792 | ALB         | 14              | 0              |
| ILMN_1650062 | RGD1563903_PREDICTED | 0          | 0              |
| ILMN_1370031 | LOC362068   | 0               | 0              |
| ILMN_1371684 | LOC499103   | 0               | 0              |
| ILMN_1358480 | LOC365566   | 0               | 0              |
| ILMN_1376663 | LOC287167   | 10              | 0              |
| ILMN_1349205 | RGD1562905_PREDICTED | 1           | 0              |
| ILMN_1359180 | MRPL18_PREDICTED | 1             | 0              |

**NDUFAF5** Homo sapiens NADH dehydrogenase (ubiquinone) complex I, assembly factor 5 (NDUFAF5), nuclear gene encoding mitochondrial protein, transcript variant 1, mRNA. Similar to oocyte-testis gene 1 (I can’t find a record of this gene).

**RT1-A1** is HLA-B.

**Ndufa5** Homo sapiens NADH dehydrogenase (ubiquinone) complex I, assembly factor 5 (NDUFAF5), nuclear gene encoding mitochondrial protein, transcript variant 1, mRNA.

**Pmp4** Homo sapiens peroxisomal membrane protein 4, 24kDa (PMP4), transcript variant 1, mRNA.

**Colq** Homo sapiens collagen-like tail subunit (single strand of homotrimer) of asymmetric acetylcholinesterase (COLQ), transcript variant 1, mRNA.

**Alb** Homo sapiens albumin (ALB), mRNA.

**Gpr98** Homo sapiens G protein-coupled receptor 98 (GPR98), transcript variant 1, mRNA.

**Hba1** Homo sapiens hemoglobin, alpha 1 (HBA1), mRNA.

**Rpl17** Homo sapiens ribosomal protein L17 (L23) (predicted).
Table III All candidate SHRSP gene-based results from LBC1936 and CHARGE with WMH variables. Chr is chromosome. nSNPs is the number of SNPs in the gene (+/- 50kb). Please note that the gene boundaries are overlapping as SNPs can be allocated to multiple genes, so the same SNP could be driving the signal in different genes. The results are ordered by significance.

| Chr | Gene   | Start Position | Stop Position | Genotyped SNPs | Discovery: LBC1936 | Imputed SNPs | Replication: CHARGE |
|-----|--------|----------------|---------------|----------------|--------------------|--------------|----------------------|
|     |        |                |               | nSNPs | P-value | P-value | nSNPs | P-value | P-value | nSNPs | P-value |
| 4   | AFP    | 74,520,796     | 74,540,356    | 13    | 0.0021 | 0.0009 | 77    | 0.0037 | 0.0037 | 67    | 0.83    |
| 4   | ALB    | 74,488,869     | 74,505,834    | 11    | 0.0027 | 0.0017 | 61    | 0.0063 | 0.0068 | 53    | 0.75    |
| 7   | GNAI1  | 79,602,075     | 79,686,661    | 42    | 0.034  | 0.033  | 181   | 0.014  | 0.015  | 166   | 0.79    |
| 1   | RBM8A  | 144,218,994    | 144,222,801   | 13    | 0.038  | 0.057  | 26    | 0.029  | 0.024  | 21    | 0.54    |
| 2   | INPP5D | 233,633,279    | 233,781,288   | 69    | 0.041  | 0.78   | 198   | 0.044  | 0.868  | 162   | 0.98    |
| 10  | XPNPEP1| 111,614,513    | 111,673,192   | 18    | 0.042  | 0.14   | 130   | 0.145  | 0.23   | 120   | 0.000101|
| 9   | NR4A3  | 101,623,957    | 101,668,994   | 13    | 0.045  | 0.16   | 62    | 0.106  | 0.245  | 56    | 0.48    |
| 13  | FARPI  | 97,593,434     | 97,900,024    | 154   | 0.049  | 0.25   | 550   | 0.18   | 0.507  | 468   | 0.025   |
| 18  | RPL17  | 45,268,853     | 45,272,904    | 16    | 0.058  | 0.13   | 79    | 0.076  | 0.218  | 74    | 0.28    |
| 6   | MRPL18 | 160,131,481    | 160,139,451   | 24    | 0.059  | 0.039  | 89    | 0.16   | 0.048  | 76    | 0.24    |
| 4   | RAB28  | 12,978,479     | 13,095,054    | 11    | 0.076  | 0.49   | 84    | 0.264  | 0.61   | 40    | 0.18    |
| 8   | CYP11B1| 143,950,774    | 143,958,238   | 30    | 0.080  | 0.49   | 103   | 0.15   | 0.613  | 87    | 0.13    |
| 2   | SDPR   | 192,407,280    | 192,420,226   | 15    | 0.081  | 0.10   | 68    | 0.24   | 0.314  | 55    | 0.25    |
| 1   | SIPA1L2| 230,600,334    | 230,717,866   | 80    | 0.087  | 0.00934 | 340   | 0.20   | 0.018  | 285   | 0.87    |
| 6   | PDE10A | 165,690,452    | 165,995,575   | 159   | 0.094  | 0.065  | 594   | 0.15   | 0.168  | 534   | 0.57    |
| 13  | SLAIN1 | 71,710,470     | 71,736,378    | 26    | 0.10   | 0.30   | 137   | 0.19   | 0.341  | 124   | 0.51    |
| 2   | KIF5C  | 149,349,288    | 149,591,519   | 31    | 0.10   | 0.66   | 148   | 0.14   | 0.711  | 128   | 0.35    |
| 12  | IGFBP6 | 51,777,702     | 51,782,395    | 16    | 0.13   | 0.08   | 67    | 0.17   | 0.065  | 60    | 0.24    |
| 21  | COL6A1 | 46,226,090     | 46,249,391    | 27    | 0.14   | 0.28   | 96    | 0.080  | 0.20   | 93    | 0.66    |
| 2   | FHL2   | 105,343,714    | 105,421,392   | 47    | 0.15   | 0.56   | 179   | 0.11   | 0.55   | 167   | 0.37    |
| 5   | SLC1A3 | 36,642,213     | 36,724,193    | 71    | 0.15   | 0.45   | 211   | 0.27   | 0.74   | 186   | 0.16    |
| 19  | RPS16  | 44,615,686     | 44,618,458    | 20    | 0.16   | 0.43   | 60    | 0.38   | 0.69   | 53    | 0.56    |
| Gene Symbol | Symbol | Chromosome | Position | p Value | DEG | Adj p Value |
|-------------|--------|------------|----------|---------|-----|-------------|
| SCN3A       | C21orf56 | 165,652,275 | 165,768,823 | 0.17 | 0.83 | 0.043 | 0.036 |
| RNF40       | 30,681,130 | 30,694,039 | 5 | 0.17 | 0.58 | 0.21 | 0.66 |
| FOS         | 1,541,286 | 201,545,352 | 19 | 0.18 | 0.33 | 0.37 | 0.55 |
| TM9SF4      | 20,160,969 | 20,178,722 | 16 | 0.20 | 0.29 | 0.28 | 0.17 |
| TTF1        | 134,240,757 | 134,272,042 | 37 | 0.20 | 0.23 | 0.61 | 0.40 |
| CYR61       | 85,819,047 | 85,821,978 | 19 | 0.21 | 0.88 | 0.28 | 0.10 |
| RPS9        | 59,396,537 | 59,403,327 | 22 | 0.21 | 0.77 | 0.37 | 0.55 |
| RPL31       | 100,985,122 | 101,002,587 | 28 | 0.21 | 0.27 | 0.066 | 0.16 |
| C6orf136    | 30,722,779 | 30,728,961 | 38 | 0.32 | 0.92 | 0.37 | 0.47 |
| VPS13C      | 59,931,881 | 60,139,939 | 56 | 0.30 | 0.76 | 0.37 | 0.47 |
| CTGF        | 132,311,009 | 132,314,211 | 33 | 0.30 | 0.16 | 0.78 | 0.48 |
| HLA-DMA     | 33,024,372 | 33,028,831 | 95 | 0.31 | 0.54 | 0.41 | 0.73 |
| C6orf136    | 30,722,779 | 30,728,961 | 38 | 0.32 | 0.92 | 0.39 | 0.94 |
| GUCY1A3     | 156,807,327 | 156,871,226 | 56 | 0.33 | 0.49 | 0.51 | 0.49 |
| ZNF597      | 3,426,110 | 3,433,491 | 19 | 0.33 | 0.70 | 0.20 | 0.62 |
| HBA1        | 166,678 | 167,520 | 16 | 0.34 | 0.26 | 0.37 | 0.34 |
| EGR1        | 137,829,079 | 137,832,903 | 16 | 0.34 | 0.91 | 0.62 | 0.94 |
| ZNF575      | 48,729,168 | 48,732,124 | 18 | 0.34 | 0.80 | 0.232 | 0.089 |
| DHX40       | 54,997,667 | 55,040,484 | 10 | 0.36 | 0.19 | 0.55 | 0.27 |
| PER1        | 7,984,512 | 7,996,478 | 27 | 0.38 | 0.49 | 0.55 | 0.57 |
| C20orf7     | 13,713,681 | 13,745,874 | 28 | 0.39 | 0.11 | 0.83 | 0.17 |
| ZCCHC9      | 80,633,177 | 80,644,872 | 22 | 0.39 | 0.14 | 0.57 | 0.30 |
| KANK1       | 494,702 | 736,103 | 169 | 0.40 | 0.19 | 0.52 | 0.43 |
| MEX3C       | 46,954,917 | 46,977,688 | 15 | 0.41 | 0.22 | 0.38 | 0.31 |
| TCF7L2      | 114,699,998 | 114,916,060 | 49 | 0.42 | 0.73 | 0.79 | 0.87 |
| PXMP4       | 31,754,210 | 31,771,797 | 9 | 0.43 | 0.44 | 0.40 | 0.59 |
| C12orf65    | 122,283,415 | 122,308,459 | 4 | 0.43 | 0.73 | 0.62 | 0.39 | 0.90 |
|   | Gene   | Chromosome | Position | z-score | p-value | mDS   | pDS  |
|---|--------|------------|----------|---------|---------|-------|------|
| 1 | PLA2G2A | 20,174,517 | 20,179,496 | 0.44    | 0.46    | 26    | 49   |
| 16| NECAB2  | 82,559,737 | 82,593,880 | 0.44    | 0.48    | 54    | 259  |
| 19| FKB8    | 18,503,567 | 18,515,383 | 0.44    | 0.74    | 9     | 49   |
| 19| ZNF582  | 61,586,459 | 61,596,701 | 0.46    | 0.52    | 22    | 128  |
| 3 | RYBP    | 72,506,438 | 72,578,464 | 0.46    | 0.44    | 29    | 166  |
| 14| STRN3   | 30,432,755 | 30,565,358 | 0.48    | 0.52    | 39    | 166  |
| 19| SYMPK   | 51,010,539 | 51,058,388 | 0.44    | 0.37    | 23    | 79   |
| 19| PRMT5   | 22,459,572 | 22,468,501 | 0.51    | 0.99    | 17    | 68   |
| 15| ADPGK   | 70,830,760 | 70,863,179 | 0.51    | 0.76    | 19    | 73   |
| 16| FAM173A | 711,158    | 712,591   | 0.51    | 0.68    | 16    | 39   |
| 21| FTCD    | 46,380,603 | 46,399,909 | 0.52    | 0.33    | 31    | 104  |
| 6 | HLA-B   | 31,429,627 | 31,432,968 | 0.52    | 0.67    | 84    | 383  |
| 12| TUBA1B  | 47,807,832 | 47,811,571 | 0.52    | 0.71    | 3     | 30   |
| 19| POLR2I  | 51,010,539 | 51,058,388 | 0.44    | 0.44    | 29    | 124  |
| 6 | HMGN3   | 79,967,680 | 80,001,174 | 0.54    | 0.44    | 12    | 42   |
| 3 | GPR149  | 155,538,154| 155,630,198| 0.54    | 0.50    | 32    | 167  |
| 8 | TCEB1   | 75,021,187 | 75,046,900 | 0.45    | 0.68    | 24    | 91   |
| 16| HAGHL   | 716,958    | 719,716   | 0.55    | 0.50    | 15    | 42   |
| 2 | COL3A1  | 189,547,343| 189,585,717| 0.56    | 0.18    | 27    | 143  |
| 2 | EGR4    | 73,371,564 | 73,374,811 | 0.57    | 0.43    | 8     | 43   |
| 16| ERAF    | 31,446,703 | 31,447,625 | 0.57    | 0.75    | 11    | 48   |
| 14| NFKBIA  | 34,940,466 | 34,943,711 | 0.57    | 0.65    | 23    | 95   |
| 19| ZNF582  | 61,607,529 | 61,628,212 | 0.58    | 0.41    | 28    | 88   |
| 6 | SGK1    | 134,532,076| 134,537,727| 0.59    | 0.79    | 21    | 78   |
| 9 | FAM129B | 129,307,438| 129,381,089| 0.59    | 0.37    | 31    | 119  |
| 2 | PLCL1   | 198,377,777| 198,721,365| 0.60    | 0.99    | 49    | 270  |
| 11| MED17   | 93,157,052 | 93,186,144 | 0.60    | 0.46    | 15    | 65   |
| 5 | FAM151B | 79,819,555 | 79,873,962 | 0.60    | 0.51    | 17    | 96   |
| 20| TPX2    | 29,790,564 | 29,853,264 | 0.62    | 0.35    | 16    | 112  |
| 6 | RPS18   | 33,347,829 | 33,352,259 | 0.64    | 0.66    | 14    | 63   |
| 19| CEACAM1 | 47,703,297 | 47,724,479 | 0.64    | 0.55    | 4     | 29   |
| 3 | COLQ    | 15,466,643 | 15,538,262 | 0.65    | 0.38    | 37    | 145  |
|   | Gene | Start (chr1) | End (chr1) | R1   | R2   | p1  | p2   | Delta p |
|---|------|--------------|------------|------|------|-----|------|---------|
|10 | RAB18| 27,833,254    | 27,869,105 | 22   | 0.66 | 0.70| 103  | 0.79  | 0.74 | 96   | 0.80 |
|19 | IER2 | 13,122,281    | 13,126,718 | 15   | 0.67 | 0.71| 33   | 0.45  | 0.66 | 21   | 0.19 |
|3  | LYZLA| 42,413,578    | 42,427,069 | 28   | 0.68 | 0.72| 138  | 0.88  | 0.88 | 112  | 0.99 |
|3  | ATP11B| 183,993,984   | 184,122,115| 26   | 0.69 | 0.21| 135  | 0.89  | 0.37 | 134  | 0.50 |
|5  | DUSP1| 172,127,706   | 172,130,809| 34   | 0.69 | 0.98| 107  | 0.79  | 0.98 | 90   | 0.29 |
|11 | HTATIP2| 20,341,806   | 20,361,905 | 24   | 0.69 | 0.44| 88   | 0.77  | 0.41 | 85   | 0.30 |
|2  | RNFL49| 101,258,984   | 101,291,584| 15   | 0.72 | 0.72| 80   | 0.68  | 0.94 | 77   | 0.17 |
|19 | FAM32A| 16,157,234    | 16,163,857 | 15   | 0.72 | 0.72| 74   | 0.72  | 0.95 | 71   | 0.11 |
|19 | ZNF317| 9,112,088     | 9,135,089  | 30   | 0.72 | 0.96| 328  | 0.89  | 0.45 | 284  | 0.31 |
|1  | GALNT2| 228,269,650   | 228,484,569| 99   | 0.74 | 0.18| 59   | 0.91  | 0.82 | 59   | 0.26 |
|19 | ZNF566| 41,630,421    | 41,672,177 | 16   | 0.76 | 0.52| 1036 | 0.75  | 0.73 | 967  | 0.89 |
|14 | MDGA2| 46,378,577    | 47,213,738 | 164  | 0.76 | 0.63| 90   | 0.67  | 0.79 | 81   | 0.061|
|19 | ZNF45 | 49,108,620    | 49,121,398 | 16   | 0.76 | 0.88| 90   | 0.67  | 0.79 | 81   | 0.061|
|1  | RPL22 | 6,167,666     | 6,182,266  | 19   | 0.78 | 0.89| 75   | 0.99  | 0.73 | 68   | 0.98 |
|1  | PTG52 | 184,907,591   | 184,916,179| 16   | 0.79 | 0.44| 98   | 0.61  | 0.66 | 97   | 0.94 |
|7  | MLL5 | 104,441,872   | 104,541,768| 21   | 0.79 | 0.39| 98   | 0.61  | 0.66 | 97   | 0.94 |
|10 | USMG5 | 101,380,803   | 101,461,213| 10   | 0.79 | 0.88| 31   | 0.92  | 0.96 | 28   | 0.000142|
|14 | MMP14 | 22,375,632    | 22,386,643 | 40   | 0.79 | 0.90| 99   | 0.87  | 0.97 | 75   | 0.046|
|19 | PPP2R1A| 57,385,045   | 57,421,483 | 47   | 0.80 | 0.66| 141  | 0.94  | 0.84 | 132  | 0.41|
|6  | OOEPE | 74,135,000    | 74,136,236 | 9    | 0.80 | 0.93| 65   | 0.76  | 0.92 | 55   | 0.87 |
|8  | PPP1R16A| 145,692,916 | 145,698,312| 6    | 0.81 | 0.11| 24   | 0.78  | 0.25 | 20   | 0.27 |
|8  | ARC | 143,689,411   | 143,692,835| 17   | 0.81 | 0.94| 81   | 0.22  | 0.98 | 78   | 0.033|
|5  | GPR98 | 89,890,372    | 90,495,789 | 142  | 0.82 | 0.75| 657  | 0.91  | 0.88 | 591  | 0.74 |
|9  | ZNF189 | 103,200,983   | 103,212,763| 25   | 0.82 | 0.30| 135  | 0.72  | 0.28 | 128  | 0.98 |
|10 | POLL | 103,328,628   | 103,337,963| 12   | 0.83 | 0.90| 79   | 0.90  | 0.76 | 73   | 0.69 |
|10 | EGR2 | 64,241,762    | 64,246,133 | 21   | 0.84 | 0.93| 100  | 0.88  | 0.95 | 97   | 0.84 |
|2  | NABI | 191,222,092   | 191,265,737| 17   | 0.85 | 1.00| 88   | 0.87  | 1.00 | 86   | 0.12 |
|14 | ZFP361L | 68,324,127   | 68,329,538 | 31   | 0.86 | 0.47| 87   | 0.76  | 0.41 | 84   | 0.20 |
|2  | PER2 | 238,817,417   | 238,861,946| 18   | 0.87 | 0.96| 92   | 0.20  | 0.66 | 70   | 0.37 |
|14 | FAM14A| 93,663,870    | 93,665,710 | 34   | 0.87 | 0.98| 96   | 0.98  | 1.00 | 89   | 0.35 |
|18 | ZNF519| 14,094,723    | 14,122,429 | 13   | 0.88 | 0.75| 49   | 0.87  | 0.83 | 41   | 0.86 |
|20 | CSNK2AI| 411,337      | 472,482    | 34   | 0.90 | 0.53| 115  | 0.94  | 0.29 | 100  | 0.41 |
| Rank | Gene  | Start Position | End Position | Length | Coverage 1 | Coverage 2 | Coverage 3 | Coverage 4 |
|------|-------|----------------|--------------|--------|------------|------------|------------|------------|
| 4    | GPM6A | 176,791,081    | 177,160,642  | 95     | 0.90       | 0.95       | 383        | 0.94       |
| 13   | NBEA  | 34,414,455     | 35,144,873   | 114    | 0.93       | 0.72       | 621        | 0.99       |
| 11   | IGF2  | 2,106,922      | 2,127,409    | 21     | 0.94       | 0.64       | 83         | 0.92       |
| 1    | RGS2  | 191,044,793    | 191,048,026  | 11     | 0.94       | 0.82       | 84         | 0.88       |
| 1    | BAI2  | 31,965,304     | 32,002,235   | 17     | 0.95       | 0.65       | 61         | 0.91       |
| 20   | PLCB1 | 8,061,295      | 8,813,547    | 290    | 0.95       | 0.84       | 1123       | 0.94       |
| 19   | ZNF461| 41,820,122     | 41,849,579   | 9      | 0.98       | 0.55       | 54         | 0.90       |
| 19   | JUNB  | 12,763,309     | 12,765,125   | 5      | 0.99       | 0.48       | 29         | 1.00       |

Sequence 1: 34414455-35144873
Coverage 1: 0.93
Sequence 2: 2106922-2127409
Coverage 2: 0.94
Sequence 3: 191044793-191048026
Coverage 3: 0.94
Sequence 4: 31965304-32002235
Coverage 4: 0.95
Sequence 5: 8061295-8813547
Coverage 5: 0.95
Sequence 6: 41820122-41849579
Coverage 6: 0.98
Sequence 7: 12763309-12765125
Coverage 7: 0.99
Table IV Top hits from genome wide association study with WMH variables in LBC1936 (P < 1x10^{-5}).

| SNP         | Chromosome | Position    | Nearest Gene       | Risk Allele | Allele Frequency | Beta       | P           |
|-------------|------------|-------------|--------------------|-------------|------------------|------------|-------------|
| **WMH**     |            |             |                    |             |                  |            |             |
| rs7312545   | 12         | 104,877,937 | NUAK1              | A           | 0.48             | 0.1775     | 4.70x10^{-6} |
| rs1908311   | 1          | 164,698,889 | FMO9P              | T           | 0.34             | 0.1761     | 6.27x10^{-6} |
| rs1778193   | 1          | 94,886,005  | SLC44A3/F3         | T           | 0.10             | -          | 7.50 x10^{-6} |
| rs1344567   | 12         | 104,869,369 | NUAK1              | T           | 0.49             | 0.1735     | 7.71 x10^{-6} |
| rs10439220  | 2          | 235,653,271 | SH3BP4             | C           | 0.27             | 0.1735     | 8.05 x10^{-6} |
| **Fazekas** |            |             |                    |             |                  |            |             |
| score       |            |             |                    |             |                  |            |             |
| rs1156440   | 4          | 106,244,538 | TET2               | A           | 0.26             | 0.1931     | 1.45 x10^{-6} |
| rs764275    | 4          | 106,236,076 | TET2               | G           | 0.26             | 0.1858     | 3.67 x10^{-6} |
| rs1991979   | 4          | 147,368,197 | LSM6/SLC10A7       | C           | 0.37             | 0.1794     | 8.27 x10^{-6} |
| rs9905906   | 17         | 7,626,473   | DNAH2              | T           | 0.35             | 0.1778     | 8.58 x10^{-6} |
| rs1923416   | 6          | 89,348,510  | RNGTT              | C           | 0.11             | 0.1771     | 8.92 x10^{-6} |
| rs13250792  | 8          | 16,784,710  | FGF20              | G           | 0.27             | 0.1773     | 9.50 x10^{-6} |
| rs6561615   | 13         | 50,595,795  | GUGY1B2            | A           | 0.13             | 0.1773     | 9.87 x10^{-6} |
Table V Top gene-based results for the LBC1936 WMH variables analysed in Vegas (P<0.001). Chr is chromosome. nSNPs is the number of SNPs in the gene (+/- 50kb). Please note that the gene boundaries are overlapping as SNPs can be allocated to multiple genes, so the same SNP could be driving the signal in different genes. The results are ordered by significance.

| Chr | Gene   | nSNPs | Start  | Stop     | P       | Chr | Gene   | nSNPs | Start  | Stop     | P         |
|-----|--------|-------|--------|----------|---------|-----|--------|-------|--------|----------|-----------|
| 17  | FBF1   | 17    | 71,418,212 | 71,448,714 | 8.00x10^-6 | 17  | TRIM65 | 10    | 71,396,635 | 71,404,649 | 5.7x10^-3 |
| 17  | MRPL38 | 10    | 71,406,318 | 71,413,069 | 1.20x10^-5 | 17  | FBF1   | 17    | 71,418,212 | 71,448,714 | 6.10x10^-5 |
| 17  | TRIM65 | 10    | 71,396,635 | 71,404,649 | 1.80x10^-5 | 19  | HDGF2  | 21    | 4,423,254  | 4,453,222  | 6.70x10^-5 |
| 17  | TRIM47 | 10    | 71,381,839 | 71,386,251 | 1.90x10^-5 | 17  | MRPL38 | 10    | 71,406,318 | 71,413,069 | 6.70x10^-5 |
| 17  | WBP2   | 11    | 71,353,374 | 71,363,096 | 2.20x10^-5 | 17  | TRIM47 | 10    | 71,381,839 | 71,386,251 | 7.40x10^-5 |
| 17  | UNC13D | 13    | 71,334,901 | 71,352,393 | 3.50x10^-5 | 19  | UBXD1  | 16    | 4,396,260  | 4,408,790  | 8.30x10^-5 |
| 17  | ACOX1  | 26    | 71,449,186 | 71,487,039 | 5.00x10^-5 | 19  | LSM6   | 17    | 147,316,284 | 147,330,663 | 1.00x10^-5 |
| 15  | SLC12A1| 27    | 46,285,789 | 46,383,568 | 7.10x10^-5 | 17  | MRPL24 | 17    | 154,973,717 | 154,977,547 | 0.00011   |
| 15  | DUT    | 20    | 46,401,912 | 46,422,862 | 8.40x10^-5 | 1    | HDGF   | 17    | 154,978,522 | 154,988,864 | 0.00012   |
| 12  | CD55   | 21    | 71,292,275 | 71,333,481 | 0.00012  | 17  | WBP2   | 11    | 71,353,374 | 71,363,096 | 0.00014   |
| 4   | LSM6   | 17    | 147,316,284 | 147,330,663 | 0.00024  | 19  | LSM6   | 17    | 154,959,036 | 154,964,329 | 0.00016   |
| 9   | ODF2   | 8     | 130,258,252 | 130,303,060 | 0.00035  | 19  | LS8D5  | 24    | 4,473,543  | 4,486,208  | 0.00025   |
| 1   | MRPL24 | 17    | 154,973,717 | 154,977,547 | 0.00037  | 1   | HDGF   | 17    | 154,978,522 | 154,988,864 | 0.00012   |
| 9   | C1orf66| 19    | 154,964,901 | 154,973,365 | 0.00035  | 1    | HDGF   | 17    | 154,978,522 | 154,988,864 | 0.00012   |
| 17  | TMEM106A| 2  | 38,719,419  | 38,727,115  | 0.00054  | 19  | LS8D5  | 24    | 4,473,543  | 4,486,208  | 0.00025   |
| 1    | C1orf66| 19    | 154,964,901 | 154,973,365 | 0.00059  | 16  | SSTR5  | 30    | 1,068,869  | 1,069,964  | 0.00031   |
| 1    | TMEM106A| 2  | 38,719,419  | 38,727,115  | 0.00059  | 19  | LS8D5  | 24    | 4,473,543  | 4,486,208  | 0.00025   |
| 1    | C1orf66| 19    | 154,964,901 | 154,973,365 | 0.00059  | 19  | LS8D5  | 24    | 4,473,543  | 4,486,208  | 0.00025   |
| 19                                                                 | 17  | TLE3   | 34    | 68,127,596 | 68,177,310 | 0.00081   |
| 9   | CERCAM | 8     | 130,222,579 | 130,239,451 | 0.00090  | 17  | ACOX1  | 26    | 71,449,186 | 71,487,039 | 0.00094   |
Figure I Genome-wide association study results of WMH volume (a) and Fazekas score (b) using genotyped data on 542,050 SNPs in LBC1936. QQ and Manhattan plots are shown.
References

(1) Houlihan LM, Davies G, Tenesa A, Harris SE, Luciano M, Gow AJ et al. Common variants of large effect in F12, KNG1, and HRG are associated with activated partial thromboplastin time. *Am J Hum Genet* 2010;86:626-31.

(2) Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 2005;102:15545-50.

(3) Wang K, Li M, Hakonarson H. Analysing biological pathways in genome-wide association studies. *Nat Rev Genet* 2010;11:843-54.

(4) Bailey EL, McBride MW, Crawford W, McClure JD, Graham D, Dominiczak AF et al. Differential gene expression in multiple neurological, inflammatory and connective tissue pathways in a spontaneous model of human small vessel stroke. *Neuropathol Appl Neurobiol* 2014;40:855-72.