Genetic basis of lacunar stroke: a pooled analysis of individual patient data and genome-wide association studies

Citation for published version:
Traylor, M, Persyn, E, Tomppo, L, Klasson, S, Abedi, V, Bakker, MK, Torres, N, Li, L, Bell, S, Rutten-jacobs, L, Tozer, DJ, Griessenauer, CJ, Zhang, Y, Pedersen, A, Sharma, P, Jimenez-conde, J, Rundek, T, Grewal, RP, Lindgren, A, Meschia, JF, Salomaa, V, Havulinna, A, Kourkoulis, C, Crawford, K, Marini, S, Mitchell, BD, Kittner, SJ, Rosand, J, Dichgans, M, Jern, C, Strobian, D, Fernandez-cadenas, I, Zand, R, Ruigrok, Y, Rost, N, Lemmens, R, Rothwell, PM, Anderson, CD, Wardlaw, J, Lewis, CM & Markus, HS 2021, 'Genetic basis of lacunar stroke: a pooled analysis of individual patient data and genome-wide association studies', The Lancet Neurology. https://doi.org/10.1016/S1474-4422(21)00031-4

Digital Object Identifier (DOI):
10.1016/S1474-4422(21)00031-4

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published In:
The Lancet Neurology

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
GENETIC BASIS OF LACUNAR STROKE: A POOLED ANALYSIS OF INDIVIDUAL PATIENT DATA AND GENOME-WIDE ASSOCIATION STUDIES

Matthew Taylor, Elodie Persyn, Lisa Tomppo, Sofia Klasson, Vida Abedi, Mark K Bakker, Nuria Torres, Linxin Li, Steven Bell, Loes Rutten-Jacobs, Daniel J Tozer, Christoph J Griesenauer, Yanfei Zhang, Annie Pedersen, Pankaj Sharma, Jordi Jimenez-Conde, Tatjana Rundek, Raji P Grewal, Arne Lindgren, James F Meschia, Veikko Salomaa, Aki Havulinna, Christina Koukkouli, Katherine Crawford, Sandro Marini, Braxton D Mitchell, Steven J Kittner, Jonathan Rosand, Martin Dichgans, Christina Jern, Daniel Strbian, Israel Fernandez-Cadenas, Ramin Zand, Ynte Ruigrok, Natalia Rost, Robin Lemmens, Peter M Rothwell, Christopher D Anderson, Joanna Wardlaw, Cathryn M Lewis, Hugh S Markus, Helsinki Stroke Study, Dutch Parelsnoer Institute-Cerebrovascular Accident (CVA) Study Group, National Institute of Neurological Disorders and Stroke (NINDS) Stroke Genetics Network, UK DNA Lacunar Stroke Study Investigators, International Stroke Genetics Consortium

Summary

Background The genetic basis of lacunar stroke is poorly understood, with a single locus on 16q24 identified to date. We sought to identify novel associations and provide mechanistic insights into the disease.

Methods We did a pooled analysis of data from newly recruited patients with an MRI-confirmed diagnosis of lacunar stroke and existing genome-wide association studies (GWAS). Patients were recruited from hospitals in the UK as part of the UK DNA Lacunar Stroke studies 1 and 2 and from collaborators within the International Stroke Genetics Consortium. Cases and controls were stratified by ancestry and two meta-analyses were done: a European ancestry analysis, and a transethnic analysis that included all ancestry groups. We also did a multi-trait analysis of GWAS, in a joint analysis with a study of cerebral white matter hyperintensities (an aetiologically related radiological trait), to find additional genetic associations. We did a transcriptome-wide association study (TWAS) to detect genes for which expression is associated with lacunar stroke; identified significantly enriched pathways using multi-marker analysis of genomic annotation; and evaluated cardiovascular risk factors causally associated with the disease using mendelian randomisation.

Findings Our meta-analysis comprised studies from Europe, the USA, and Australia, including 7338 cases and 254798 controls, of which 2987 cases (matched with 29540 controls) were confirmed using MRI. Five loci (ICA1L-WDR12-CARF-NBEAL1, ULK4, SPI1-SLC39A13-PSCMC3-RAPSN, ZCCHC14, ZBTB14-EPB41L3) were found to be associated with lacunar stroke in the European or transethnic meta-analyses. A further seven loci (SLC25A44-PMF1-BGLAP, LOX-ZNF474-LOC100508584, FOXX2-FOXQ1, VTA1-GPR126, SH3PD2A, HTRA1-ARMS2, COL4A2) were found to be associated in the multi-trait analysis with cerebral white matter hyperintensities (n=4130). Two of the identified loci contain genes (COL4A2 and HTRA1) that are involved in monogenic lacunar stroke. The TWAS identified associations between the expression of six genes (SCL25A44, ULK4, CARF, FAM117B, ICA1L, NBEAL1) and lacunar stroke. Pathway analyses implicated disruption of the extracellular matrix, phosphatidylinositol 5 phosphate binding, and roundabout binding (false discovery rate <0.05). Mendelian randomisation analyses identified positive associations of elevated blood pressure, history of smoking, and type 2 diabetes with lacunar stroke.

Interpretation Lacunar stroke has a substantial heritable component, with 12 loci now identified that could represent future treatment targets. These loci provide insights into lacunar stroke pathogenesis, highlighting disruption of the vascular extracellular matrix (COL4A2, LOX, SH3PD2A, GPR126, HTRA1), pericyte differentiation (FOXX2, GPR126), TGF-β signalling (HTRA1), and myelination (ULK4, GPR126) in disease risk.

Funding British Heart Foundation.

Copyright © 2021 The Author(s). Published by Elsevier Ltd. This is an Open Access article under the CC BY 4.0 license.

Introduction

Lacunar strokes are small subcortical infarcts that arise from ischaemia in the territory of the deep perforating arteries of the brain. They account for up to a quarter of all ischaemic strokes and are usually due to cerebral small vessel disease, which is also the most common pathological underlyi
Because studies of other MRI markers of cerebral small vessel disease,7 and sporadic lacunar stroke has been strongly associated with a family history of stroke.4 Additionally, studies of other MRI markers of cerebral small vessel disease have shown a substantial genetic component; a further seven associated with lacunar stroke jointly with white matter hyperintensities. These loci highlight several key mechanisms in lacunar stroke pathogenesis, including extracellular matrix dysfunction, myelination, and pericyte differentiation. The current findings also show that individuals with increased genetic predisposition to elevated blood pressure, smoking, and type 2 diabetes are at increased risk of lacunar stroke, suggesting a causal role for these factors.

Implications of all the available evidence
No treatments for prevention of lacunar stroke are available, aside from management of vascular risk factors, such as blood pressure. This is due in part to poor understanding of the mechanisms underlying the disease. Our findings highlight novel mechanisms underlying lacunar stroke pathogenesis, and therefore point to pathways that have potential to be targeted by therapeutics. Improved treatment of elevated blood pressure and type 2 diabetes, as well as prevention of smoking in the population, is likely to reduce the burden of lacunar stroke.

For the GWAS, lacunar stroke cases were recruited from a combination of acute stroke admissions and outpatient services from Europe, the USA, South America, and Australia. Study inclusion criteria are detailed in the appendix (pp 99–114). Cases and controls were stratified by ancestry (European, south Asian, African, Hispanic) and analysed separately. We did two meta-analyses: a European ancestry analysis, and a transethnic analysis that included all ancestry groups. For each contributing study, approval for inclusion in this analysis complied with local ethical standards and with local institutional review board or ethics committee oversight. All participants provided written informed consent for genetic studies.

Lacunar stroke samples were divided into two strata: the MRI-confirmed group and the standard phenotyping group. In the MRI-confirmed group, lacunar stroke was defined as a clinical lacunar syndrome22 with an anatomically compatible lesion on MRI (subcortical infarct, ≤15 mm in diameter), either as a region of high intensity on diffusion-weighted imaging for acute infarcts or as an ischemically compatible lesion on MRI (subcortical infarct,

Methods
Study design and phenotype definitions
We did a GWAS of lacunar stroke to identify novel associations and provide mechanistic insights into lacunar stroke. We recruited cases of lacunar stroke from hospitals across the UK as part of the UK DNA Lacunar Stroke studies 1 and 2,10 and from collaborators within the International Stroke Genetics Consortium; and re-analysed data from previous studies.11–14 Because MRI confirmation of lacunar stroke is more reliable than standard phenotyping,15–16 we focused on recruiting MRI-confirmed cases. First, we did a GWAS of these data to identify novel genetic loci associated with lacunar stroke. Second, we used a multi-trait approach to detect additional genetic variation associated with lacunar stroke in a joint analysis with cerebral white matter hyperintensities from a large-scale GWAS.17 Third, we did a transcriptome-wide association study (TWAS) to identify transcribed genes for which expression is associated with lacunar stroke, and used mendelian randomisation to assess common cardiovascular risk factors that contribute to the disease. We did separate analyses in MRI-confirmed and standard phenotyping groups to assess whether MRI confirmation improves power to detect genetic associations.
infarcts); and any other specific cause of stroke (eg, lupus anticoagulant, cerebral vasculitis, dissection, and monogenic cause of stroke).

In the standard phenotyping group, lacunar stroke was defined according to the TOAST criteria, based on a clinical lacunar syndrome and the absence of other causes of stroke, or non-lacunar infarction on CT.

Genotyping and imputation

Genotyping arrays, quality control filters, and imputation reference panels are listed in the appendix (pp 7–8). All studies inferred the genetic ancestry of samples by comparison with reference populations using principal components analysis. European samples in this study are defined as those that segregated with European ancestry reference samples. The majority of studies were imputed to the Haplotype Reference Consortium build. If this was not possible due to logistical or ethical reasons, imputation to 1000 Genomes Phase 3 (all ancestry groups) panels was used.14,18

Statistical analysis

All studies used logistic regression to assess the association of single nucleotide polymorphism (SNP) allele dosages with lacunar stroke, including ancestry informative principal components as covariates as appropriate. All studies included cases with geographically matched controls, as confirmed by principal components analysis. Some studies had a combination of cases based on TOAST diagnosis of lacunar stroke, which were re-analysed for MRI confirmation of lacunar stroke, and cases based on TOAST diagnosis of lacunar stroke only, for which MRI was either not acquired or was inaccessible. In these circumstances, we analysed the MRI-confirmed and TOAST-only (standard phenotyping) groups separately and divided the study controls between the two groups to avoid any sample overlap. Any cases with subsequent MRI confirmation of lacunar stroke were omitted from the TOAST only group; all individuals were analysed only once.

Meta-analysis was done based on a fixed-effects inverse-variance weighted procedure using the METAL tool. Meta-analysis was done in the MRI-confirmed and standard phenotyping groups separately, and in all studies combined. We used the principles described by Winkler and colleagues to scrutinise datasets used in the meta-analysis. For each study, we filtered out SNPs with imputation INFO scores of less than 0.7 or minor allele frequency (MAF) of less than 0.01. Additionally, we removed low frequency or poorly imputed SNPs in small studies by removing variants for which INFO × MAF × number of cases equalled less than 2. Genomic control correction based on genomic inflation factor \( \lambda \) was used for each study to adjust for any residual inflation. Linkage disequilibrium (LD) score intercept values were used to assess whether population structure had been sufficiently resolved at the meta-analysis level. After meta-analysis, we excluded SNPs that were not present in at least 50% of cases.

We defined significant loci as those containing SNPs with a p value of less than 5 × 10^{-8} and in linkage equilibrium (\( r^2 > 0.1 \)) with other lead SNPs. If multiple loci met these criteria within 1 Mb we used conditional and joint multiple SNP analysis (GCTA-cojo) to evaluate whether these SNPs remained genome-wide significant in a joint modelling scenario. We used Nagelkerke \( R^2 \) values to calculate the proportion of variance explained by genome-wide significant SNPs using the Nagelkerke R2 function in the R package BMASS. To obtain an estimate of \( R^2 \), we subtracted the \( R^2 \) estimate for the model that included only principal components from the model that also contained the genome-wide significant SNPs. We used a genome-based restricted maximum likelihood (GREML) approach, implemented in GCTA software, and LD score regression, to estimate the heritability of lacunar stroke (MRI-confirmed and non-confirmed).

We did multi-trait analysis of GWAS (MTAG), performing a joint analysis with a large study of cerebral white matter hyperintensities on MRI (n=42 310), which shares a common cause with lacunar stroke through cerebral small vessel disease, to uncover additional genetic variation associated with lacunar stroke. We considered associations significant if they had a p value of less than 5 × 10^{-8} in MTAG, had a p value of less than 0.05 for association with white matter hyperintensities and lacunar stroke in univariate analysis, and showed greater significance in MTAG than in univariate analyses for white matter hyperintensities or lacunar stroke. To confirm our associations, we used an alternative approach, Bayesian multivariate analysis of summary statistics (using R package BMASS). We used the TWAS approach to identify genes for which genetically altered expression was associated with lacunar stroke. Analyses were done using FUSION, from gene expression models derived from the Genotype-Tissue Expression (GTEx) portal V7, CommonMind Consortium.

Figure 2: Analysis pipeline
GWAS=genome-wide association study.

Allison McCance Center for Brain Health (C Kourkoulis, Prof J Rosand, C D Anderson); Department of Neurology (Prof J Rosand, C D Anderson), and (Phil) Kistler Stroke Research Center (Prof N Rost MD), Massachusetts General Hospital, Boston, MA, USA; Program in Medical & Population Genetics, Broad Institute of Harvard and MIT, Cambridge, MA, USA (C Kourkoulis, K Crawford, S Marinis, Prof J Rosand, C D Anderson); Division of Endocrinology, Diabetes and Nutrition, Department of Medicine (Prof B D Mitchell PhD) and Department of Neurology (Prof S J Kittner MD), University of Maryland School of Medicine, Baltimore, MD, USA; Geriatric Research and Education Clinical Center, Baltimore Veterans Administration Medical Center, Baltimore, MD, USA (Prof B D Mitchell, Prof S J Kittner); Institute for Stroke and Dementia Research (ISD), LMU Munich, Munich, Germany (Prof M Dichgans MD); Munich Cluster for Systems Neurology (SyNergy), Munich, Germany (Prof M Dichgans); Clinical Neurosciences, University of Helsinki, Helsinki, Finland (O Stibran); Neurovascular Research Laboratory and Neurovascular Unit, Institut de Recerca, Hospital Vall d’Hebron, Universitat Autonoma de Barcelona, Barcelona, Spain
| Chromosome | Base Genomic position | Identifier | Genomic context | MTAG Genomic context | Locus | Stroke (European: 6020 cases; 239 cases) | Lacunar stroke (transethnic: 7338 cases, 258 controls) |
|------------|----------------------|------------|----------------|----------------------|--------|-----------------------------------|--------------------------------------------------|
| 2          | 203964973            | rs225744 C/T | Intergenic     | 0.77 1.11 (0.03)     | 1.4    | 156 2.4                         | 2.0 1.7                                           |
| 16         | 51253464             | rs2293576 G/A | Exonic         | 0.67 1.14 (0.02)     | 3.7    | 9.9                           | 6.0                                              |
| 20         | 15664813             | rs11838776 A/G | Intronic       | 0.29 1.11 (0.02)     | 3.2    | 1.7                           | 4.0                                              |
| 4           | 2303655              | rs2303655 T/C | Downstream     | 0.81 1.14 (0.03)     | 41     | 2.5                           | 1.8 1.6                                          |
| 11          | 72934535             | rs72934535 T/C | Intronic       | 0.89 1.25 (0.04)     | 3.5    | 7.9                           | 105 1.9                                           |
| 7           | 79404314             | rs79043147 T/C | Intronic       | 0.07 1.21 (0.04)     | 6.4    | 3.7                           |                                                  |
| 22          | 64234567             | rs2984613 C/T | Intronic       | 0.64 1.10 (0.02)     | 7.2    | 87                           | 7.1 3.1                                           |
| 12          | 12445022             | rs12445022 A/G | Intergenic     | 0.34 1.10 (0.02)     | 7.2    | 87                           | 7.1 3.1                                           |
| 5           | 22518537             | rs22518537 T/C | Intergenic     | 0.81 1.14 (0.02)     | 3.6    | 10.0                          | 13 4.0 1.4                                       |
| 2           | 15919378             | rs9842261 C/T | Genomic context | 0.64 1.10 (0.02)     | 3.6    | 10.0                          | 13 4.0 1.4                                       |
| 16          | 22485022             | rs22485022 A/G | Genomic context | 0.34 1.10 (0.02)     | 3.6    | 10.0                          | 13 4.0 1.4                                       |
| 16          | 54858532             | rs54858532 T/C | Intergenic     | 0.10 1.10 (0.02)     | 9.9    | 10.0                          | 13 4.0 1.4                                       |
| 22          | 22450022             | rs22450022 A/G | Genomic context | 0.34 1.10 (0.02)     | 3.6    | 10.0                          | 13 4.0 1.4                                       |
| 16          | 54858532             | rs54858532 T/C | Intergenic     | 0.10 1.10 (0.02)     | 9.9    | 10.0                          | 13 4.0 1.4                                       |
| 22          | 22450022             | rs22450022 A/G | Genomic context | 0.34 1.10 (0.02)     | 3.6    | 10.0                          | 13 4.0 1.4                                       |
| 22          | 22450022             | rs22450022 A/G | Genomic context | 0.34 1.10 (0.02)     | 3.6    | 10.0                          | 13 4.0 1.4                                       |

The CMC gene expression tissues (labelled as CMC brain) were collected from the dorsolateral prefrontal cortex of individuals with schizophrenia or controls (TWAS n=452). The YFS study (labelled as YFS whole blood), peripheral blood gene expression was analysed (TWAS n=1264). Among the available GTEx tissues, we focused our TWAS analysis on the aortic artery (TWAS n=267), coronary artery (TWAS n=152), tibial artery (TWAS n=388), and whole blood (TWAS n=369), based on the assumption that these tissues would be the most relevant for lacunar stroke pathogenesis. Bonferroni correction for multiple testing was applied, taking into account the total number of tested genes across the tissues (cutoff for significance: p<1.5×10⁻⁶). TWAS results were further investigated with colocalisation analysis of expression quantitative trait loci and GWAS signals with the R package COLOC, to assess whether the observed expression quantitative trait loci and GWAS associations were consistent with a shared association.

In a bioinformatics analysis for novel associations, we used Phenoscanner to query whether genome-wide significant SNPs were associated with DNA methylation, metabolite levels, or protein levels from genome-wide studies at genome-wide significance (p<5×10⁻⁸) in other GWAS. We scanned DrugBank and the Drug Gene Interaction Database to assess the therapeutic potential of targeting associated genes. To identify biological pathways significantly associated with risk of lacunar stroke, we used the multi-marker analysis of genomic annotation (MAGMA) tool. We first used MAGMA to calculate the significance of each gene based on association results, and then used these gene-level statistics to estimate enrichment of Gene Ontology pathways from the Molecular Signatures Database using a gene-set enrichment analysis approach. We investigated only Gene Ontology terms containing at least four and fewer than 200 genes and considered pathways attaining a false discovery rate (FDR) of less than 0.05 as being significantly associated with lacunar stroke.

To validate these findings, we used Mendel randomisation analyses to evaluate whether any lipid (LDL, HDL, or triglycerides), blood pressure (systolic blood pressure, diastolic blood pressure, pulse pressure), metabolic (type 2 diabetes, body-mass index), or lifestyle risk factors (ever smoking) have a causal impact on lacunar stroke based on genetics. Instrumental variables were independent (LD r²<0.01) of tested genes across the tissues (cutoff for significance: p<1.5×10⁻⁶). We investigated whether any lipid (LDL, HDL, or triglycerides), blood pressure (systolic blood pressure, diastolic blood pressure, pulse pressure), metabolic (type 2 diabetes, body-mass index), or lifestyle risk factors (ever smoking) have a causal impact on lacunar stroke based on genetics. Instrumental variables were independent (LD r²<0.01) of tested genes across the tissues (cutoff for significance: p<1.5×10⁻⁶). We investigated whether any lipid (LDL, HDL, or triglycerides), blood pressure (systolic blood pressure, diastolic blood pressure, pulse pressure), metabolic (type 2 diabetes, body-mass index), or lifestyle risk factors (ever smoking) have a causal impact on lacunar stroke based on genetics. Instrumental variables were independent (LD r²<0.01) of tested genes across the tissues (cutoff for significance: p<1.5×10⁻⁶). We investigated whether any lipid (LDL, HDL, or triglycerides), blood pressure (systolic blood pressure, diastolic blood pressure, pulse pressure), metabolic (type 2 diabetes, body-mass index), or lifestyle risk factors (ever smoking) have a causal impact on lacunar stroke based on genetics. Instrumental variables were independent (LD r²<0.01) of tested genes across the tissues (cutoff for significance: p<1.5×10⁻⁶). We investigated whether any lipid (LDL, HDL, or triglycerides), blood pressure (systolic blood pressure, diastolic blood pressure, pulse pressure), metabolic (type 2 diabetes, body-mass index), or lifestyle risk factors (ever smoking) have a causal impact on lacunar stroke based on genetics. Instrumental variables were independent (LD r²<0.01) of tested genes across the tissues (cutoff for significance: p<1.5×10⁻⁶). We investigated whether any lipid (LDL, HDL, or triglycerides), blood pressure (systolic blood pressure, diastolic blood pressure, pulse pressure), metabolic (type 2 diabetes, body-mass index), or lifestyle risk factors (ever smoking) have a causal impact on lacunar stroke based on genetics. Instrumental variables were independent (LD r²<0.01) of tested genes across the tissues (cutoff for significance: p<1.5×10⁻⁶). We investigated whether any lipid (LDL, HDL, or triglycerides), blood pressure (systolic blood pressure, diastolic blood pressure, pulse pressure), metabolic (type 2 diabetes, body-mass index), or lifestyle risk factors (ever smoking) have a causal impact on lacunar stroke based on genetics. Instrumental variables were independent (LD r²<0.01) of tested genes across the tissues (cutoff for significance: p<1.5×10⁻⁶). We investigated whether any lipid (LDL, HDL, or triglycerides), blood pressure (systolic blood pressure, diastolic blood pressure, pulse pressure), metabolic (type 2 diabetes, body-mass index), or lifestyle risk factors (ever smoking) have a causal impact on lacunar stroke based on genetics. Instrumental variables were independent (LD r²<0.01) of tested genes across the tissues (cutoff for significance: p<1.5×10⁻⁶). We investigated whether any lipid (LDL, HDL, or triglycerides), blood pressure (systolic blood pressure, diastolic blood pressure, pulse pressure), metabolic (type 2 diabetes, body-mass index), or lifestyle risk factors (ever smoking) have a causal impact on lacunar stroke based on genetics. Instrumental variables were independent (LD r²<0.01) of tested genes across the tissues (cutoff for significance: p<1.5×10⁻⁶). We investigated whether any lipid (LDL, HDL, or triglycerides), blood pressure (systolic blood pressure, diastolic blood pressure, pulse pressure), metabolic (type 2 diabetes, body-mass index), or lifestyle risk factors (ever smoking) have a causal impact on lacunar stroke based on genetics. Instrumental variables were independent (LD r²<0.01) of tested genes across the tissues (cutoff for significance: p<1.5×10⁻⁶). We investigated whether any lipid (LDL, HDL, or triglycerides), blood pressure (systolic blood pressure, diastolic blood pressure, pulse pressure), metabolic (type 2 diabetes, body-mass index), or lifestyle risk factors (ever smoking) have a causal impact on lacunar stroke based on genetics. Instrumental variables were independent (LD r²<0.01) of tested genes across the tissues (cutoff for significance: p<1.5×10⁻⁶).
procedure. As secondary analyses, we used median, weighted median, and MR-Egger approaches to aggregate across SNPs. We used the MR-Egger intercept to assess evidence of directional pleiotropy. In all analyses, we used the MendelianRandomization package in R. Results are presented as odds ratios (ORs) per genetically predicted increase in each risk factor (original scale).

Role of the funding source
The funder had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

Results
We meta-analysed studies from Europe, the USA, and Australia, including previous GWAS and additional cases and controls from the UK DNA Lacunar Stroke studies and the International Stroke Genetics Consortium, comprising a total of 6030 cases and 248929 controls of European ancestry, and 7338 cases and 254798 controls in the transethnic analysis (figure 1). 2987 (40.7%) cases (matched with 29540 controls) had confirmation by MRI. Study cohorts, including genomic inflation λ values, are described in the appendix (pp 5–6, 100–11). Following meta-analysis, the LD score intercept value in the European analysis was 1.046 (SE 0.008) and the λ1000 value was 1.007, whereas in the transethnic analysis, the LD score intercept value in the European analysis was 1.046 (SE 0.008) and the λ1000 value was 1.005, indicating no substantial inflation.

SNP heritability (h²) was substantial for lacunar stroke and cerebral white matter hyperintensities, calculated using LD score regression, was 0.065, and slightly lower when including all lacunar strokes (0.37, 0.09, p=4.0×10⁻⁵). In the joint analysis with cerebral white matter hyperintensities, variants in seven additional loci reached genome-wide significance for lacunar stroke overall (table, figure 2). Four of these loci (SLC25A44-PMF1-BGLAP, LOX-ZNF474-LOC100505841, SH3PXD2A, and COL4A2) have previously been associated with white matter hyperintensities.14 Regional association plots and forest plots for these loci are provided in the appendix (pp 120–134).

None of the 12 loci reaching genome-wide significance showed evidence of heterogeneity (p=0.05 to p=0.98; appendix pp 123–134). In two regions (SH3PXD2A and HTRA1-ARMS2) multiple apparently independent (LD r²<0.1) SNPs reached genome-wide significance. However, in a joint modelling scenario that was done using GCTA-cojo, only a single SNP at each of these regions had genome-wide significance, showing that a single variant remains the most parsimonious explanation of the association at this locus.13 We discarded two regions according to our protocol: one region on chromosome 1 (Fernandez-Cadenas); Experimental Neurology, Department of Neurosciences, KU Leuven, Leuven, Belgium (R Lemmens MD); VIB Center for Brain & Disease Research, Department of Neurology, University Hospitals Leuven, Leuven, Belgium (R Lemmens); Centre for Clinical Brain Sciences, UK Dementia Research Institute and Row Fogo Centre for Research into the Ageing Brain, University of Edinburgh, Edinburgh, UK (Prof J Wardlaw FRCR).

Correspondence to: Prof Hugh S Markus, Stroke Research Group, Clinical Neurosciences, University of Cambridge, Cambridge CB2 0QQ, UK hm32@medschl.cam.ac.uk See Online for appendix
Articles

Evidence of colocalisation of gene expression and lacunar stroke signals is shown by triangle size, with larger triangles indicating stronger evidence of colocalisation. CMC=CommonMind Consortium. COLOC.PP4=posterior probability of hypothesis 4 in colocalisation analysis (that there is a consistent association between lacunar stroke and expression of the given gene). GTEx=Genotype-Tissue Expression portal. WMH=white matter hyperintensities.

Figure 3: Genes for which expression is associated with lacunar stroke in six tissues from transcriptome-wide association analysis

For FUSION software see https://github.com/gusevlab/fusion_twas

17q25 showed an association solely with white matter hyperintensities, with no association with lacunar stroke (lead SNP p=0·39); a second region on chromosome 14 (EVL-DEGS2), was not as significant in MTA analysis (p=1·2×10⁻⁴) as in white matter hyperintensities alone (p=1·2×10⁻¹²), so an independent contribution of lacunar stroke to the association could not be determined. Further evidence is required to determine whether these two regions are associated with lacunar stroke, so they were discarded from this analysis. The ZBTB14-EPB41L3 locus that was associated with lacunar stroke was not associated with white matter hyperintensities (p=0·33 and effect in the opposite direction). Similarly, for the ULK4 locus associated with lacunar stroke, the lead SNP did not reach significance for white matter hyperintensities (p=0·12), but it was in the consistent effect direction and thus could reflect a lack of study power.

The 12 loci showed stronger effects in the MRI-confirmed group than in the standard phenotyping group (in European ancestry analysis, appendix p 86), although the difference was not statistically significant (one-tailed p=0·07), with a median proportional increase in OR of 3·4%. The 12 loci explained 1·4% of the overall heritability, and 6·5–8·1% of the lacunar stroke heritability from GWAS arrays, as calculated in this study.

We did a TWAS to identify genes for which expression was associated with lacunar stroke (figure 3). Genetically elevated expression of SLC25A44 was associated with lacunar stroke in multi-trait analysis in arterial tissues, whereas genetically decreased expression of ULK4 was associated with lacunar stroke in arterial tissues, whole blood, and the brain. At the 2q33·2 locus, genetically elevated expression of CARF, FAM117B, ICA1L, and NBEAL1 were all associated with lacunar stroke. All associations were confirmed by colocalisation analysis between the gene expression and lacunar stroke associations (posterior probability >0·7). Five other associations were identified in the TWAS, but these were not confirmed by colocalisation analysis (posterior probability <0·7, figure 3).

We used Phenoscanner to evaluate whether the 12 lead SNPs were associated with DNA methylation, metabolite levels, or protein levels from large-scale studies.9–11 Of the 12 lead SNPs, 11 of the 12 lead SNPs showed associations with DNA methylation at genome-wide significance, which was more than expected by chance based on randomly selected SNPs across the genome (p<0·01). Ten of the SNPs were associated in multiple independent studies (appendix pp 87–97). Conversely, none of the 12 SNPs were associated with metabolite or protein levels.

Querying databases that catalogue drug-gene relationships showed that 11 of the genes listed in table are categorised as druggable, indicating that they have potential for therapeutic development (appendix p 98). However, no existing drugs target any of the genes identified in this study.

A pathway analysis based on the multi-trait analysis results using MAGMA identified five significantly associated Gene Ontology gene sets: phosphatidyl-inositol 5 phosphate binding (p=2·2×10⁻⁶, FDR 0·020), extracellular matrix structural constituent (p=6·2×10⁻⁶, 0·027), extracellular matrix constituent conferring elasticity (p=8·9×10⁻⁶, 0·027), middle ear morphogenesis (p=2·3×10⁻⁵, 0·049), and roundabout binding (p=2·7×10⁻⁵, 0·049). No pathways were significant when based solely on lacunar stroke results. Results for all pathways with FDR of less than 0·5 are presented in the appendix (pp 75–76).

www.thelancet.com/neurology Published online March 25, 2021 https://doi.org/10.1016/S1474-4422(21)00031-4
Mendelian randomisation analyses using an inverse variance weighted approach found positive associations of diastolic, systolic, and pulse pressure, type 2 diabetes, and ever smoking with lacunar stroke (figure 4). No significant finding showed any evidence of pleiotropy, as assessed using the MR-Egger intercept. There was some evidence of a negative association between HDL and lacunar stroke, but this result did not reach Bonferroni-corrected significance. There was no evidence of an association with body-mass index, low density lipoprotein or triglycerides. Secondary analysis for all risk factors using median, weighted median, and MR-Egger approaches are presented in the appendix (pp 135–143).

Discussion

Despite the public health importance of lacunar stroke as the cause of a quarter of all strokes, previous GWAS studies have only identified one genetic locus associated with the disease, in contrast with the 35 identified for ischaemic stroke and its major subtypes.1 We did a GWAS of lacunar stroke, including the largest number of cases with MRI confirmation to date, identifying 11 novel loci in addition to replicating the one previously reported locus.

The primary analysis identified four novel loci. One association on chromosome 11, encompassing SPI1-SLC39A13-PSMC3-RAPSN, was identified in both European and transethnic analyses. The lead SNP is a synonymous variant in SLC39A13, which encodes for solute carrier family 39 member 13, a transmembrane protein with roles in zinc transport. Mutations in this gene cause a form of Ehlers-Danlos syndrome, a group of connective tissues disorders that affect the vasculature and can cause stroke; vascular abnormalities have been reported in SLC39A13-knockout mice.11 We additionally identified a locus for which the lead SNP resides in an intron of ULK4 (encoding for serine/threonine-protein kinase ULK4) on chromosome 3. The TWAS analysis suggests ULK4 is the most likely implicated gene, with genetically decreased expression of ULK4 being associated with lacunar stroke. Variants in close LD with the lead SNP have been implicated in diastolic blood pressure in large-scale GWAS.12 However, the direction of effect was opposite to that for lacunar stroke, suggesting this likely reflects pleiotropy rather than a causal pathway. Variants in close LD have also been associated with another cardiovascular disease, acute aortic dissection.13 ULK4 belongs to the family of serine-threonine protein kinases, a group of phosphorylating kinases involved in diverse processes including cell proliferation and differentiation, apoptosis, and embryonic development. Its deficiency leads to hypomyelination,14 and it has been associated with neuropsychiatric traits.15 Finally, we report a novel association on chromosome 18, located between ZBTB14, encoding a zinc finger transcription factor (zinc finger and BTB domain-containing protein 14), and EPB41L3, encoding for a membrane protein that inhibits cell proliferation and promotes apoptosis (erythrocyte membrane protein band 4·1-like 3).

In multi-trait analysis we identified seven further loci, all of which are reported as being associated with lacunar stroke at genome-wide significance for the first time. Two have not been reported as being associated with any cerebrovascular disease previously. One locus lies in an intergenic region between the VTA1 (vacuolar protein sorting-associated protein VTA1 homolog) and GPR126 (G protein-coupled receptor 126) genes. GPR126 is activated by type IV collagen and has an important role in myelination.16 GPR126 also binds laminin-211,19 an extracellular matrix protein produced by astrocytes and present in the brain, with key roles in development and function of the blood-brain barrier,40 in part through regulation of pericyte differentiation, a mechanism previously implicated through the FOXF2 gene.41 An extra-cellular matrix protein produced by astrocytes and present in the brain, with key roles in development and function of the blood-brain barrier,40 in part through regulation of pericyte differentiation, a mechanism previously implicated through the FOXF2 gene.41 The presence of both rare and common risk variants in HTRA1 points to it being a key molecule in lacunar stroke pathogenesis, and is a feature shared with another gene identified in this study, COL4A2, in which rare variants also cause monogenic forms of cerebral small vessel disease.7 Candidate gene studies have previously shown associations not reaching genome-wide significance in COL4A2 with lacunar stroke and the same region has also previously been associated with intracerebral haemorrhage in multi-trait analysis, and coronary artery disease.42 Four other loci identified (SH3PD2A, LOX-ZNF474-LOC10050841,
SLC25A44-PMF1-BGLAP, FOXF2-FOXQ1] were associated with broad stroke in MEGASTROKE (associations of all SNPs in MEGASTROKE are shown in the appendix pp 11–18) or a previous meta-analysis, although, to our knowledge, this is the first study to confirm their association specifically with lacunar stroke. At the SLC25A44-PMF1-BGLAP locus, the TWAS results point to an association of genetically elevated expression of SLC25A44 with lacunar stroke, which was validated in colocalisation analysis. SLC25A44 has an important role in catabolism of branched-chain amino acids in brown adipose tissue by transporting them into mitochondria, and thus has potential as a mediating factor in the relationship between metabolic disease and lacunar stroke. However, variants in close LD have also been associated with mosaic Y chromosome loss, highlighting mosaicism as an alternative mechanism. Further functional studies will be required to untangle these relationships with lacunar stroke. The strength of the association for all associated variants was moderate to large in the context of GWAS (OR of 1·10 to 1·25 in Europeans) and notably larger than effects previously reported for variants associated with broad stroke phenotypes. This observation is consistent with the variants acting specifically on the lacunar stroke subtype rather than on stroke as a whole.

We also found that 11 of the 12 lead SNPs influence DNA methylation at genome-wide significance, pointing to epigenetic changes being one source through which risk of lacunar stroke is conferred. Whether this genetically altered DNA methylation influences transcription of nearby genes, and which genes are affected, should be the focus of further study. A pathway analysis implicated several biological processes in lacunar stroke pathophysiology. Two pathways involved the ECM, the network of extracellular molecules that provide scaffolding and biochemical support to surrounding tissues. Disruption of the vascular ECM has been hypothesised to be a key component in pathogenesis of cerebral small vessel disease, particularly in monogenic forms, and several of the genes implicated in this study (COL4A2, LOX, SH3PXD2A, GPR126, HTRA1) have key roles in the ECM. Therefore, our findings support this hypothesis and suggest ECM dysfunction also has a key role in sporadic cerebral small vessel disease.

We did Mendelian randomisation to assess whether cardiovascular risk factors showed evidence of a causal association with lacunar stroke. We found support for genetically predicted elevated blood pressure (systolic, diastolic, and pulse pressure), type 2 diabetes, and smoking being associated with lacunar stroke. The results are consistent with those from observational studies and suggest that targeting these factors would reduce risk of lacunar stroke. There was evidence, not reaching Bonferroni-corrected significance, for a protective effect of increased HDL on risk of lacunar stroke, and no association with LDL, replicating findings in previous studies. Overall these findings show that the impact of the direct effects of LDL-lowering medications, such as statins, on the incidence of lacunar stroke is likely to be small.

Our study emphasises the benefit of accurate phenotyping using MRI. Using this approach, the heritability of lacunar stroke using GREML was substantial, and larger than previous estimates based on TOAST subtyping. Using LD score, the heritability was larger in the MRI-confirmed group, but estimates were considerably lower than for GREML. The use of MRI subtyping also increased the strength of association for the lacunar stroke-associated variants, although this increase was not statistically significant. These results suggest that further genetic risk factor studies in lacunar stroke are likely to be more successful if MRI subtyping is used.

Our study has limitations. The analysis was done in a predominantly European ancestry population. Large studies including diverse ancestries should be done to assess the generalisability of our findings to all ethnic groups. The MTAG approach relies on the relatively strong assumption that associated variants act on both traits, which might not always be the case for white matter hyperintensities or lacunar stroke alone, as being significant. However, independent replication will remain the gold standard for confirming these and all other reported associations in this Article. Previous studies have suggested that a more conservative threshold of a p value less than 1×10⁻⁸ should be considered in GWAS using larger imputation panels such as this study. If we were to use this threshold, one locus (ZBTB14-EPB41L3) would no longer be significant. Additional caution should therefore be applied when interpreting this finding, particularly as it was not significant in MTAG analysis. To increase sample size and study power, we used publicly available controls in analyses. As such, we could not determine whether these individuals had a history of lacunar stroke. Our analyses did not adjust for age and sex, and there is an ongoing debate about the importance of including such covariates in genetic studies. In analyses with substantial differences between case and control populations, subtle biases could occur.

In summary, these findings represent substantial progress in identifying the genetic mechanisms underlying lacunar stroke, a disease for which little is known about the molecular causes. Our findings highlight diverse mechanisms contributing to the disease, implicating disruption of the vascular ECM (COL4A2, LOX, SH3PXD2A, GPR126, HTRA1), pericyte differentiation (FOXF2, GPR126), TGF-β signalling (HTRA1), and myelination (ULK4, GPR126) in disease risk. These results provide novel insights into the pathogenesis of lacunar stroke, and highlight multiple candidates for functional
experiments to identify the specific mechanisms concerning the risk of lacunar stroke that could be targeted therapeutically.

Contributors
MT, CML, and HSM designed the experiment. MT did the meta-analysis and downstream analyses. HSM provided oversight. EF analysed the TWAS, L2, SK, VA, MRB, NT, SB, and YZ did the statistical analysis. LL, Omicron Trust (WT072952), with additional support from the Stroke Genetics Network study. The Edinburgh Mild Stroke Study (MS2) was funded by the Wellcome Trust (WT088134/Z/09/A). MT was supported by The Barts Charity and the National Institute of Health Research (NIHR) Barts Biomedical Research Centre. RL is a senior clinical investigator of FWO Flanders. PMR and the Oxford Vascular Study are funded by the NIHR Oxford Biomedical Research Centre and the Wellcome Trust. CML is supported by the NIHR Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King’s College London. HSM is supported by an NIHR Senior Investigator award, and his work is supported by the Cambridge Universities NIHR Comprehensive Biomedical Research Centre.

Declaration of interests
Al reports grants from the Swedish Heart and Lung Foundation, Region Skåne, Skåne University Hospital, Freemasons Lodge of Instruction Eos in Lund, Lund University, the Fär & Frosta Foundation (a Sparbanken Skåne Foundation), and the Swedish Research Council, during the conduct of the study; and personal fees from AstraZeneca, Bristol Myers Squibb/Pfizer, Portola, and Bayer, outside the submitted work. VS reports honoraria for consultation from Novo Nordisk and Sanofi and grants from Bayer, outside the submitted work. AH reports grants from the Academy of Finland, outside the submitted work. JR reports grants from the US National Institutes of Health and OneMind, during the conduct of the study, and personal fees from Boehringer Ingelheim, Pfizer, and New Beta Innovation, outside the submitted work. PMR reports personal fees from Bayer and Bristol Myers Squibb, outside the submitted work. CML reports grants from the US National Institutes of Health, the American Heart Association, Massachusetts General Hospital, and Bayer, and personal fees from ApoPharma, outside the submitted work. JW reports grants from the Wellcome Trust, during the conduct of the study. HSM reports personal fees from BIBA Medical, outside the submitted work. All other authors declare no competing interests.

Data sharing
GWAS summary statistics from these analyses are available at GWAS Catalog and on the Cerebrovascular Disease Knowledge Portal. Individual-level data from the NINDS Stroke Genetics Network study are available to researchers through dbGAP.

Acknowledgments
The UK Household Longitudinal Study is led by the Institute for Social and Economic Research at the University of Essex and funded by the Economic and Social Research Council. The survey was done by NatCen and the genome-wide scan data were analysed and deposited by the Wellcome Trust Sanger Institute. Information on how to access the data can be found on the Understanding Society website https://www.understandingsociety.ac.uk/. This research made use of the UK Biobank Resource under application number 36509. Ethical approval for UK Biobank was received from the research ethics committee (REC reference 11/NW/0382). We are grateful to deCODE genetics for providing data for this analysis. We acknowledge the contribution of Giorgio Boncoraglio (Fondazione IRCCS Institute Neurologico Carlo Besta, Milan, Italy). This work, including collection and genotyping of the UK Young Lacunar Stroke DNA Study 2 (DNA Lacunar 2), was supported by a British Heart Foundation Programme Grant (RG/16/4/32218). The NINDS Stroke Genetics Network study was funded by the US National Institute of Neurological Disorders and Stroke, National Institutes of Health (U01 NS092008 and R01 NS007078). Collection of the UK Young Lacunar Stroke DNA Study 1 (DNA Lacunar 1) was primarily supported by the Wellcome Trust (WT072952), with additional support from the Stroke Association (TSA 2010/01). Genotyping of the DNA Lacunar samples was supported by a Stroke Association Grant (TSA 2013/01). The principal funding for the WTCCC2 stroke study was provided by the Wellcome Trust, as part of the Wellcome Trust Case Control Consortium 2 project (GWAS5: B/08/Z and 98545/Z/08/Z and 98545/Z/08/Z and 98545/Z/08/Z). Collection and genotyping of the Sahlgrenska Academy Study on Ischemic Stroke were primarily supported by the Swedish Research Council (grant #2018-02534), the Swedish Heart and Lung Foundation (2019023), and the Swedish state under the agreement between the Swedish government and the county councils (the ALF agreement, ALFGBG-720008). The genetic data from Geisinger was made available through the collaboration with Regeneron Genetic Centre. The Edinburgh Mild Stroke Study (MS2) was funded by the Wellcome Trust (WT088134/Z/09/A). MT was supported by The Barts Charity and the National Institute of Health Research (NIHR) Barts Biomedical Research Centre. RL is a senior clinical investigator of FWO Flanders. PMR was primarily supported by the Swedish Research Council and the Skåne Foundation, and the Swedish Research Council in Lund, Lund University, the Färs & Frosta Foundation (a Sparbanken Skåne Foundation), and the Swedish Research Council (grant #2018–02543), the Swedish Heart and Lung Foundation.

References
1. Pantoni L. Cerebral small vessel disease: from pathogenesis and clinical characteristics to therapeutic challenges. Lancet Neurol 2010; 9: 689–701.
2. Wardlaw J, Smith C, Dichgans M. Small vessel disease: mechanisms and clinical implications. Lancet Neurol 2019; 18: 684–96.
3. Wardlaw J, Smith EE, Biessels GJ, et al. Neuroimaging standards for research into small vessel disease and its contribution to ageing and neurodegeneration. Lancet Neurol 2013; 12: 822–38.
4. Malik R, Rannikmäe K, Traylor M, et al. Genome-wide meta-analysis identifies 3 novel loci associated with stroke. Ann Neurol 2018; 84: 934–39.
5. Malik R, Chauhan G, Traylor M, et al. Multiancestry genome-wide association study of 520,000 subjects identifies 32 loci associated with stroke and stroke subtypes. Nat Genet 2018; 50: 524–37.
6. Traylor M, Malik R, Nalls MA, et al. Genetic variation at IQc24.2 is associated with small vessel stroke. Ann Neurol 2017; 83: 383–94.
7. Tan R, Traylor M, Rutten-Jacobs L, Markus H. New insights into mechanisms of small vessel disease stroke from genetics. Clin Sci (Lond) 2017; 131: 515–31.
8. Jerrard-Dunne P, Cloud G, Hassan A, Markus HS. Evaluating the genetic component of ischemic stroke subtypes: a family history study. Stroke 2003; 34: 1364–69.
9. Persyn E, Hanscom BE, Howson JMM, Lewis CM, Traylor M, Markus HS. Genome-wide association study of MRI markers of cerebral small vessel disease in 42,310 participants. Nat Commun 2020; 11: 2175.
10. Kilarski LL, Rutten-Jacobs LCA, Bevan S, et al. Prevalence of CADASIL and Fabry disease in a cohort of MRI defined younger onset lacunar stroke. PLoS One 2015; 10: e0136352.
11. Adams HP Jr, Bendixen BH, Kappelle LJ, et al. Classification of subtype of acute ischemic stroke. Definitions for use in a multicenter clinical trial. TOAST. Trial of Org 10172 in acute stroke treatment. Stroke 1993; 24: 35–41.
12. Belleguerce C, Bevan S, Grechwindt A, et al. Genome-wide association study identifies a variant in HDAC9 associated with large vessel ischemic stroke. Nat Genet 2011; 44: 328–33.
13. NINDS Stroke Genetics Network (SGN), International Stroke Genetics Consortium (ISGC). Loci associated with ischemic stroke and its subtypes (SGN): a genome-wide association study. Lancet Neurol 2016; 15: 174–84.
14. McCarthy S, Das S, Kretzschmar W, et al. A reference panel of 64,976 haplotypes for genotype imputation. Nat Genet 2016; 48: 1279–83.
15. Rajajee V, Kidwell C, Starkman S, et al. Diagnosis of lacunar infarcts within 6 hours of onset by clinical and CT criteria versus MRI. J Neuroimaging 2008; 18: 66–72.
16. Markus HS, Khan U, Birns J, et al. Differences in stroke subtypes and clinical characteristics to therapeutic challenges. Lancet Neurol 2010; 9: 689–701.
19 Wilner CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics 2010; 26: 2190–91.
20 Winkler TW, Day FR, Croteau-Chonka DC, et al. Quality control and conduct of genomewide association meta-analyses. Nat Protoc 2014; 9: 1192–212.
21 Devlin B, Roeder K. Genomic control for association studies. Biometrics 1999; 55: 997–1004.
22 Bulik-Sullivan BK, Loh PR, Finucane HK, et al. LD score regression distinguishes confounding from polygenicity in genomewide association studies. Nat Genet 2015; 47: 291–95.
23 Yang J, Ferreira T, Morris AP, et al. Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. Nat Genet 2012; 44: 369–75, S1–3.
24 Nagelkerke NJ. A note on a general definition of the coefficient of determination. Biometrika 1991; 78: 691–92.
25 Lee SH, DeCandia TR, Ripke S, et al. Estimating the proportion of variation in susceptibility to schizophrenia captured by common SNPs. Nat Genet 2012; 44: 247–50.
26 Yang J, Lee SH, Goddess ME, Vischer PM. GCTA: a tool for genomewide complex trait analysis. Am J Hum Genet 2011; 88: 76–82.
27 Turley P, Walters RK, Maghazian O, et al. Multi-variant analysis of genomewide association summary statistics using MTAG. Nat Genet 2018; 50: 229–37.
28 Turchin MC, Stephens M. Bayesian multivariate reanalysis of large genetic studies identifies many new associations. PLoS Genet 2019; 15: e1008411.
29 Gusse A, Ko A, Shi H, et al. Integrative approaches for large-scale transcriptome-wide association studies. Nat Genet 2016; 48: 245–52.
30 Battle A, Brown CD, Engelhardt BE, Montgomery SB. Genetic effects on gene expression across human tissues. Nature 2017; 556: 204–13.
31 Fromer M, Roussos P, Sieberts SK, et al. Gene expression elucidates functional impact of polygenic risk for schizophrenia. Nat Neurosci 2016; 19: 1442–53.
32 Raitakari OT, Joumala M, Rönnemaa T, et al. Cohort profile: the cardiovascular risk in Young Finns Study. Int J Epidemiol 2008; 37: 1220–26.
33 Giambartolomei C, Vuksic D, Schadt EE, et al. Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. PLoS Genet 2014; 10: e1004383.
34 Staley JR, Blackshaw J, Kamat MA, et al. PhenoScanner: a database of human genotype-phenotype associations. Bioinformatics 2016; 32: 3207–09.
35 Kamat MA, Blackshaw JA, Young R, et al. PhenoScanner V2: an expanded tool for searching human genotype-phenotype associations. Bioinformatics 2019; 35: 4851–53.
36 Wishart DS, Feunang YD, Guo AC, et al. DrugBank 5.0: a major update to the DrugBank database for 2018. Nucleic Acids Res 2018; 46: D1074–82.
37 Cotto KC, Wagner AH, Feng YY, et al. DGIdb 3.0: a redesign and expansion of the drug-gene interaction database. Nucleic Acids Res 2018; 46: D1068–73.
38 de Leeuw CA, Mouji JM, Heskes T, Posthuma D. MAGMA: generalized gene-set analysis of GWAS data. PLoS Comput Biol 2011; 7: e1002419.
39 Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci USA 2005; 102: 15545–50.
40 Teslovich TM, Musunuru K, Smith AV, et al. Biological, clinical and population relevance of 95 loci for blood lipids. Nat Genet 2010; 42: 199–206.
41 Hoffmann TJ, Ehret GB, Nandakumar P, et al. Genome-wide association studies for height and body mass index in 70,000 individuals of European ancestry. Hum Mol Genet 2018; 27: 3641–49.
42 Mahajan A, Talukd A, Thurner M, et al. Fine-mapping type 2 diabetes loci to single-variant resolution using high-density imputation and sLetter-specific epigenome maps. Nat Genet 2018; 50: 1209–13.
43 Liu M, Jiang Y, Wedow R, et al. Association studies of up to 1.2 million individuals yield new insights into the genetic etiology of tobacco use and dependence. Nat Genet 2019; 51: 237–44.
44 Haycock PC, Burgess S, Wade KH, Bowden J, Relton C, Davey Smith G. Best (but oft-forgotten) practices: the design, analysis, and interpretation of Mendelian randomization studies. Am J Clin Nutr 2016; 103: 965–79.
45 Yavorska OO, Burgess S. MendelianRandomization: an R package for performing Mendelian randomization analyses using summarized data. Int J Epidemiol 2017; 46: 1734–39.
46 Chung J, Martini S, Perera J, et al. Genome-wide association study of cerebral small vessel disease reveals established and novel loci. Brain 2019; 142: 1176–89.
47 Verhaaren BF, Debette S, Bis JC, et al. Multiethnic genome-wide association study of cerebral white matter hyperintensities on MRI. Circ Cardiovasc Genet 2015; 8: 398–409.
48 Bonder MJ, Luijk R, Zherkova DV, et al. Disease variants alter transcription factor levels and methylation of their binding sites. Nat Genet 2017; 49: 131–9.
49 Gaunt TR, Shihaib HA, Hemani G, et al. Systematic identification of genetic influences on methylation across the human life course. Genome Biol 2016; 17: 61.
50 Chen L, Ge B, Casale FP, et al. Genetic drivers of epigenetic and transcriptional variation in human immune cells. Cell 2016; 167: 1398–1414.e214.
51 Bir BH, Hoyou S, Hosaka T, et al. Molecular pathogenesis of spondylohyperostotic Ehlers-Danlos syndrome caused by mutant ZIP13 proteins. EMBO Mol Med 2014; 6: 1028–42.
52 Hirose T, Shimazaki T, Takahashi N, et al. Morphometric analysis of thoracic aorta in Slc39a13/Zip13-KO mice. Cell Tissue Res 2019; 376: 137–41.
53 Ehret GB, Munroe PB, Rice KM, et al. Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. Nature 2011; 478: 101–09.
54 Guo DC, Grove ML, Prakash SK, et al. Genetic variants in LRP1 and ULK4 are associated with acute aortic dissections. Am J Hum Genet 2016; 99: 762–69.
55 Liu M, Xu P, Guan Z, et al. ULK4 deficiency leads to hypomyelination in mice. Glia 2018; 66: 175–90.
56 Liu M, Xu P, O’Brien T, Shen S. Multiple roles of ULK4 in neurogenesis and brain function. Neurongenesis (Austin) 2017; 4: e131364.
57 Paavola KJ, Sidik H, Zuchero JB, Eckart M, Talbot WS. Type IV collagen is an activating ligand for the adhesion G protein-coupled receptor GPR126. Sci Signal 2014; 7: ra76.
58 Mehta P, Piao X. Adhesion G-protein coupled receptors and extracellular matrix proteins: Roles in myelination and glial cell development. Dev Dyn 2017; 246: 275–81.
59 Menezes MJ, McClenahan FK, Leiton CV, Aranamolte A, Shan X, Colognato H. The extracellular matrix protein laminin α2 regulates the maturation and function of the blood-brain barrier. J Neurosci 2014; 34: 15260–80.
60 Yao Y, Chen ZL, Norris EH, Strickland S. Astrocytic laminin regulates pericyte differentiation and maintains blood brain barrier integrity. Nat Commun 2014; 5: 3413.
61 Chauhan G, Arnold CR, Chu AV, et al. Identification of additional risk loci for stroke and small vessel disease: a meta-analysis of genomewide association studies. Lancet Neurol 2016; 15: 695–707.
62 Rajani RM, Quick S, Ruigrok SR, et al. Reversal of endothelial dysfunction reduces white matter vulnerability in cerebral small vessel disease in rats. Sci Transl Med 2018; 10: eaar9507.
63 Harra K, Shiga A, Fukutake T, et al. Association of HTRA1 mutations and familial ischemic cerebral small-vessel disease. N Engl J Med 2009; 360: 1729–39.
64 Beaufort N, Scharer E, Kremmer E, et al. Cerebral small vessel disease-related protease HTRA1 processes latent TGF-β binding protein 1 and facilitates TGF-β signaling. Proc Natl Acad Sci USA 2014; 111: 16496–501.
65 Rannikmae K, Sivakumar V, Millar H, et al. COJ4A2 is associated with lacunar ischemic stroke and deep ICH: meta-analyses among 21,500 cases and 40,600 controls. Neurology 2017; 88: 1829–39.
67 Nikpay M, Goel A, Won HH, et al. A comprehensive 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease. Nat Genet 2015; 47:1121–30.
68 Yoneshiro T, Wang Q, Tajima K, et al. BCAA catabolism in brown fat controls energy homeostasis through SLC25A44. Nature 2019; 572:614–19.
69 Wright DJ, Day FR, Kerrison ND, et al. Genetic variants associated with mosaic Y chromosome loss highlight cell cycle genes and overlap with cancer susceptibility. Nat Genet 2017; 49:674–79.
70 Joutel A, Haddad I, Ratelade J, Nelson MT. Perturbations of the cerebrovascular matrixome: a convergent mechanism in small vessel disease of the brain? J Cereb Blood Flow Metab 2016; 36:143–57.
71 Rutten-Jacobs LCA, Markus HS. Vascular risk factor profiles differ between magnetic resonance imaging-defined subtypes of younger-onset lacunar stroke. Stroke 2017; 48:2405–11.
72 Hindy G, Engström G, Larsson SC, et al. Role of blood lipids in the development of ischemic stroke and its subtypes: a mendelian randomization study. Stroke 2018; 49:820–27.
73 Georgakis MK, Malik R, Anderson CD, Parhofer KG, Hopewell JC, Dichgans M. Genetic determinants of blood lipids and cerebral small vessel disease: role of high-density lipoprotein cholesterol. Brain 2020; 143:597–610.
74 Bevan S, Traylor M, Adib-Samii P, et al. Genetic heritability of ischemic stroke and the contribution of previously reported candidate gene and genomewide associations. Stroke 2012; 43:3161–67.
75 Wu Y, Zheng Z, Visscher PM, Yang J. Quantifying the mapping precision of genomewide association studies using whole-genome sequencing data. Genome Biol 2017; 18:86.
76 Pirinen M, Donnelly P, Spencer CC. Including known covariates can reduce power to detect genetic effects in case-control studies. Nat Genet 2012; 44:848–51.