Original Contribution

Genetic Architecture of Lacunar Stroke

Matthew Traylor, PhD; Steve Bevan, PhD; Jean-Claude Baron, ScD; Ahamad Hassan, MD; Cathryn M. Lewis, PhD; Hugh S. Markus, DM

Background and Purpose—Lacunar strokes comprise ≈20% of all strokes. Despite this frequency, their pathogenesis is poorly understood. Previous genome-wide association studies in lacunar stroke have been disappointing, which may be because of phenotypic heterogeneity. Pathological and radiological studies suggest that there may be different pathologies underlying lacunar strokes. This has led to the suggestion of 2 subtypes: isolated lacunar infarcts and multiple lacunar infarcts and leukoaraiosis.

Methods—We performed genome-wide analyses in a magnetic resonance imaging–verified cohort of 1012 younger onset lacunar stroke cases and 964 controls. Using these data, we first estimated the heritability of lacunar stroke and its 2 hypothesized subtypes, and secondly, we determined whether this is enriched for regulatory regions in the genome, as defined by data from Encyclopedia of DNA Elements (ENCODE) and other sources. Finally, we determine the evidence for a polygenic contribution from rare variation to lacunar stroke and its subtypes.

Results—Our results indicate a substantial heritable component to magnetic resonance imaging–verified lacunar stroke (20%–25%) and its 2 subtypes (isolated lacunar infarct, 15%–18%; multiple lacunar infarcts/leukoaraiosis, 23%–28%). This heritable component is significantly enriched for sites affecting expression of genes. In addition, we show that the risk of the 2 subtypes of lacunar stroke in isolation, but not in combination, is associated with rare variation in the genome.

Conclusions—Lacunar stroke, when defined on magnetic resonance imaging, is a highly heritable complex disease. Much of this heritability arises from regions of the genome affecting gene regulation. Rare variation affects 2 subtypes of lacunar in isolation, suggesting that they may have distinct genetic susceptibility factors. (Stroke. 2015;46:00-00. DOI: 10.1161/STROKEAHA.115.009485.)

Key Words: genetics ■ genetic association studies ■ magnetic resonance imaging ■ polymorphism, single nucleotide ■ stroke, lacunar

Lacunar infarcts resulting from occlusion of the small penetrating arteries of the brain comprise ≈20% of all ischemic strokes, a proportion similar to those resulting from cardioembolism or large artery atherosclerosis.1 Despite this, comparatively little is known about their underlying pathogenesis. Epidemiological studies have established hypertension and diabetes mellitus as important risk factors;2,3 but pathological studies have been hampered by methodological inconsistencies and inadequate classification of the disease, as well as limited pathological tissue because of the low early mortality rate.5

One approach to identifying the underlying causes of complex diseases, such as lacunar stroke, is through genetic studies. In the recent past, genome-wide association studies (GWAS) have transformed our understanding of complex diseases and have begun to identify the common genetic component to ischemic and hemorrhagic strokes.5,7 Despite these advances, no genetic variants have yet been identified that specifically confer risk of lacunar stroke in white populations. Indeed, although family history data suggest that genetic predisposition may be particularly important for lacunar stroke,8 estimates of the heritability of lacunar stroke from GWAS have been low compared with those of the other subtypes.9,10 Multiple factors might explain this disparity.

One factor which might be important is disease heterogeneity. Pathological studies have shown different vascular lesions in patients presenting with lacunar stroke, with 2 main pathologies reported, namely focal microatheroma and a diffuse small-vessel arteriopathy.11 The former has been associated with larger single lacunar infarcts and the latter with multiple...
smaller lacunes and leukoaraiosis. These 2 subtypes have also been shown to have differing risk factor profiles. These sources point to the existence of pathophysiological subtypes of lacunar stroke, each of which might be presumed to have distinct genetic susceptibility factors. Another factor might be inadequate disease classification. Lacunar infarcts are small and frequently not seen on computed tomography. Despite this, in GWAS, to date, lacunar stroke has been often been diagnosed based on computed tomography, with a clinical lacunar syndrome and no infarcts visible on computed tomography being used as criteria for lacunar stroke. It has been shown that this can lead to a marked overdiagnosis of lacunar stroke.

Diagnostic accuracy is much improved with magnetic resonance imaging (MRI).

In this study, we use a large, highly phenotyped cohort of MRI-confirmed lacunar stroke cases and controls to investigate the genetic architecture of lacunar stroke and its subtypes. We investigate whether MRI-confirmed lacunar stroke and its subtypes are heritable and whether this variation in enriched sites in the genome affecting the expression and regulation of genes. Finally, we investigate whether the risk of the 2 subtypes of lacunar stroke is conferred through rare genetic variation.

Materials and Methods

Study Population

A total of 1029 white patients with lacunar stroke, aged ≤70 years, were recruited from 72 specialist’s stroke centers throughout the United Kingdom (online-only Data Supplement), between 2002 and 2012, as part of the Young Lacunar Stroke DNA Resource. This study was approved by the Multi-Center Research Ethics Committee for Scotland (04/MRE00/36), and informed consent was obtained from all participants. Lacunar stroke was defined as a clinical lacunar syndrome, with an anatomically compatible lesion on MRI (subcortical infarct, ≤15 mm in diameter). Time from event to MRI was variable; the median time to MRI was 5 days. All patients underwent full stroke investigation, including brain MRI, imaging of the carotid arteries, and ECG. Echocardiography was performed when appropriate. All MRIs and clinical histories were reviewed centrally by 1 physician (H.S.M.). Exclusion criteria were stroke >50% in the extra- or intracranial cerebral vessels, or previous carotid endarterectomy; cardioembolic source of stroke, defined according to the Trial of Org 10172 in Acute Stroke Treatment (TOAST) criteria as high or moderate probability; cortical infarct on MRI; subcortical infarct >15 mm in diameter, as these can be caused by embolic mechanisms (striatocapsular infarcts); and any other specific cause of stroke (eg, lupus anticoagulant, cerebral vasculitis, dissection, and monogenic cause of stroke). All cases were screened for NOTCH3 cerebral autosomal-dominant arteriopathy with subcortical infarcts and leukoencephalopathy and Fabry disease mutations, and positive cases were excluded. An additional 82 white patients of all ages with lacunar stroke were recruited from St George’s Hospital, London. The same investigations and exclusions were made as in the DNA-lacunar study.

Unrelated white controls, free of clinical cerebrovascular disease, were obtained by random sampling, stratified for age and sex, from general practice lists from the same geographical location as the patients. All patients and controls underwent a standardized clinical assessment and completed a standardized study questionnaire. MRI was not performed in controls.

The data set was genotyped on the Illumina HumanExomeCore array, which contains both exome content (≈250000 single-nucleotide polymorphisms [SNPs]) and common tag SNPs (≈250100 SNPs) found on conventional GWAS arrays, and imputed to 1000-genome phase 1. Full details are provided in the online-only Data Supplement.

Risk Factors

Hypertension was defined as elevation of systolic blood pressure of >140 mm Hg or diastolic blood pressure of >90 mm Hg persisting >7 days after stroke onset or before stroke treatment with antihypertensive drugs. Diabetes mellitus was defined as a previous diagnosis of type I or type II diabetes mellitus or at least 2 random glucose readings of >11.1 mmol/L or fasting blood glucose readings of >7.0 mmol/L after the acute phase of stroke. Hypercholesterolemia was defined as serum cholesterol of >5.2 mmol/L or prestroke treatment with a cholesterol-lowering agent. A positive smoking history was recorded in those who had smoked at any time in their lives.

Subtyping of Lacunar Stroke

Leukoaraiosis was graded on MRI using the semiquantitative Fazekas scale, which has been shown to reflect pathological severity of small-vessel disease in a postmortem validation study. On the basis of the leukoaraiosis grade, patients were subtyped into 2 groups: (1) isolated lacunar infarct (ILI): single lacunar infarct with absent or mild leukoaraiosis (equivalent to Fazekas periventricular score of <2); (2) multiple lacunar infarcts (MLI) or lacunar infarct with moderate or severe confluent leukoaraiosis (equivalent to Fazekas grade of ≥2) according to a previously validated method. Twenty MRI scans were randomly selected on a second occasion by the same rater, and there was perfect agreement in assignment of subtype (κ=1).

Heritability Estimates

To assess the heritability of lacunar stroke, we first set to missing all imputed genotypes with a probability of <0.9 and discarded all SNPs that met criteria in <90% of individuals. We then calculated the genetic relationships between all individuals across all 8122203 remaining SNPs using the GCTA package. After removing distantly related individuals (>0.125), we used genetic restricted maximum likelihood (GREML) methods to estimate the proportion of phenotypic variance on the liability scale explained by the genetic relationships between individuals based on common SNPs (here termed heritability), as implemented in the GCTA package. We performed the analysis for 3 phenotypes: first for all MRI-defined lacunar stroke cases versus controls, then for cases with ILI versus controls, and for cases with MLI or extensive leukoaraiosis versus controls. We included the first principal component as a covariate in the model in all analyses. We calculated the heritability for prevalence of stroke (K) of 1% and 3%, assuming that lacunar strokes comprise 20% of all cases.

Heritability Estimates, Partitioned on Functional Status

Recently, the Encyclopedia of DNA Elements (ENCODE) project has generated a huge weight of data describing functional sites in the human genome. The project aimed to identify all functional sites in the genome through a series of experiments in many tissue types, including chromatin immunoprecipitation sequencing, DNase I hypersensitive sites sequencing, and formaldehyde-assisted isolation of regulatory elements sequencing. Each technique uses a distinct approach to identify the location of regulatory regions in the genome. This is of particular interest to genetic studies because previous analyses have shown that GWAS associations are enriched for such functional sites. In addition, genotype-tissue expression studies provide a complementary approach to identifying SNP variants that affect expression of genes. In such studies, mRNA expression levels of genes are compared with SNP genotypes, thereby determining SNPs that affect expression. Such SNPs are often termed expression quantitative trait loci (eQTLs).

In this experiment, we investigated whether the heritability of lacunar stroke was enriched for regulatory sites in the autosome. To do this, we used information from the RegulomeDB database, which catalogues data from the ENCODE project and others, determining the evidence that each SNP in the genome affects the regulation of genes. The database separates SNPs into categories based...
on the available evidence. The group with the strongest evidence, which we term eQTLs, includes SNPs that have been shown to affect the levels of an mRNA molecule in any tissue and overlap any transcriptional factor–binding site, transcription factor motif, DNase footprint, or DNase peak from ENCODE or other experiments. We first partitioned our data on this group, including all such SNPs, as well as tagging SNPs with $r^2 > 0.98$, based on linkage disequilibrium from European samples from the 1000-genome data set. A total of 24,722 SNPs were included.

Secondly, we partitioned our data on regulatory regions from RegulomeDB. This group includes all SNPs that overlap a transcription factor–binding site and DNase peak, as well as either having a matched transcription factor motif or a matched DNase footprint. Therefore, these SNPs represent regions where regulatory factors are thought to bind to the genome. As before, we included all such SNPs, as well as tagging SNPs with $r^2 > 0.98$. A total of 938,693 SNPs were included.

We then used the GCTA package to calculate the heritability, partitioned on (1) eQTLs and (2) regulatory regions for lacunar stroke and its 2 subtypes (ILI and MLI/leukoaraiosis). We compared our estimates with the proportion of overall heritability that would be expected for the number of SNPs analyzed.

Polygenic Contribution From Rare Variation

If we define protective variants as those where the major allele is associated with disease and risk variants as those where the minor allele is associated with disease, then under the null, an equal proportion of genetic associations should be from either protective or risk variants. An increase in the ratio of risk to protective variants at low allele frequencies can indicate a polygenic contribution from low-frequency variants. We calculated the ratio of risk to protective variants at low allele frequencies between 1% to 5% and 5% to 10% separately. In addition, we evaluated the enrichment for risk allele frequencies between 30% and 50% and 50% to 100% separately. We tabulated the number of simulations in which the risk to protective ratio was greater than that observed in our data and divided by the number of simulations to generate an empirical $P$ value.

We first analyzed all lacunar stroke cases versus controls and then performed analyses for the MLI/leukoaraiosis and ILI subtypes versus controls. We evaluated the enrichment for risk allele frequencies between 1% to 5% and 5% to 10% separately. In addition, we evaluated the enrichment for risk allele frequencies between 30% and 50% and 50% to 100% separately. All analyses were performed using the pisa java package.

Results

Study Characteristics

After quality control steps, a total of 1012 cases and 964 controls remained for analysis. Characteristics of this cohort are given in Table 1.

Heritability Estimates

We estimated the proportion of variance of lacunar stroke status explained by the relatedness between individuals using GREML methods, as implemented in the GCTA package. All analyses showed that lacunar stroke and its subtypes were significantly heritable (Table 2). We determined the heritability of lacunar stroke to be 0.20 (0.064) assuming prevalence of 0.2% and 0.25 (0.080) assuming prevalence of 0.6%. For the subtypes of lacunar stroke, heritability estimates were higher for the MLI/leukoaraiosis subtype and slightly lower for the ILI subtype, although this difference was not significant ($P > 0.05$). We performed sensitivity analyses to determine the

Table 1. Cohort Characteristics

<table>
<thead>
<tr>
<th></th>
<th>All Lacunar</th>
<th>ML/LA Subtype</th>
<th>ILI Subtype</th>
<th>Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>1012</td>
<td>502</td>
<td>501</td>
<td>964</td>
</tr>
<tr>
<td>Age, mean (SD)</td>
<td>57.2 (9.6)</td>
<td>59.5 (9.1)</td>
<td>54.7 (9.2)</td>
<td>59.7 (4.3)</td>
</tr>
<tr>
<td>Men, %</td>
<td>69.8</td>
<td>71.6</td>
<td>68.3</td>
<td>52.5</td>
</tr>
<tr>
<td>Hypertension, %</td>
<td>70.8</td>
<td>79.2</td>
<td>62.9</td>
<td>53.3</td>
</tr>
<tr>
<td>Diabetes mellitus, %</td>
<td>16.7</td>
<td>19.0</td>
<td>14.2</td>
<td>7.7</td>
</tr>
<tr>
<td>Hyperlipidemia, %</td>
<td>68.1</td>
<td>70.6</td>
<td>65.4</td>
<td>58.1</td>
</tr>
<tr>
<td>Ever smoker, %</td>
<td>69.6</td>
<td>75.0</td>
<td>64.1</td>
<td>56.0</td>
</tr>
<tr>
<td>Fazekas LA grade, mean (SD)</td>
<td>0.96 (1.07)</td>
<td>1.72 (1.00)</td>
<td>0.21 (0.42)</td>
<td>NA</td>
</tr>
</tbody>
</table>

ILI indicates isolated lacunar infarct; LA, leukoaraiosis; and MLI, multiple lacunar infarcts.

Table 2. Heritability Estimates, Based on Assumption of Prevalence of Lacunar Stroke ($K$)

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>$K=0.2%$</th>
<th>$K=0.6%$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$h^2$ (SE)</td>
<td>$h^2$ (SE)</td>
</tr>
<tr>
<td>Lacunar stroke</td>
<td>0.20 (0.064)</td>
<td>0.25 (0.080)</td>
</tr>
<tr>
<td>ML/LA subtype</td>
<td>0.23 (0.087)</td>
<td>0.28 (0.11)</td>
</tr>
<tr>
<td>ILI subtype</td>
<td>0.15 (0.084)</td>
<td>0.18 (0.10)</td>
</tr>
</tbody>
</table>

ILI indicates isolated lacunar infarct; LA, leukoaraiosis; and MLI, multiple lacunar infarcts.
influence of the genetic relatedness threshold and number of principal components included in the model. Both parameters had minimal influence on the results (Tables I and II in the online-only Data Supplement).

Heritability Estimates, Partitioned on Functional Status
We next estimated the heritability of lacunar stroke and its subtypes explained by eQTLs and regulatory regions. An excess of heritability was explained by either eQTLs or regulatory regions when compared with what would be expected by chance. Partitioning on eQTLs, we found that 2.3% of the heritability of lacunar stroke was explained by these SNPs (Figure). This value is equivalent to 11.5% of the total heritability, meaning that although eQTLs only make up 3.0% of all SNPs, they explain a considerable proportion of the heritability. Similar results were obtained for the subtypes.

Similarly, when partitioning on regulatory regions, 6.0% of the heritability lacunar stroke status was explained by the SNPs, equating to 30.0% of the heritability from only 11.6% of the SNPs. Similar results were again obtained for the subtypes.

Polygenic Contribution From Rare Variation
We used the pisa package to calculate the excess of risk to protective variants with $P<0.05$ at low allele frequencies (1%–5% and 5%–10%, separately).31 Our simulations indicate an excess of risk to protective variants at $P<0.05$ for frequencies between 5% and 10% in the 2 subtypes of lacunar stroke but not in all lacunar stroke itself (ILI subtype, $P=0.001$; MLI/leukoaraiosis subtype, $P=0.004$; lacunar, $P=0.22$; Table 3). We also found a significant association with the MLI/leukoaraiosis subtype at frequencies between 1% and 5% but not with the ILI subtype or all lacunar strokes (ILI, $P=0.19$; MLI/leukoaraiosis, $P=0.007$; lacunar, $P=0.15$). The results indicate that there is a polygenic contribution from rare variants to the 2 subtypes of lacunar stroke, but this cannot be detected when all lacunar strokes are considered together. This suggests that distinct pathophysiological mechanisms lead to the 2 diseases.

Discussion
Using a genome-wide approach from a large younger onset lacunar stroke population, our results show that lacunar stroke, when verified by MRI, is highly heritable. Our estimates are greater than those previously reported from GWAS data for lacunar stroke,9,10 suggesting that detailed phenotyping of cases, including MRI, is important for identification of genetic associations with the disease. In addition, we show that 2 subtypes of lacunar stroke, ILI and MLI/leukoaraiosis, are also highly heritable. Estimates of heritability for the MLI/leukoaraiosis subtype were higher, which may indicate a stronger genetic component, although more evidence is needed to determine this. The estimates of heritability for each of the analyses are comparable with those for Alzheimer disease (24%),36 schizophrenia (23%),37 and multiple sclerosis (30%),38 in which large-scale GWAS have been highly successful.

We also show that a significant proportion of the heritability of lacunar stroke, and each of its subtypes, is from SNPs affecting the regulation of genes. This was true for both

Table 3. Evidence for Contribution of Rare Variants to Disease Risk

<table>
<thead>
<tr>
<th>Allele Frequency, %</th>
<th>$P$ Value Cutoff</th>
<th>Lacunar Stroke</th>
<th>MLI/LA Subtype</th>
<th>ILI Subtype</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$O$ (R/P)</td>
<td>$E$ (R/P)</td>
<td>$P$ Value</td>
</tr>
<tr>
<td>1–5</td>
<td>0.05</td>
<td>1.02</td>
<td>0.99</td>
<td>0.15</td>
</tr>
<tr>
<td>5–10</td>
<td>0.05</td>
<td>0.99</td>
<td>0.99</td>
<td>0.56</td>
</tr>
<tr>
<td>30–50</td>
<td>0.05</td>
<td>1.00</td>
<td>1.00</td>
<td>0.51</td>
</tr>
</tbody>
</table>
eQTLs, where a SNP has been shown to directly affect the expression of a given gene, and for regulatory regions, where experiments from ENCODE indicate a region where regulatory factors bind to the genome. This is an important finding because it indicates a mechanism by which genetic variation leads to increased risk of lacunar stroke. Our findings mirror those from other complex diseases\textsuperscript{26-28} and support the notion that much of the genetic variation in liability to lacunar stroke is through subtle differences in gene expression and regulation.

Finally, our results indicate that rare genetic variants contribute to risk of lacunar stroke subtypes, ILI and MLI/leukoaraiosis, but do not have a significant effect on lacunar stroke as a whole. The strongest evidence was for the MLI/leukoaraiosis subtype, where significant enrichment of risk variants was observed for SNPs with a minor allele frequency of \textless10\%. A significant result was only observed in the ILI subtype for SNPs with minor allele frequency between 5\% and 10\%. This might be because of power, as there is a greater number of SNPs in this frequency band. This result is important for several reasons. Firstly, it suggests that there exist genetic associations that are specific to these subtypes of lacunar stroke and, therefore, that some of the genetic contribution is not shared between the 2 diseases. This has important consequences for future studies. First of all, it shows that for genetic studies of lacunar stroke, detailed phenotyping is important. By splitting analyses into the 2 subtypes of lacunar stroke, novel associations might be identified. Secondly, as the focus of genetic studies turns from common to rare variants, our results show that the greatest benefits will be reaped from detailed phenotyping of lacunar stroke populations.

Our study has several strengths. All data in this multicenter study were prospectively collected using uniform data collection proformas. MRI was used in all cases to confirm lacunar stroke and was centrally reviewed by a single rater. Twenty randomly selected scans showed perfect agreement for determination of lacunar stroke subtype, indicating high reliability. Similarly, our study has limitations. The sample size used was relatively small for genetic studies, meaning that confidence intervals around estimates of heritability were moderately large. Similarly, we were underpowered to directly estimate the genetic correlation between the 2 subtypes based on all SNPs using GREML approaches (18\% power to detect genetic correlation=0.5, using GCTA-GREML power calculator).\textsuperscript{38} An important extension of this work would be to perform such an analysis in a larger population with sufficient power to determine the degree to which the 2 subphenotypes share pathogenesis. In addition, our assessment of the proportion of heritability explained by eQTLs is limited by the information currently available. As eQTL studies grow in size, more SNPs will be identified that affect mRNA expression, and this will likely affect our results. Another important point to consider is that we were unable to obtain MRIs for the controls. It is possible that a proportion of these might have had silent subcortical infarcts, which may have a small effect of the results. Finally, the GREML approach used to estimate heritability has limitations. The estimates of heritability are derived from common SNPs meaning that the contributions from rare and structural variations are underestimated because of incomplete tagging of causal variation. Similarly, a proportion of the heritability might be because of susceptibility to risk factors for lacunar stroke, such as hypertension. An extension to this work would be to estimate the proportion of the observed heritability that acts through susceptibility to such risk factors.

In summary, we show that lacunar stroke, when diagnosed using MRI and detailed phenotyping, is highly heritable and that much of this heritability can be partitioned on regions of the genome affecting the regulation of genes. Our results suggest that rare variation affects 2 subtypes of lacunar isolation, but not with lacunar stroke as a whole, suggesting that these 2 subtypes might have distinct genetic susceptibility factors.

## Acknowledgments

A full list of the centers from which patients were recruited for DNA-lacunar study is given in the online-only Data Supplement.

## Sources of Funding

H.S. Markus was supported by a National Institute for Health Research (NIHR) Senior Investigator award. H.S. Markus and S. Bevan were supported by the NIHR Cambridge University Hospitals Comprehensive Biomedical Research Centre. Collection of the UK Young Lacunar Stroke Resource was primarily supported by the Wellcome Trust with additional support from the Stroke Association. Genotyping and M. Traylor were supported by a project grant from the Stroke Association (TSA 2013/01).

## Disclosures

None.

## References


Neuroimaging wide complex trait analysis.

2012;44:1147–1151. doi: 10.1038/ng.2397.

Common variants at 6p21.1 are associated with large artery atherosclerosis.


JAMA evaluation, and treatment of high blood cholesterol in adults (adult treatment panel iii).


2012;491:57–74. doi: 10.1038/nature10994.


2012;43161–3167. doi: 10.1161/STROKEAHA.112.665760.


Genetic Architecture of Lacunar Stroke
Matthew Traylor, Steve Bevan, Jean-Claude Baron, Ahamad Hassan, Cathryn M. Lewis and Hugh S. Markus

Stroke. published online August 4, 2015;
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/early/2015/08/04/STROKEAHA.115.009485
Free via Open Access

Data Supplement (unedited) at:
http://stroke.ahajournals.org/content/suppl/2015/08/04/STROKEAHA.115.009485.DC1
http://stroke.ahajournals.org/content/suppl/2016/04/06/STROKEAHA.115.009485.DC2

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/
Online Supplemental Data
Genotyping and imputation

The genetic dataset, which included other individuals not eligible for this analysis, was genotyped on the Illumina HumanExomeCore array. This array contains both exome content (~250,000 SNPs) and common tag SNPs (~250,000 SNPs) found on conventional GWAS arrays. Post-imputation, this array provides comparable coverage of common SNPs to larger arrays (MAF>5%, 78.2% of SNPs covered with $r^2 \geq 0.8$ compared to 86.6% for Illumina OmniExpress).\(^1\) SNPs were excluded with MAF<0.01, genotype missingness>3%, HWE p<1e-6 in controls, strand ambiguity (A/T or C/G) or evidence of differential missingness by case-control status (p<0.05). Individuals were excluded if they had missingness>3%, excess or reduced heterozygosity, showed evidence of relatedness with another individual (pi-hat>0.1875), or failed a “sex-check” in PLINK. EIGENSTRAT was used to remove non-caucasian individuals, and was then repeated to calculate ancestry-informative principal components.\(^2\) The remaining 269,691 autosomal SNPs and 2,603 individuals were then imputed to 1000 Genomes phase 1 (March 2012):\(^3\) SHAPEIT v2 was used to phase the haplotypes and IMPUTE v2.2.2 was used to perform the imputation,\(^4\)\(^,\)^\(^5\) resulting in 9,289,526 single nucleotide polymorphisms (SNPs).
**Supplementary Table I** - Sensitivity Analyses, estimating heritability for different numbers of principal components (K=0.2% throughout)

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>N (case/control)</th>
<th>$h^2$ (SE), 1 PC</th>
<th>$h^2$ (SE), 5 PCs</th>
<th>$h^2$ (SE), 10 PCs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lacunar stroke</td>
<td>1012 / 964</td>
<td>0.205 (0.064)</td>
<td>0.208 (0.065)</td>
<td>0.208 (0.065)</td>
</tr>
<tr>
<td>MLI / LA subtype</td>
<td>502 / 964</td>
<td>0.233 (0.087)</td>
<td>0.243 (0.087)</td>
<td>0.247 (0.087)</td>
</tr>
<tr>
<td>ILI subtype</td>
<td>501 / 964</td>
<td>0.152 (0.084)</td>
<td>0.158 (0.084)</td>
<td>0.153 (0.085)</td>
</tr>
</tbody>
</table>

SE, standard error; K, prevalence of lacunar stroke; ILI, isolated lacunar infarct; MLI, multiple lacunar infarcts; LA, leukoaraiosis
**Supplementary Table II** - Sensitivity Analyses, estimating heritability for different thresholds of genetic relatedness (K=0.2% throughout)

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>$h^2$ (SE) for given GRM relatedness threshold</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.125</td>
</tr>
<tr>
<td></td>
<td>0.10</td>
</tr>
<tr>
<td></td>
<td>0.075</td>
</tr>
<tr>
<td></td>
<td>0.05</td>
</tr>
<tr>
<td>Lacunar stroke</td>
<td>0.205 (0.064)</td>
</tr>
<tr>
<td></td>
<td>0.203 (0.065)</td>
</tr>
<tr>
<td></td>
<td>0.203 (0.065)</td>
</tr>
<tr>
<td></td>
<td>0.201 (0.065)</td>
</tr>
<tr>
<td>MLI / LA subtype</td>
<td>0.233 (0.087)</td>
</tr>
<tr>
<td></td>
<td>0.232 (0.087)</td>
</tr>
<tr>
<td></td>
<td>0.232 (0.087)</td>
</tr>
<tr>
<td></td>
<td>0.228 (0.087)</td>
</tr>
<tr>
<td>ILI subtype</td>
<td>0.152 (0.084)</td>
</tr>
<tr>
<td></td>
<td>0.150 (0.084)</td>
</tr>
<tr>
<td></td>
<td>0.150 (0.084)</td>
</tr>
<tr>
<td></td>
<td>0.156 (0.084)</td>
</tr>
</tbody>
</table>

SE, standard error; K, prevalence of lacunar stroke; ILI, isolated lacunar infarct; MLI, multiple lacunar infarcts; LA, leukoaraiosis
References
Appendix

UK Young Lacunar Stroke DNA Study (DNA Lacunar)

Study managers: Josie Monaghan; Alan Zanich, Samantha Febrey, Eithne Smith, Jenny Lennon, St George’s University of London

Database cleaning: Loes Rutten-Jacobs, University of Cambridge

Participating centres (number of enrolled patients per centre; local investigators):
Aberdeen Royal Infirmary, Aberdeen (12; Mary Macleod). Addenbrooke’s Hospital, Cambridge (54; Jean-Claude Baron, Elizabeth Warburton, Diana J Day, Julie White). Airedale General Hospital, Steeton (4; Samantha Mawer). Barnsley Hospital, Barnsley (3; Mohammad Albazzaz, Pravin Torane, Keith Elliott, Kay Hawley). Bart’s and the London, London (2; Patrick Gompertz). Basingstoke and North Hampshire Hospital, Basingstoke (13; Elio Giallombardo, Deborah Dellafera). Blackpool Victoria Hospital, Blackpool (11; Mark O'Donnell). Bradford Royal Infirmary, Bradford (1; Chris Patterson). Bristol Royal Infirmary, Bristol (8; Sarah Caine). Charing Cross Hospital, London (12; Pankaj Sharma). Cheltenham General and Gloucester Royal Hospitals, Cheltenham and Gloucester (10; Dipankar Dutta). Chesterfield Royal Hospital, Chesterfield (4; Sunil Punnoose, Mahmod Sajid). Countess of Chester Hospital, Chester (22; Kausik Chatterjee). Derriford Hospital, Plymouth (4; Azlisham Mohd Nor). Dorset County Hospital NHS Foundation Trust, Dorchester (6; Rob Williams). East Kent Hospitals University NHS Foundation Trust, Kent (22; Hardeep Baht, Guna Gunathilagan). Eastbourne District General Hospital, Eastbourne (4; Conrad Athulathmudali). Frenchay Hospital, Bristol (1; Neil Baldwin). Frimley Park Hospital NHS Foundation Trust, Frimley (6; Brian Clarke). Guy’s and St Thomas’ Hospital, London (14; Tony Rudd). Institute of Neurology, London (25; Martin Brown). James Paget University Hospital, Great Yarmouth (1; Peter Harrison). King’s College Hospital, London (16; Lalit Kalra). Leeds Teaching Hospitals NHS Trust, London (125; Ahamad Hassan). Leicester General Hospital and Royal Infirmary, Leicester (9; Tom Robinson, Amit Mistri). Luton and Dunstable NHSFT University Hospital, Luton (16; Lakshmanan Sekaran, Sakthivel Sethuraman, Frances Justin). Maidstone and Tunbridge Wells NHS Trust (3; Peter Maskell). Mayday University Hospital, Croydon (14; Enas Lawrence). Medway Maritime Hospital, Gillingham (5; Sam Sammuganathan). Milton Keynes Hospital, Milton Keynes (1; Yaw Duodu). Musgrove Park Hospital, Taunton (9; Malik Hussain). Newcastle Hospitals NHS Foundation Trust, Newcastle upon Tyne (12; Gary Ford). Ninewells Hospital, Dundee (5; Ronald MacWalter). North Devon District Hospital, Barnstaple (8; Mervyn Dent). Nottingham University Hospitals, Nottingham (17; Philip Bath, Fiona Hammond). Perth Royal Infirmary, Perth (2; Stuart Johnston). Peterborough City Hospital, Peterborough (1; Peter Owusu-Agyei). Queen Elizabeth Hospital, Gateshead (5; Tim Cassidy, Maria Bokhari). Radcliffe Infirmary, Oxford (5; Peter Rothwell). Rochdale Infirmary, Rochdale (4; Robert Namushi). Rotherham General Hospital, Rotherham (1; James Okwera). Royal Cornwall Hospitals NHS Trust, Truro (11; Frances Harrington, Gillian Courtauld). Royal Devon and Exeter Hospital, Exeter (22; Martin James). Royal Hallamshire Hospital, Sheffield (1; Graham Venables). Royal Liverpool University Hospital and Broadgreen Hospital, Liverpool (9; Aravind Manoj). Royal Preston Hospital, Preston (18; Shuja Punekar). Royal Surrey County Hospital, Guildford (23; Adrian Blight, Kath Pasco). Royal Sussex County Hospital, Brighton (14; Chakravarthi Rajkumar, Joanna Breeds). Royal United Hospital, Bath (6; Louise Shaw, Barbara Madigan). Salford Royal Hospital, Salford (16; Jane Molloy). Southampton General Hospital, Southampton (1; Giles Durward). Southend Hospital, Westcliff-on-Sea (26; Paul Guyler).
Southern General Hospital, Glasgow (34; Keith Muir, Wilma Smith). St George’s Hospital, London (108; Hugh Markus). St Helier Hospital, Carshalton (10; Val Jones). Stepping Hill Hospital, Stockport (4; Shivakumar Krishnamoorthy). Sunderland Royal Hospital, Sunderland (1; Nikhil Majumdar). The Royal Bournemouth Hospital, Bournemouth (15; Damian Jenkinson). The Walton Centre, Liverpool (15; Richard White). Torbay Hospital, Torquay (19; Debs Kelly). University Hospital Aintree, Liverpool (19; Ramesh Durairaj). University Hospital of North Staffordshire, Stoke-on-trent (16; David Wilcock). Wansbeck General Hospital and North Tyneside Hospital, Ashington and North Shields (6; Christopher Price). West Cumberland Hospital, Whitehaven (6; Olu Orugun, Rachel Glover). West Hertfordshire Hospital, Watford (20; David Collas). Western General Hospital, Edinburgh (12; Cathie Sudlow). Western Infirmary, Glasgow (33; Kennedy R. Lees, Jesse Dawson). Wycombe Hospital and Stoke Mandeville, High Wycombe (20; Dennis Briley and Matthew Burn). Yeovil District Hospital, Yeovil (46; Khalid Rashed). York Teaching Hospital, York (1; John Coyle).
Genetic Architecture of Lacunar Stroke

Matthew Traylor, PhD; Steve Bevan, PhD; Jean-Claude Baron, ScD, et al.

1 Department of Clinical Neurosciences, University of Cambridge, Cambridge, United Kingdom; and 2 Department of Neurology, Université Paris Descartes Sorbonne Paris Cité, Centre de Psychiatrie et Neurosciences, Institut National de la Santé et de la Recherche Médicale S894, Centre Hospitalier Sainte-Anne, Paris, France.

Abstract

Background and Objectives: Lacunar stroke is a common cause of stroke that accounts for approximately 20% of strokes. The aetiologies of lacunar strokes are numerous and complex, highlighting the need for further research to elucidate the genetic risk factors. We aimed to utilise the genetic variants available in the Genomes-Wide Association Studies (GWAS) database to determine whether candidate traits were associated with the risk of lacunar stroke.

Methods: We performed a genome-wide association analysis of single nucleotide polymorphisms (SNPs) in the GWAS database that were associated with risk of lacunar stroke. We then used these associations to estimate the effect of each SNP on the risk of lacunar stroke.

Results: We identified 123 significant associations between SNPs and the risk of lacunar stroke. Of these, 105 were previously reported in the GWAS database and 18 were novel. The results of our analysis suggest that genetic factors play a role in the development of lacunar stroke.

Conclusion: Our findings support the hypothesis that genetic factors contribute to the risk of lacunar stroke. Further research is needed to identify the specific genetic variants that are associated with the risk of lacunar stroke.