ARTICLE OPEN ACCESS

How common are single gene mutations as a cause for lacunar stroke?

A targeted gene panel study

Rhea Y.Y. Tan, PhD, Matthew Traylor, PhD, Karyn Megy, PhD, Daniel Duarte, MSc, Sri V.V. Deevi, PhD, Olga Shamardina, PhD, and Rutendo P. Mapeta, BSc, NIHR BioResource: Rare Diseases Consortium Willem H. Ouwehand, FMedSci, Stefan Gräf, PhD, Kate Downes, PhD, and Hugh S. Markus, FMedSci

Correspondence
Dr. Rhea Y.Y. Tan
yyrt2@medschl.cam.ac.uk

Neurology® 2019;93:e2007-e2020. doi:10.1212/WNL.0000000000008544

Abstract

Objectives

To determine the frequency of rare and pertinent disease-causing variants in small vessel disease (SVD)-associated genes (such as *NOTCH3*, *HTRA1*, *COL4A1*, *COL4A2*, *FOXC1*, *TREX1*, and *GLA*) in cerebral SVD, we performed targeted gene sequencing in 950 patients with younger-onset apparently sporadic SVD stroke using a targeted sequencing panel.

Methods

We designed a high-throughput sequencing panel to identify variants in 15 genes (7 known SVD genes, 8 SVD-related disorder genes). The panel was used to screen a population of 950 patients with younger-onset (≤70 years) MRI-confirmed SVD stroke, recruited from stroke centers across the United Kingdom. Variants were filtered according to their frequency in control databases, predicted effect, presence in curated variant lists, and combined annotation dependent depletion scores. Whole genome sequencing and genotyping were performed on a subset of patients to provide a direct comparison of techniques. The frequency of known disease-causing and pertinent variants of uncertain significance was calculated.

Results

We identified previously reported variants in 14 patients (8 cysteine-changing *NOTCH3* variants in 11 patients, 2 *HTRA1* variants in 2 patients, and 1 missense *COL4A1* variant in 1 patient). In addition, we identified 29 variants of uncertain significance in 32 patients.

Conclusion

Rare monogenic variants account for about 1.5% of younger onset lacunar stroke. Most are cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy variants, but the second most common gene affected is *HTRA1*. A high-throughput sequencing technology platform is an efficient, reliable method to screen for such mutations.

From the Stroke Research Group, Department of Clinical Neurosciences (R.Y.Y.T., M.T., H.S.M.), Department of Haematology (K.M., D.D., S.V.V.D., O.S., R.P.M., W.H.O., S.G., K.D.), and Division of Respiratory Medicine, Department of Medicine (S.G.), University of Cambridge; and NIHR BioResource: Rare Diseases (K.M., S.V.V.D., O.S., R.P.M., W.H.O., S.G., H.S.M.), Biomedical Campus, Cambridge, UK.

Go to Neurology.org/N for full disclosures. Funding information and disclosures deemed relevant by the authors, if any, are provided at the end of the article.

Coinvestigators are listed at links.lww.com/WNL/A1000.

The Article Processing Charge was funded by University of Cambridge.

Glossary

CADASIL = cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy; **CARASIL** = cerebral autosomal recessive arteriopathy with subcortical infarcts and leukoencephalopathy; **CI** = confidence interval; **CNV** = copy number variant; **DHPLC** = denaturing high-performance liquid chromatography; **FLAIR** = fluid-attenuated inversion recovery; **gnomAD** = genome aggregation database; **HTS** = high-throughput sequencing; **ICH** = intracerebral hemorrhage; **NIHR** = National Institute for Health Research; **SVD** = small vessel disease; **WES** = whole-exome sequencing; **WGS** = whole-genome sequencing; **WMH** = white matter hyperintensities.

Cerebral small vessel disease (SVD) accounts for around 25% of strokes in the form of lacunar strokes and deep intracerebral hemorrhages (ICH),[1] and is the primary pathology underlying vascular cognitive impairment.[2] In the majority of cases, it is a sporadic disease of aging related to hypertension and subsequent arteriosclerosis, but a minority of cases are due to rare genetic variants.[3] The most common inherited form of SVD is cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy (CADASIL) due to *NOTCH3* variants.[4] More recently, other genes have been reported to cause similar phenotypes, including *HTRA1*, *COL4A1*, *COL4A2*, *TREX1*, *GLA*, and *FOXC1*.[5] However, the frequency of these variants in populations with presumed sporadic SVD is unknown.

Identification of disease-causing variants currently largely relies on Sanger sequencing of the gene of interest. Often for cost reasons this involves sequencing a subset of exons, such as in CADASIL, where exons 3 and 4 are most frequently affected, and therefore preferentially screened.[6] As the spectrum of monogenic SVD expands, testing on a gene-by-gene basis is not cost- or time-effective. High-throughput sequencing (HTS) panels using next-generation sequencing technologies allow simultaneous testing in multiple genes underlying a single disease phenotype in a more cost-effective manner and are increasingly being used in clinical practice.

In this study, we developed a HTS panel comprising 15 genes linked to the SVD phenotype. We evaluated the platform for disease diagnosis and to determine the frequency of monogenic disease-causing variants in a well-defined population with MRI-confirmed younger-onset lacunar stroke. This study evaluates both known disease-causing mutations and novel, potentially disease-causing variants.

Methods

Platform design

The gene panel was developed to include 7 genes known to be causal of SVD (*NOTCH3*, *HTRA1*, *FOXC1*, *COL4A1*, *COL4A2*, *TREX1*, *GLA*) as well as 8 genes associated with disorders with SVD-related phenotypes. These include familial cerebral amyloid angiopathy (*APP*, *CST3*, *ITM2B*), familial hemiplegic migraine (*ATP1A2*, *CACNA1A*, *SCN1A*), and connective tissue disorders (*ABCC6*, *COL3A1*). These disorders share clinical manifestations with monogenic forms of SVD (for example, lacunar stroke, MRI white matter hyperintensities [WMH], dementia, migraine with aura, and encephalopathy) and could therefore present similarly. For each gene, the transcript on which to report variants was selected based on size, RefSeq information, and previously reported variants, and submitted to the Locus Reference Genomic database.[7] The capture design has previously been described by Simeoni et al.[8]

Study population

The study population consisted of patients from the UK DNA Lacunar Stroke Study.[9] A total of 72 specialist centers across the United Kingdom recruited unrelated patients of European ancestry with MRI-confirmed lacunar stroke occurring at or before the age of 70. The study was approved by the Multicentre Research Ethics Committee for Scotland (04/MRE00/36) and informed consent was obtained from participants. Stored DNA was available for 950 patients, all of whom were included in this study.

Lacunar stroke was defined as a clinical lacunar syndrome, with an anatomically compatible lesion on MRI (subcortical infarct ≤15 mm in diameter). All patients underwent full stroke investigations including brain MRI, carotid artery imaging, and ECG. Echocardiography was performed when appropriate. Patients were excluded if the cause of stroke was not SVD, including stenosis >50% in the extracranial or intracranial vessels; previous carotid endarterectomy; cardioembolic source of stroke defined according to Trial of Org 10172 in Acute Stroke Treatment criteria[10] 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 (e.g., lupus anticoagulant, vasculitis, dissection, known monogenic cause).

All MRI scans and clinical histories were reviewed centrally by one physician (H.S.M.). The presence and extent of WMH was graded on T2-weighted or fluid-attenuated inversion recovery (FLAIR) scans using the Fazekas scale: 0 = none, 1 = mild, 2 = early confluent, 3 = severe confluent, as previously described.[11] Lacunar infarcts were identified as a high signal lesion on acute diffusion-weighted imaging performed within 3 weeks of acute stroke, or as a cavitated hypodense lesion on T1 or FLAIR sequences. Cerebral microbleeds were identified

on gradient echo sequences. Family history was collected for first-degree relatives (table 1).

Sample processing

DNA samples were processed as previously described.[8]

Clinical bioinformatics

Sequence reads were processed as previously described.[8] The filtering step was adapted to SVD: single nucleotide variants and indels were prioritized in the following: (1) minor allele frequencies in the genome aggregation database (gnomAD)[12] <0.0001; (2) predicted impact according to SnpEff[15] is high, moderate, or splice region; (3) presence in HGMD Pro (2017.2) or in curated locus-specific databases (in particular the Leiden Open-Source Variation Database [May 2017 version][13]); (4) degree of deleteriousness according to Combined Annotation Dependent Depletion score[14] ≥15. *NOTCH3* variants that resulted in the gain or loss of a cysteine residue in the EGF-like repeats were automatically prioritized, as they are known to cause CADASIL.[15] The resulting variants were assessed according to the American College of Medical Genetics and Genomics guidelines[16] and retained if classified as pathogenic, likely pathogenic, or of unknown significance.

Large copy number variants (CNVs) were called using a custom pipeline based on ExomeDepth 1.1.10[19] as previously described.[17]

Table 1 Study population

	Values
Age at first stroke, y	
Mean (SD)	56.3 (8.7)
Median	57.4
Male sex, n (%)	674 (70.9)
BMI, mean (SD)	28.7 (6.2)
Number (%) of patients with the following:	
Hypertension	679 (71.5)
Diabetes	157 (16.5)
Hyperlipidemia	643 (67.7)
Smoker (current or previous)	663 (69.8)
Alcohol excess (≥20 units/wk)	272 (28.6)
Migraine	189 (19.8)
Migraine with aura	106 (11.1)
Myocardial infarction/coronary artery bypass graft or angioplasty	31 (3.3)
Peripheral vascular disease	26 (2.7)
Previous or recurrent strokes	70 (7.4)
White matter hyperintensities (Fazekas grade ≥2)	309 (32.5)
Microbleeds present[a]	58 (18.1)

Abbreviation: BMI = body mass index.
[a]A total of 321 patients had gradient echo sequences performed to evaluate the presence of cerebral microbleeds.

Panel validation

The panel was assessed by comparing the results with those obtained by 2 independent methods: whole-genome sequencing (WGS) and *NOTCH3* and *GLA* sequencing.

Thirty-four samples sequenced using the panel were also sequenced using WGS, as part of the National Institute for Health Research (NIHR) BioResource–Rare Disease study. The NIHR BioResource projects were approved by Research Ethics Committees in the United Kingdom and appropriate national ethics authorities in non–United Kingdom enrollment centers. WGS was performed by Illumina (San Diego, CA) on HiSeqXTen generating 150 bp paired-end reads per lane with minimum coverage of 15X for at least 95% of the genome (30X on average). Reads were aligned to the GRCh37 build of the human genome reference using the Isaac Aligner, and variants were called using the Isaac VariantCaller.[18] Variants in the 15 genes were analyzed following the same criteria as for the HTS panel.

Samples from all 950 patients had been previously screened for disease-causing *NOTCH3* and *GLA* variants.[9] Exons in 3, 4, 5, 6, 11, 18, 19, and 22 of *NOTCH3* were screened using denaturing high-performance liquid chromatography (DHPLC), and in addition, exons 3 and 4 were screened using Sanger sequencing.[9] Five patients with typical CADASIL-causing variants were identified. *GLA* was screened using high-resolution melt-curve analysis, covering all exons and intron/exon junctions, and one deep intronic region containing a known pathogenic variant. No Fabry-causing variants were identified.[9]

Statistical analysis

Comparisons of variant frequency between patients with and without a family history of stroke and with and without confluent WMH were performed using the Fisher exact test. Analyses were performed using R statistical software (version 3.5.1).

Data availability

Anonymized data will be made available upon reasonable request.

Results

Overall frequency of potentially disease-causing variants

Previously reported disease-causing variants and novel rare variants are shown in table 2.

In the 7 known SVD genes, known disease-causing variants were identified in 14 individuals (1.5%); this represented 11 different mutations. The proportion of patients with a reported family history of stroke found to have mutations was higher (8 of 372 patients [2.2%]) than that in patients without a reported family history of stroke (6 of 578 patients [1.0%]), although this difference was not significant (p = 0.18) (figure 1A).

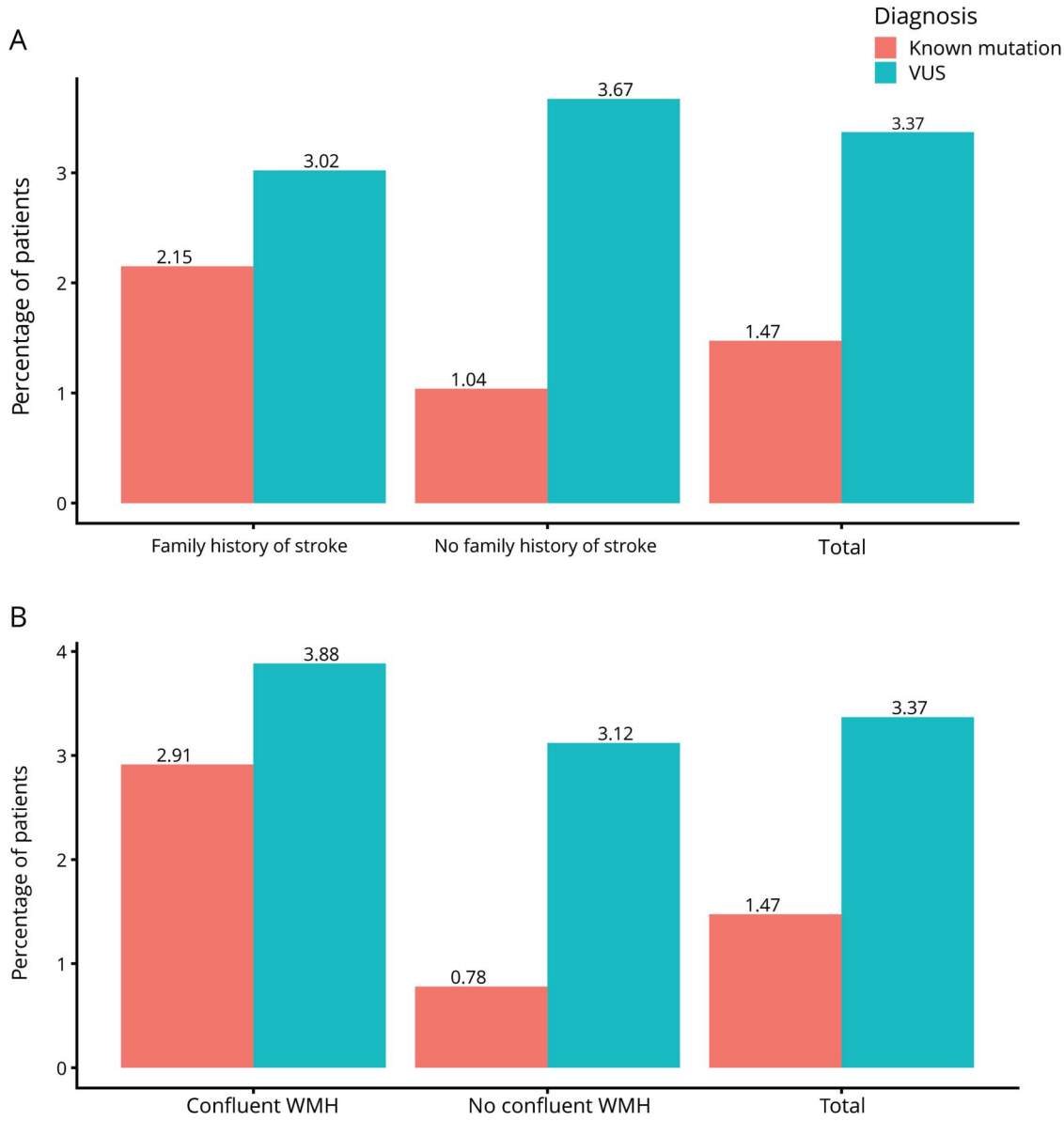
In addition, we identified 31 novel rare variants in the 7 known SVD genes. These were identified in 35 individuals. Excluding 2 *COL4A1* variants in 3 patients that were predicted to be benign in ClinVar (table 3), the overall frequency

Table 2 Known pathogenic *NOTCH3* (ENST00000263388), *HTRA1* (ENST00000368984), and *COL4A1* (ENST00000375820) variants

Gene	Sex	Age, y	Risk factors	Imaging features	FHx	Variant	MAF in gnomAD	References/remarks
NOTCH3	M	61	HTN, HLD, DM, S, BMI 34	WMH FS 3, bilateral ATP involvement, no GE	N	c.227G>A, p.Cys76Tyr	—	41 p.Cys76Arg[42] and p.Cys76Trp[43,44] also reported Not identified on DHPLC/Sanger sequencing Also detected on WGS
	M	46	HLD, S, BMI 30, MA	WMH FS 3, no GE	Y	c.505C>T, p.Arg169Cys	—	6,45 Not identified on DHPLC/Sanger sequencing Also detected on WGS
	M	47	HLD, EtOH, BMI 30	WMH FS 3, minor left ATP involvement, no GE	Y			
	M	58	HTN, HLD, S	WMH FS 3, no GE	Y	c.619C>T, p.Arg207Cys	8.1×10^{-6}	42 Identified on DHPLC and Sanger sequencing
	M	35	S, EtOH	WMH FS 2, no microbleeds on GE	N	c.967T>A, p.Cys323Ser	—	Internal CADASIL database Identified on DHPLC
	F	67	HTN, HLD, EtOH, BMI 45, Dep, MA, TIA	WMH FS 3, bilateral ATP involvement	Y	c.1162T>C, p.Cys388Arg	—	46 Not sequenced on DHPLC/Sanger
	F	65	HTN, HLD, S	WMH, FS 3, no microbleeds on GE	N	c.1759C>T, p.Arg587Cys	3.6×10^{-5}	45,47 Identified on DHPLC
	M	53	BMI 35	ILI, FS 1, no microbleeds	Y	c.3356G>A, p.Cys1119Tyr	8.1×10^{-6}	Internal CADASIL database Not sequenced on DHPLC/Sanger
	M	50	HLD, S	MLI, FS 1, no GE	Y	c.3664T>G, p.Cys1222Gly,	1.1×10^{-4}	48 Not identified on DHPLC[20,21]
	F	67	HTN, HLD, S, BMI 29	WMH FS 3, microbleeds	N			
	M	55	HTN, HLD, S, TIAs	WMH FS 2, no microbleeds	N			48 Identified on DHPLC
HTRA1	M	43	HTN, HLD, DM, S, BMI 30, Dep	MLI, FS 0, no microbleeds	Y	c.904C>T, p.Arg302Ter	1.1×10^{-5}	Previously reported in autosomal recessive and dominant disease
	F	46	HTN, S, BMI 29	ILI, FS 0, no microbleeds	N	c.1348G>C, p.Asp450His	7.2×10^{-6}	Variant in PDZ domain previously reported in autosomal dominant disease[21]
COL4A1	F	70	HTN, MA	MLI, FS 1, no GE	Y	c.1055C>T, p.Pro352Leu	6.5×10^{-5}	Previously associated with ICH,[24] impairs COL4A1 secretion[24]

Abbreviations: ATP = anterior temporal pole; BMI = body mass index; CADASIL = cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy; Dep = depression history; DHPLC = denaturing high-performance liquid chromatography; DM = diabetes mellitus; EtOH = alcohol excess; FHx = family history of stroke; FS = Fazekas score; GE = gradient echo imaging for microbleeds; gnomAD = genome aggregation database; HLD = hyperlipidemia; HTN = hypertension; ICH = intracerebral hemorrhage; ILI = isolated lacunar infarcts; MA = migraine with aura; MAF = minor allele frequencies; MLI = multiple lacunar infarcts; S = smoking history; WGS = whole-genome sequencing; WMH = white matter hyperintensities.
FHx was collected for first-degree relatives (parents, siblings, offspring).

Figure 1 Family history of stroke and severity of white matter disease



(A) Number (percentage of subjects tested) of known variants and variants of uncertain significance (VUS) identified in patients with and without a family history of stroke and (B) with and without confluent white matter hyperintensities (WMH) on MRI.

of novel variants was 3.4% (32 of 950 patients). There was no difference in the proportion of novel rare variants among patients with and without a family history of stroke (11 of 364 vs 21 of 572, respectively; $p = 0.71$) (figure 1A).

To determine whether variants were more common in a particular phenotype of SVD, or in more severe cases, we examined variant frequency in those with confluent WMH. Of the 309 patients with confluent WMH on MRI (Fazekas score[10] ≥2), 9 (2.9%) had a known disease-causing variant, compared with 5 of 641 (0.8%) in those without confluent WMH ($p = 0.018$). The proportions for rare novel variants of uncertain significance were 12 of 309 (3.9%) for those with WMH and 20 of 641 (3.1%) for those without ($p = 0.57$) (figure 1B).

CADASIL

CADASIL is caused by cysteine-altering *NOTCH3* mutations in the epidermal growth factor–like repeat domains encoded by exons 2 to 24.[15] Eight different cysteine-changing variants in exons 2–24 of *NOTCH3* were identified in 11 individuals (table 2). Previous screening had identified 5 disease-causing variants,[9] and these were again identified using this platform.

Of the 6 additional individuals with cysteine-changing *NOTCH3* variants, 2 variants in 2 individuals were previously missed by both DHPLC and Sanger sequencing.

The overall frequency of CADASIL-causing variants was 1.2% (95% confidence interval [CI] 0.6%–2.1%). Of patients with confluent WMH (Fazekas score ≥2) the frequency was 2.9% (9 of 309, 95% CI 1.5%–5.4%) compared to 0.3% of patients without confluent WMH (Fazekas score <2) (2 of 641, 95% CI 0.1%–1.1%) ($p = 0.001$).

Comparing age groups, the overall frequency of CADASIL-causing variants was 1.2% (95% CI 0.6%–2.5%) in patients ≤60 years and 1.1% (95% CI 0.4%–0.7%) in patients aged >60 years. Among patients with confluent WMH, the mutation frequency was 3.7% (95% CI 1.6%–8.3%) in patients ≤60 years and 2.3% (95% CI 0.9%–5.8%) in patients aged >60 years.

Non-CADASIL monogenic small vessel arteriopathies

HTRA1

Missense and nonsense *HTRA1* variants have been reported in cerebral autosomal recessive arteriopathy with subcortical infarcts and leukoencephalopathy (CARASIL).[19] Recently, heterozygous variants have also been identified in patients with an autosomal dominant form of SVD.[20,21] Eight heterozygous missense variants and 1 nonsense *HTRA1* variant were identified in 12 individuals (1.3%, 95% CI 0.7%–2.2%). Two of these have previously been reported as disease-causing: p.Arg302Ter was reported in both recessive and dominant disease[19,20]; p.Asp450His was reported in autosomal dominant disease.[21] p.Arg302Ter has been demonstrated to result in a mutant protein with 21%–50% of normal protease activity.[19] There were no individuals with compound heterozygous or homozygous *HTRA1* variants.

Of the remaining novel *HTRA1* variants, 5 resided in the trypsin domain, with 1 (p.Arg227Trp) identified in 4 individuals. This variant is close to the trypsin active site in position 220. The clinical features of patients with *HTRA1* variants are provided in table 3.

Among patients with confluent WMH (Fazekas score ≥2), the frequency of rare *HTRA1* variants passing filters was 1.3% (4 of 309, 95% CI 0.5%–3.3%), similar to that in those without confluent WMH (1.2%, 8 of 641, 95% CI 0.6%–2.4%, nonsignificant difference). In younger patients (≤60 years), the frequency was 1.2% (7 of 574, 95% CI 0.6%–2.5%), and this value was similar in those older than 60 (5 of 376, 1.3%, 95% CI 0.5%–3.1%).

COL4A1 and *COL4A2*

COL4A1 and *COL4A2* encode the α1 and α2 chains of collagen IV, respectively. Missense variants, typically but not always affecting the glycine residue in the repetitive Gly-X-Y regions, are associated with SVD and ischemic and hemorrhagic subcortical lacunar strokes.[22] Variants affecting the C4 domain associated with tropocollagen assembly have also been associated with SVD.[23] In 10 individuals (1.1%, 95% CI 0.6%–1.9%), we identified 9 missense *COL4A1* variants (figure 2B). There were no compound heterozygous individuals. One variant (p.Pro352Leu) was previously described in an individual with presumed sporadic ICH, and was also demonstrated to impair COL4A1 secretion into the extracellular space.[24] Two other variants have been reported as likely benign in ClinVar (p.Gly332Arg and p.Pro1337Leu).

We identified 9 heterozygous missense *COL4A2* variants in 9 individuals (0.9%, 95% CI 0.5%–1.8%). None of these have previously been reported (figure 2C).

Variants identified in both *COL4A1* and *COL4A2* were found in various regions including the Gly-X-Y regions and the C4 domain associated with disease (table 3).

FOXC1

The *FOXC1* gene encodes the Forkhead box C1 transcription factor. Autosomal dominant missense, nonsense, and frameshift variants, as well as deletion or duplication of the locus (6p25), have been associated with ocular abnormalities described as the Axenfeld-Rieger syndrome.[25,26] Some variants are also associated with white matter abnormalities.[25] Two heterozygous predicted high-impact variants in *FOXC1* were identified in 2 individuals (0.2%, 95% CI 0.06%–0.8%): 1 novel in-frame insertion (p.Thr68_Pro69insThrProGln) and 1 frameshift variant (p.Ala381_Gly382fs) (table 3). The in-frame insertion was in a low-complexity region near the forkhead box domain, and has previously been reported in ClinVar as likely pathogenic for Axenfeld-Rieger syndrome (p.Ala381GlyfsTer147). Loss-of-function variants in *FOXC1* are known to be pathogenic and the frameshift variant is predicted to lead to premature stop codon 147 residues downstream.

TREX1

Retinal vasculopathy with cerebral leukodystrophy and systemic manifestations arises due to frameshift variants near the C-terminus of *TREX1*.[27] Missense *TREX1* variants are also associated with Aicardi-Goutières syndrome,[28] a form of pediatric-onset encephalopathy, and familial chilblain lupus.[29] Three novel missense variants (p.Gly197Ala, p.Arg229Gly, p.Lys230Asn), 1 novel in-frame deletion, and 1 previously reported frameshift variant were found in 5 individuals (0.5%, 95% CI 0.2%–1.2%). The frameshift variant (p.Ala194fs) was also identified in the same patient in a separate study using Sanger sequencing (reported as p.Ala139Valfs*21 on a different isoform), and was found to nearly completely abolish TREX1 nuclease activity in vitro.[30] This variant has been reported in a compound heterozygous case of Aicardi-Goutières syndrome in the DECIPHER database (patient 303873).[31] (table 3).

Fabry disease: *GLA* variants

No pathogenic or likely pathogenic Fabry variants were identified.

Table 3 Rare, novel, or presumed benign variants identified in *HTRA1* (*ENST00000368984*), *COL4A1* (*ENST00000375820*), *COL4A2* (*ENST00000360467*), *FOXC1* (*ENST00000380874*), and *TREX1* (*ENST00000422277*) genes

Gene	Sex	Age, y	Risk factors	Imaging features	Family history of stroke	Gene/variant	MAF in gnomAD	Pathogenicity
HTRA1	F	68	HLD, S	ILI, FS 1, no microbleeds	N	c.521A>C, p.Asp174Ala	4.1×10^{-6}	Novel variant[a]
	M	56	HLD, S, EtOH, Dep	MLI, FS 1, no microbleeds	N	c.632A>C, p.Glu211Ala	—	Novel variant[a]
	M	59	HTN, HLD, DM	WMH FS 3, no GE	N	c.679C>T, p.Arg227Trp	7.9×10^{-5}	Novel variant[b]
	M	62	HTN, HLD, S, BMI 30, Mig	ILI, FS 0, no microbleeds	Y			
	M	60	HTN, HLD, EtOH, BMI 40	WMH, FS 3, no GE	N			
	M	61	HTN, HLD, S, EtOH	ILI, FS 0, no microbleeds	N			
	M	50	HTN, S	WMH FS 3, no GE	Y	c.834C>A, p.Phe278Leu	—	Novel variant[b]
	M	63	HTN, HLD, MA	ILI, FS 1, microbleeds	N	c.940A>G, p.Met314Val	8.1×10^{-6}	Novel variant[b]
	M	41	None	ILI, FS 0, no GE	N	c.958G>A, p.Asp320Asn	1.2×10^{-5}	Novel variant[b]
	F	62	HLD, BMI 36	WMH, FS 2, no GE	Y	c.1103A>G, p.His368Arg	—	Novel variant found between trypsin and PDZ domains
COL4A1	F	53	HTN, HLD, S, EtOH, D	WMH FS 2, no GE	Y	c.994G>C, p.Gly332Arg	4.3×10^{-5}	ClinVar: likely benign variant; ncbi.nlm.nih.gov/clinvar/variation/311067/
	F	45	BMI 37	ILI, FS 0, no GE	N	c.1246C>G, p.Pro416Ala	4.9×10^{-5}	Novel variant[d]
	F	63	HTN, HLD, S, BMI 32, Dep, MA	WMH, FS 2, no GE	N	c.2093A>G, p.Lys698Arg	5.9×10^{-5}	Novel variant[a]
	F	55	HTN, HLD, S, BMI 32	ILI, FS 0, no GE	NK	c.2174C>T, p.Pro725Leu	—	Novel variant[c]
	M	42	EtOH, BMI 33	ILI, FS 0, no GE	N	c.4010C>T, p.Pro1337Leu	7.9×10^{-5}	ClinVar: likely benign variant in context of *COL4A1*-related SVD[b]; ncbi.nlm.nih.gov/clinvar/variation/311030/
	F	68	HLD	ILI, FS 0, no microbleeds	N			
	M	38	HTN, S, BMI 30	ILI, FS 0, no GE	NK	c.4423T>C, p.Tyr1475His	1.6×10^{-5}	Novel variant[e]
	M	65	S	WMH, FS 3, no GE	Y	c.4678G>A, p.Val1560Met	1.2×10^{-5}	Novel variant[e]
	M	69	S, EtOH	WMH, FS 2, no microbleeds	N	c.4970C>T, p.Thr1657Met	2.4×10^{-5}	Novel variant[e]
COL4A2	M	59	HTN, HLD, S, EtOH, BMI 37	WMH, FS 2, no GE	Y	c.661C>A, p.Pro221Thr	1.5×10^{-5}	Novel variant[c]
	M	58	HLD	WMH, FS 3, no microbleeds	N	c.965G>A, p.Arg322Gln	3.6×10^{-5}	Novel variant[c]
	M	65	HTN, HLD, S	ILI, FS 0, no GE	NK	c.1396G>A, p.Gly466Ser	5.4×10^{-5}	Novel variant[c]

Continued

Table 3 Rare, novel, or presumed benign variants identified in *HTRA1* (*ENST00000368984*), *COL4A1* (*ENST00000375820*), *COL4A2* (*ENST00000360467*), *FOXC1* (*ENST00000380874*), and *TREX1* (*ENST00000422277*) genes *(continued)*

Gene	Sex	Age, y	Risk factors	Imaging features	Family history of stroke	Gene/variant	MAF in gnomAD	Pathogenicity
	M	56	HTN, HLD, BMI 29	ILI, FS 1, no GE	NK	c.1687G>A, p.Gly563Ser	—	Novel variant[d]
	M	53	HLD, S, BMI 40	ILI, FS 0, no GE	Y	c.2581C>T, p.Arg861Cys	2.7×10^{-5}	Novel variant[c]
	M	65	EtOH	WMH, FS 2, no GE	NK	c.3715G>A, p.Gly1239Ser	—	Novel variant[d]
	M	61	HTN, HLD, S, BMI 30	ILI, FS 1, no GE	Y	c.4292G>A, p.Arg1431His	2.9×10^{-5}	Novel variant[d]
	M	62	HTN, S	MLI, FS 0, no GE	N	c.4534T>C, p.Trp1512Arg	6.2×10^{-5}	Novel variant[e]
	M	58	HTN, HLD, DM, BMI 34, Dep	MLI, FS 0, no GE	Y	c.4601C>T, p.Ala1534Val	1.2×10^{-5}	Novel variant[e]
FOXC1	M	58	HLD, S, BMI 32	ILI, FS 0, no microbleeds	Y	c.204_205insGCCGCA, p.Thr68_Pro69insThr, ProGln	—	Novel in-frame insertion
	M	47	HLD, S	ILI, FS 0, no GE	Y	c.1141_1142insG, p.Ala381_Gly382fs	—	ClinVar: likely pathogenic variant for Axenfeld-Rieger syndrome; ncbi.nlm.nih.gov/clinvar/variation/537387/; frameshift variants associated with SVD.[25]
TREX1	F	48	HTN, HLD, DM, BMI 31, Dep	WMH FS 3, multiple microbleeds	N	c.581delC, p.Ala194fs	2.8×10^{-5}	Previously reported in Aicardi-Goutières syndrome[31]; also identified in the same patient on Sanger sequencing and found to impair TREX1 function[30]
	F	43	HTN, S, BMI 30, Dep	ILI, FS 0, no GE	Y	c.590G>C, p.Gly197Ala	—	Novel variant
	F	51	HTN, HLD, DM, BMI 39, Dep	WMH, FS 2, no GE	N	c.685A>G, p.Arg229Gly	1.1×10^{-5}	Novel variant
	M	65	HTN, BMI 28	ILI, FS 0, no microbleeds	N	c.690G>T, p.Lys230Asn	—	Novel variant
	F	60	HTN, HLD, BMI 32	ILI, FS 0, no GE	N	c.1024_1041del, CTGCTGGCCCCA, CTGGGT p.Leu342_Gly347del	6.1×10^{-5}	Novel in-frame deletion

Abbreviations: BMI = body mass index; Dep = depression history; DM = diabetes mellitus; EtOH = alcohol excess; FS = Fazekas score; GE = gradient echo imaging for microbleeds; gnomAD = genome aggregation database; HLD = hyperlipidemia; HTN = hypertension; ILI = isolated lacunar infarcts; MA = migraine with aura; MAF = minor allele frequencies; Mig = migraine without aura; MLI = multiple lacunar infarcts; NK = not known; S = smoking history; SVD = small vessel disease; WMH = white matter hyperintensities.
Family history was collected for first-degree relatives (parents, siblings, offspring).
[a] Found between trypsin and kazal domains of *HTRA1* gene, of uncertain significance.
[b] Found in trypsin domain of *HTRA1* gene, associated with enzyme function.
[c] In triple helix repeat region of *COL4A1* or *COL4A2* genes—other Gly-altering variants associated with disease.
[d] In low complexity region of *COL4A1* or *COL4A2* genes, of uncertain significance.
[e] In C4 domain of *COL4A1* or *COL4A2* genes, associated with tropocollagen assembly,[49] variants previously associated with disease.[23]

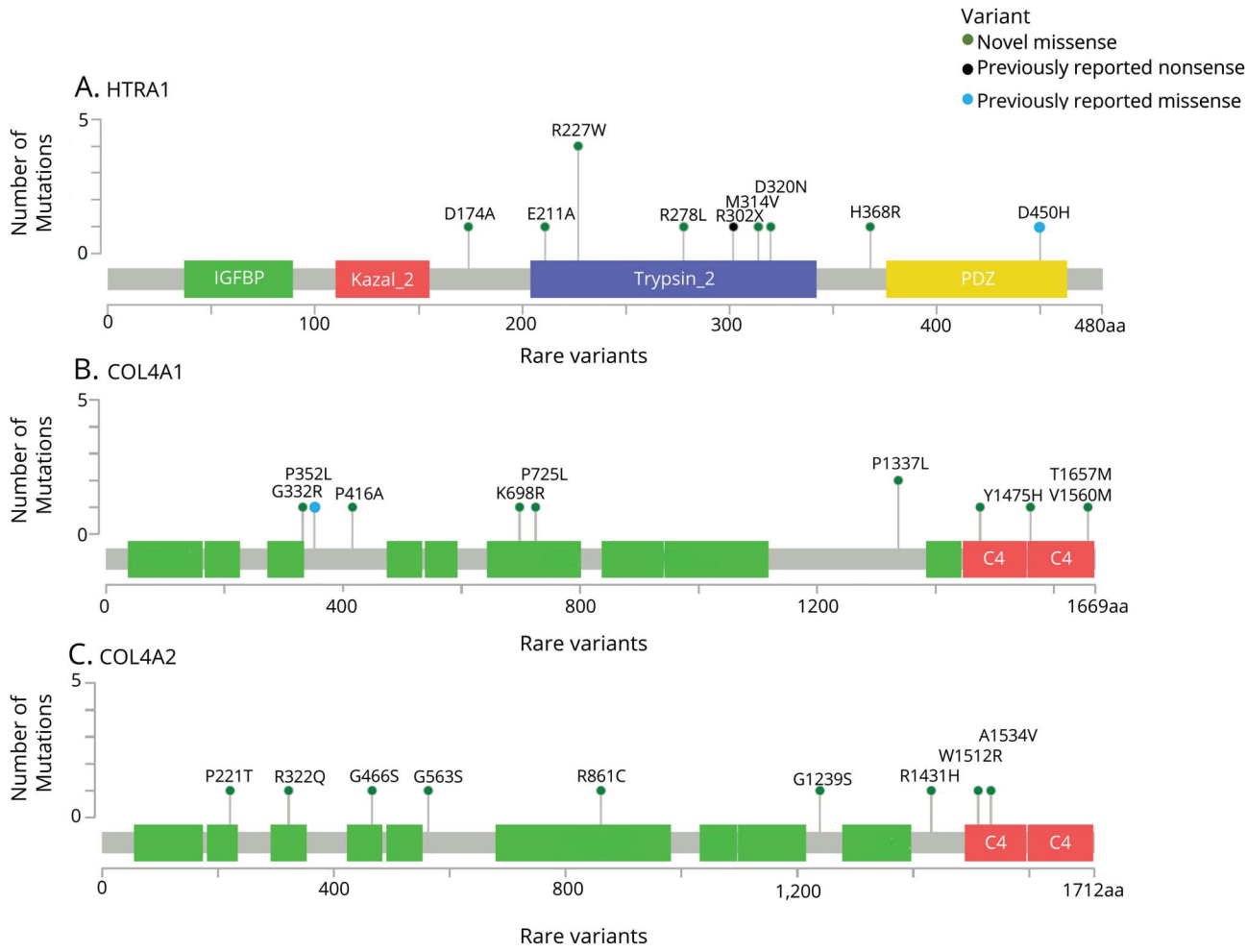
Copy number variants

Excluding CNVs with a Bayes factor lower than 20, only 1 CNV was detected in the 7 known cerebral SVD genes. This was a large duplication of both *COL4A1* and *COL4A2* in a single individual.

SVD-related genes

Forty-five heterozygous variants in 8 genes associated with SVD-related disorders were identified in 47 individuals (table 4). Of note, 2 novel variants were identified in cystatin 3 (*CST3*), where variants are known to cause familial

Figure 2 Rare variants identified in *HTRA1*, *COL4A1*, and *COL4A2*



Rare variants identified in (A) *HTRA1* (ENST00000368984), (B) *COL4A1* (ENST00000375820), and (C) *COL4A2* (ENST00000360467).

cerebral amyloid angiopathy, Icelandic type[32]: 1 stop-gained and 1 frameshift variant, both at position 93 (p.Glu93* and p.Glu93fs). Lacunar stroke is not an established manifestation of the disease, indicating that more investigations are needed to assign pathogenicity to these variants.

Validation of panel

In the 34 individuals sequenced by WGS, 2 were found to harbor CADASIL-causing *NOTCH3* variants. These were also detected by the HTS platform (table 2). No other exonic variants passing the filters could be identified in the remaining 32 individuals, either by WGS or by targeted sequencing. In addition, 5 cysteine-changing *NOTCH3* variants had been detected on prior screening, all of which were also identified by the platform.

Discussion

We developed an HTS panel for 7 genes known to be implicated in monogenic SVD and 8 genes associated with related phenotypes that can enter in the differential diagnosis of monogenic SVD. Validation against other techniques found the panel to be as sensitive, if not more sensitive than conventional screening methods, and detected all previously identified *NOTCH3* variants, including some missed on previous screening.

NOTCH3 variants, which underlie CADASIL, the most common cause of monogenic SVD, were identified in about 1% of patients with apparently sporadic SVD aged 70 or under. The yield rose to 3.7% if analysis was limited to younger (age ≤60) patients who also had confluent WMH.

Table 4 Rare variants identified in genes associated with phenotypes related to small vessel disease: *COL3A1* (ENST00000304636), *APP* (ENST00000346798), *CST3* (ENST00000376925), *ITM2B* (ENST00000378565), *ATP1A2* (ENST00000361216), and *CACNA1A* (ENST00000361216), *SCN1A* (ENST00000375405)

Related phenotype	Gene	Variant	dbSNP ID	No. of individuals	MAF in gnomAD	Remarks	Reference
Ehlers-Danlos (IV)	*COL3A1*	c.74A>G, p.Gln25Arg	—	1	—	In signal peptide	
		c.1856C>T, p.Pro619Leu	rs373838193	1	4.3×10^{-5}	In low complexity region VUS on ClinVar for aortic aneurysm/dissection and Ehlers-Danlos type IV	ncbi.nlm.nih.gov/clinvar/variation/199718/
		c.1864C>T, p.Pro622Ser	rs772638774	2	3.3×10^{-5}	In low complexity region	
		c.226A>G, p.Asn76Asp	rs142045411	1	6.2×10^{-5}	In von Willebrand factor type C domain	
		c.1165A>T, p.Asn389Tyr	rs200394946	1	9.7×10^{-5}	In collagen triple helix repeat domain VUS on ClinVar for aortic aneurysm/dissection and Ehlers-Danlos Type IV	ncbi.nlm.nih.gov/clinvar/variation/199718/
Cerebral amyloid angiopathy	*APP*	c.373G>A, p.Asp125Asn	—	1	—	In amyloid A4 N-terminal heparin binding domain	
		c.682G>A, p.Val228Ile	rs755841034	2	4.1×10^{-6}	In low complexity region, near p.Val225Ala, a VUS associated with familial Alzheimer disease	
		c.727G>A, p.Asp243Asn	rs750279232	1	3.3×10^{-6}	In coiled coil domain	
		c.736G>A, p.Glu246Lys	rs147485129	1	1.8×10^{-5}	In coiled coil domain	
		c.1859C>T, p.Pro620Leu	rs759926843	1	1.6×10^{-5}	In low complexity region	
		c.1909+5G>A, Intron and splice region	rs200271509	1	1.5×10^{-5}	In coiled coil domain	
		c.2148C>T, p.Ile716Ile	rs145564988	1	6.9×10^{-5}	In C-terminus domain; Ile716Thr reported on ClinVar (no clinical interpretation available); Ile716Val associated with familial Alzheimer disease	ncbi.nlm.nih.gov/clinvar/variation/98240/ 50
	CST3	c.277G>T, p.Glu93Ter	—	1	4.1×10^{-6}	p.Leu94Gln reported in familial Icelandic dementia	51
		c.277_296delGAGCT.GGGCCGAACCACGTG.p.Glu93fs	rs765138253	1	6.1×10^{-5}		
	ITM2B	c.92C>G, p.Pro31Arg	rs150336652	1	8.9×10^{-5}	Low complexity region	
FHM	*ATP1A2*	c.1262G>A, p.Arg421Gln	rs139499540	1	5.5×10^{-5}	Between E1-E2 and Cation ATPase domains Reported as VUS for FHM on ClinVar	ncbi.nlm.nih.gov/clinvar/variation/406194/
		c.2146G>T, p.Val716Leu	—	1	—	Near metal ion binding sites at residues 714 and 718 p.Asp718Asn reported as pathogenic in FHM	52
	CACNA1A	c.130G>A, p.Ala44Thr	rs201398669	1	2.5×10^{-5}	Low complexity region	

Continued

Table 4 Rare variants identified in genes associated with phenotypes related to small vessel disease: *COL3A1* (ENST00000304636), *APP* (ENST00000346798), *CST3* (ENST00000376925), *ITM2B* (ENST00000378565), *ATP1A2* (ENST00000361216), and *CACNA1A* (ENST00000361216), *SCN1A* (ENST00000375405) *(continued)*

Related phenotype	Gene	Variant	dbSNP ID	No. of individuals	MAF in gnomAD	Remarks	Reference
		c.633T>C, p.Ser211Ser	rs202216404	1	6.2×10^{-5}	Ion transport domain p.Ser211Asn reported as VUS for unspecified phenotype	ncbi.nlm.nih.gov/clinvar/variation/426949/
		c.1083-108_1083-101delCACACACA	—	1	—	Intron variant	
		c.2524G>C, p.Glu842Gln	—	1	—	Between ion transport domains; p.Ala841Ser reported as VUS for unspecified phenotype	ncbi.nlm.nih.gov/clinvar/variation/385371/
		c.2812G>C, p.Gly938Arg	rs771423362	1	4.1×10^{-5}	Between ion transport domains; reported as VUS for unspecified phenotype	ncbi.nlm.nih.gov/clinvar/variation/422044/
		c.2944G>A, p.Gly982Ser	—	1	—	Low complexity region	
		c.2960G>A, p.Arg987Gln	—	1	—	Low complexity region; p.Arg987Pro reported as VUS for unspecified phenotype	ncbi.nlm.nih.gov/clinvar/variation/387025/
		c.3226G>A, p.Ala1076Thr	rs554091859	1	8.6×10^{-5}	Between ion transport domains; reported as VUS (unspecified phenotype)	ncbi.nlm.nih.gov/clinvar/variation/429962/
		c.3265G>A, p.Gly1089Ser	rs201311000	1	3.3×10^{-5}	Between ion transport domains; Gly1089Cys reported as VUS for unspecified phenotype	ncbi.nlm.nih.gov/clinvar/variation/391877/
		c.3436G>A, p.Val1146Ile	rs376365775	1	1.6×10^{-5}	Between ion transport domains; reported as VUS for unspecified phenotype	ncbi.nlm.nih.gov/clinvar/variation/195470/
		c.4646A>G, p.Gln1549Arg	rs759263620	1	4.1×10^{-5}	Between ion transport domains	
		c.5494G>, p.Val1832Met	rs376815942	1	1.2×10^{-5}	GPHH (voltage-dependent L type calcium channel) domain	
		c.5795A>G, p.Asn1932Ser	—	1	4.1×10^{-6}	Voltage-gated calcium channel IQ domain	
		c.5936A>C, p.Glu1979Ala	rs774657158	1	8.3×10^{-6}	Outside of Ca channel IQ domain	
		c.6400C>T, p.Arg2134Cys	rs121908235	1	6.2×10^{-5}	Outside of Ca channel IQ domain VUS for unspecified phenotype and associated with episodic ataxia type 2 (impact uncertain)	ncbi.nlm.nih.gov/clinvar/variation/68440
		c.6644A>C, p.His2215Pro	—	1	3.8×10^{-5}	Low complexity region	
		c.6715C>T, p.Arg2239Trp	rs759576380	1	5.2×10^{-5}	Outside of Ca channel IQ domain	
		c.6758G>A, p.Arg2253Gln	rs752950486	1	1.6×10^{-5}	Outside of Ca channel IQ domain	
		c.6839G>T, p.Arg2280Leu	—	1	—	Patient also has a deletion in ABCC6 Outside of Ca channel IQ domain	
		c.7027G>A, p.Gly2343Ser	—	1	6.5×10^{-5}	Outside of Ca channel IQ domain	

Continued

Table 4 Rare variants identified in genes associated with phenotypes related to small vessel disease: *COL3A1* (ENST00000304636), *APP* (ENST00000346798), *CST3* (ENST00000376925), *ITM2B* (ENST00000378565), *ATP1A2* (ENST00000361216), and *CACNA1A* (ENST00000361216), *SCN1A* (ENST00000375405) *(continued)*

Related phenotype	Gene	Variant	dbSNP ID	No. of individuals	MAF in gnomAD	Remarks	Reference
		c.7166G>A, p.Arg2389Gln	—	1	—	Outside of Ca channel IQ domain	
	SCN1A	c.136G>A, p.Glu46Lys	rs769582667	1	3.2×10^{-5}	Low complexity region	
		c.1457C>G, p.Ala486Gly	rs777120925	1	5.8×10^{-5}	Low complexity region	
		c.1738C>G, p.Arg580Gly	—	1	—	Cytoplasmic domain of voltage-gated Na ion channel p.Arg580Gln reported as VUS for early infantile epileptic encephalopathy; p.Arg580Ter reported as pathogenic for severe myoclonic epilepsy in infancy	ncbi.nlm.nih.gov/clinvar/variation/167646/53
		c.1843G>A, p.Gly615Arg	—	1	4.1×10^{-6}	Cytoplasmic domain of voltage-gated Na ion channel	
		c.3394G>A, p.Glu1132Lys	—	1	—	Sodium ion transport-associated domain	
		c.4690C>T, p.Arg1564Cys	rs121918807	1	7.7×10^{-5}	Ion transport domain	
		c.5750G>A, p.Arg1917His	—	1	—	Outside of ion transport domain	

Abbreviations: FHM = familial hemiplegic migraine; gnomAD = genome aggregation database; MAF = minor allele frequency; VUS = variants of unknown significance.

A number of other monogenic forms of SVD have been reported, and although they appear to be rarer than CADASIL, their frequency in cases of lacunar stroke is unknown. Our study provides data on their prevalence in cases of symptomatic lacunar stroke.

We found no homozygous CARASIL-causing variants, but up to 1.3% of cases had potentially disease-causing heterozygous missense and nonsense variants, suggesting that *HTRA1*-associated autosomal dominant SVD is the second most common cause of familial SVD. These findings are consistent with previous smaller studies in which about 5% of patients with *NOTCH3*-negative familial SVD carry missense and nonsense *HTRA1* variants.[21,33,34]

In addition, we identified several rare variants predicted to be damaging in *COL4A1*, *COL4A2*, and *FOXC1*. One individual had a *COL4A1* variant (c.1055C>T, p.Pro352Leu), which has been described in an individual with sporadic ICH.[24] However, without functional support and evidence of segregation with disease, the pathogenicity of this variant remains uncertain.

We identified a frameshift *FOXC1* variant (c.1141_1142insG, p.Ala381_Gly382fs). Given that *FOXC1* has 1 exon, the transcript containing this predicted premature stop codon may escape nonsense-mediated decay. Functional studies are required to confirm pathogenicity of this high-impact variant.

Fabry disease has been suggested as an underreported cause of young stroke,[35] and has been associated with MRI features of SVD,[36] but we found no Fabry variants in this cohort. This is consistent with recent data suggesting that its importance as a cause of early-onset cryptogenic stroke may have been overestimated.[37,38]

Our results highlight a major challenge in the use of HTS panels in clinical practice; namely, determining whether a variant is pathogenic or benign. In CADASIL, variant interpretation is aided by the knowledge that all known disease-causing variants are cysteine-altering. Assigning pathogenicity of variants in genes such as *HTRA1*, however, is challenging. Variant interpretation guidelines highlight the use of allele frequency in large public databases. This presents particular challenges in late-onset diseases such as SVD. Databases such as gnomAD,[39] which are used to infer population-level frequencies of potentially disease-causing variants, include individuals drawn from different populations predominantly of European ancestry, with variable age distributions. For individuals who are young at the point of recruitment, it is unknown whether they might develop SVD in the future. In addition, such databases include affected populations from diseases such as dementia, which might include patients with

SVD due to misclassification. Determining whether variants segregate with disease in families is important but often family members are unavailable.

The use of targeted HTS panels presents several advantages. While WGS, whole-exome sequencing (WES), and HTS panels can all provide clinically acceptable coverage, the higher read depth of HTS panels allows for the detection of high-confidence CNVs. HTS panels also allow for savings in terms of cost and analysis time, compared to WGS and WES.

Our study has limitations. Our analyses did not include sequencing a cohort of MRI-phenotyped unaffected individuals. Previous studies have used age- and sex-matched controls unscreened for disease as comparison.[40] However, the value of such controls is questionable. Much of SVD is subclinical prior to cerebrovascular events and ruling out subclinical disease in such controls is not possible.

This study was performed in patients of European ancestry. It would not be possible to extrapolate these results to populations of other ancestries as variant frequencies may vary significantly in different populations. Similar studies should be performed in other populations, particularly as population databases expand to better represent these ethnicities.

Our results demonstrate that monogenic mutations account for about 1.5% of SVD stroke. Furthermore, they demonstrate the utility of targeted HTS platforms in the diagnosis of monogenic forms of SVD. We showed they could detect CADASIL-causing variants, suggesting they could replace commonly used Sanger sequencing techniques, with the added benefit of screening for other genes, such as *HTRA1*, shown in this study to be the second most common cause of monogenic SVD.

Acknowledgment

The authors thank the contributors to the DNA Lacunar Stroke Study, listed at journals.plos.org/plosone/article?id=10.1371/journal.pone.0136352. This study makes use of data generated by the NIHR BioResource–Rare Diseases. The authors thank the NIHR BioResource volunteers, the NIHR BioResource Centre, and its staff; the National Institute for Health Research; and NHS Blood and Transplant.

Study funding

Funding for the project was provided by the National Institute for Health Research (NIHR, grant RG65966). This work was supported by a BHF Programme Grant (RG/16/4/32218). Infrastructural support was provided by the Cambridge University Hospitals Trust Biomedical Research Centre. R.Y.Y.T. is supported by the Agency for Science, Technology and Research (Singapore). K.D. is a HSST trainee supported by NHS Health Education England. H.S.M. is supported by an NIHR Senior Investigator award.

Disclosure

The authors report no disclosures relevant to the manuscript. Go to Neurology.org/N for full disclosures.

Publication history

Received by *Neurology* January 10, 2019. Accepted in final form June 10, 2019.

Appendix Authors

Name	Location	Role	Contribution
Rhea Y.Y. Tan	University of Cambridge	Author	Study design, data acquisition, data analysis, drafting of manuscript
Matthew Traylor	University of Cambridge	Author	Data analysis, drafting of manuscript
Karyn Megy	University of Cambridge	Author	Data acquisition, data analysis, drafting of manuscript
Daniel Duarte	University of Cambridge	Author	Data acquisition
Sri V.V. Deevi	University of Cambridge	Author	Data acquisition, data analysis
Olga Shamardina	University of Cambridge	Author	Data acquisition, data analysis
Rutendo P. Mapeta	University of Cambridge	Author	Data acquisition
Willem H. Ouwehand	University of Cambridge	Author	Design and conceptualization of study
Stefan Gräf	University of Cambridge	Author	Data analysis
Kate Downes	University of Cambridge	Author	Design and conceptualization of study, data analysis, drafting of manuscript
Hugh S. Markus	University of Cambridge	Author	Design and conceptualization of study, data analysis, drafting of manuscript

References

1. Pantoni L. Cerebral small vessel disease: from pathogenesis and clinical characteristics to therapeutic challenges. Lancet Neurol 2010;9:689–701.
2. O'Brien JT, Thomas A. Vascular dementia. Lancet 2015;386:1698–1706.
3. Tan R, Traylor M, Rutten-Jacobs L, Markus H. New insights into mechanisms of small vessel disease stroke from genetics. Clin Sci 2017;131:515–531.
4. Chabriat H, Joutel A, Dichgans M, Tournier-Lasserve E, Bousser M-G. CADASIL. Lancet Neurol 2009;8:643–653.
5. Tan RY, Markus HS. Monogenic causes of stroke: now and the future. J Neurol 2015;262:2601–2616.
6. Joutel A, Vahedi K, Corpechot C, et al. Strong clustering and stereotyped nature of *Notch3* mutations in CADASIL patients. Lancet 1997;350:1511–1515.
7. MacArthur JAL, Morales J, Tully RE, et al. Locus Reference Genomic: reference sequences for the reporting of clinically relevant sequence variants. Nucleic Acids Res 2014;42:D873–D878.
8. Simeoni I, Stephens JC, Hu F, et al. A high-throughput sequencing test for diagnosing inherited bleeding, thrombotic, and platelet disorders. Blood 2016;127:2791–2803.

9. 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.
10. Adams HP, 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.
11. Fazekas F, Kleinert R, Offenbacher H, et al. Pathologic correlates of incidental MRI white matter signal hyperintensities. Neurology 1993;43:1683–1689.
12. Lek M, Karczewski KJ, Minikel EV, et al. Analysis of protein-coding genetic variation in 60,706 humans. Nature 2016;536:285–291.
13. Fokkema IF, Taschner PE, Schaafsma GC, Celli J, Laros JF, den Dunnen JT. LOVD v.2.0: the next generation in gene variant databases. Hum Mutat 2011;32:557–563.
14. Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants. Nat Genet 2014;46:310–315.
15. Joutel A, Corpechot C, Ducros A, et al. Notch3 mutations in CADASIL, a hereditary adult-onset condition causing stroke and dementia. Nature 1996;383:707–710.
16. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med 2015;17:405–424.
17. Simeoni I, Shamardina O, Deevi SVV, et al. GRID: Genomics of Rare Immune Disorders: a highly sensitive and specific diagnostic gene panel for patients with primary immunodeficiencies. Biorxiv Epub 2018 Oct 18:431544.
18. Raczy C, Petrovski R, Saunders CT, et al. Isaac: ultra-fast whole-genome secondary analysis on Illumina sequencing platforms. Bioinformatics 2013;29:2041–2043.
19. Hara K, Shiga A, Fukutake T, et al. Association of *HTRA1* mutations and familial ischemic cerebral small-vessel disease. N Engl J Med 2009;360:1729–1739.
20. Tateoka T, Onda H, Hirota K, et al. Unusual case of cerebral small vessel disease with a heterozygous nonsense mutation in *HTRA1*. J Neurol Sci 2016;362:144–146.
21. Verdura E, Hervé D, Scharrer E, et al. Heterozygous *HTRA1* mutations are associated with autosomal dominant cerebral small vessel disease. Brain 2015;138: 2347–2358.
22. Lanfranconi S, Markus HS. *COL4A1* mutations as a monogenic cause of cerebral small vessel disease: a systematic review. Stroke 2010;41:e513–e518.
23. Plaisier E, Chen Z, Gekeler F, et al. Novel *COL4A1* mutations associated with HANAC syndrome: a role for the triple helical CB3[IV] domain. Am J Med Genet A 2010;152A:2550–2555.
24. Weng YC, Sonni A, Labelle-Dumais C, et al. *COL4A1* mutations in patients with sporadic late-onset intracerebral hemorrhage. Ann Neurol 2012;71:470–477.
25. French CR, Seshadri S, Destefano AL, et al. Mutation of *FOXC1* and *PITX2* induces cerebral small-vessel disease. J Clin Invest 2014;124:4877–4881.
26. Kumar M, Chambers C, Dhamija R. Axenfeld-Rieger syndrome and leukoencephalopathy caused by a mutation in *FOXC1*. Pediatr Neurol 2017;66:113–114.
27. DiFrancesco JC, Novara F, Zuffardi O, et al. *TREX1* C-terminal frameshift mutations in the systemic variant of retinal vasculopathy with cerebral leukodystrophy. Neurol Sci 2015;36:323–330.
28. Crow YJ, Chase DS, Lowenstein Schmidt J, et al. Characterization of human disease phenotypes associated with mutations in *TREX1*, *RNASEH2A*, *RNASEH2B*, *RNASEH2C*, *SAMHD1*, *ADAR*, and *IFIH1*. Am J Med Genet A 2015;167A:296–312.
29. Rice G, Newman WG, Dean J, et al. Heterozygous mutations in *TREX1* cause familial chilblain lupus and dominant Aicardi-Goutières syndrome. Am J Hum Genet 2007; 80:811–815.
30. McGlasson S, Rannikmäe K, Bevan S, et al. Rare variants of the 3′-5′ DNA exonuclease *TREX1* in early onset small vessel stroke. Wellcome Open Res 2017;2:106.
31. Firth HV, Richards SM, Bevan AP, et al. DECIPHER: database of chromosomal imbalance and phenotype in humans using ensemble resources. Am J Hum Genet 2009;84:524–533.
32. Jensson O, Palsdottir A, Thorsteinsson L, Arnason A. The saga of cystatin C gene mutation causing amyloid angiopathy and brain hemorrhage: clinical genetics in Iceland. Clin Genet 1989;36:368–377.
33. Di Donato I, Bianchi S, Gallus GN, et al. Heterozygous mutations of *HTRA1* gene in patients with familial cerebral small vessel disease. CNS Neurosci Ther 2017;23:759–765.
34. Nozaki H, Sekine Y, Fukutake T, et al. Characteristic features and progression of abnormalities on MRI for CARASIL. Neurology 2015;85:459–463.
35. Rolfs A, Böttcher T, Zschiesche M, et al. Prevalence of Fabry disease in patients with cryptogenic stroke: a prospective study. Lancet 2005;366:1794–1796.
36. Crutchfield KE, Patronas NJ, Dambrosia JM, et al Quantitative analysis of cerebral vasculopathy in patients with Fabry disease. Neurology 1998;50:1746–1749.
37. Lanthier S, Saposnik G, Lebovic G, Pope K, Selchen D, Moore DF. Prevalence of Fabry disease and outcomes in young Canadian patients with cryptogenic ischemic cerebrovascular events. Stroke 2017;48:1766–1772.
38. Wozniak MA, Kittner SJ, Tuhrim S, et al. Frequency of unrecognized Fabry disease among young European-American and African-American men with first ischemic stroke.
39. Karczewski KJ, Francioli L. The Genome Aggregation Database (gnomAD) [online]. Boston: MacArthur Lab; 2017.
40. Chong M, O'Donnell M, Thijs V, et al. Mendelian genes and risk of intracerebral hemorrhage and small-vessel ischemic stroke in sporadic cases. Stroke 2017;48:2263–2265.
41. Rutten JW, Dauwerse HG, Gravesteijn G, et al. Archetypal *NOTCH3* mutations frequent in public exome: implications for CADASIL. Ann Clin Transl Neurol 2016; 3:844–853.
42. Markus HS, Martin RJ, Simpson MA, et al. Diagnostic strategies in CADASIL. Neurology 2002;59:1134–1138.
43. Peters N, Opherk C, Bergmann T, Castro M, Herzog J, Dichgans M. Spectrum of mutations in biopsy-proven CADASIL: implications for diagnostic strategies. Arch Neurol 2005;62:1091–1094.
44. Opherk C, Peters N, Herzog J, Luedtke R, Dichgans M. Long-term prognosis and causes of death in CADASIL: a retrospective study in 411 patients. Brain 2004;127:2533–2539.
45. Liao Y-C, Hsiao C-T, Fuh J-L, et al. Characterization of CADASIL among the Han Chinese in Taiwan: distinct genotypic and phenotypic profiles. PLoS One 2015;10: e0136501.
46. Ishida C, Sakajiri K, Yoshita M, Joutel A, Cave-Riant F, Yamada M. CADASIL with a novel mutation in exon 7 of *NOTCH3* (C388Y). Intern Med 2006;45:981–985.
47. Kim Y, Choi EJ, Choi CG, et al. Characteristics of CADASIL in Korea: a novel cysteine-sparing *Notch3* mutation. Neurology 2006;66:1511–1516.
48. Moreton FC, Razvi SSM, Davidson R, Muir KW. Changing clinical patterns and increasing prevalence in CADASIL. Acta Neurol Scand 2014;130:197–203.
49. Söder S, Pöschl E. The NC1 domain of human collagen IV is necessary to initiate triple helix formation. Biochem Biophys Res Commun 2004;325:276–280.
50. Eckman CB, Mehta ND, Crook R, et al. A new pathogenic mutation in the *APP* gene (I716V) increases the relative proportion of A beta 42(43). Hum Mol Genet 1997;6: 2087–2089.
51. Revesz T, Holton JL, Lashley T, et al. Genetics and molecular pathogenesis of sporadic and hereditary cerebral amyloid angiopathies. Acta Neuropathol 2009;118: 115–130.
52. Jurkat-Rott K, Freilinger T, Dreier JP, et al. Variability of familial hemiplegic migraine with novel A1A2 Na+/K+-ATPase variants. Neurology 2004;62:1857–1861.
53. Xu X, Yang X, Wu Q, et al. Amplicon resequencing identified parental mosaicism for approximately 10% of "de novo" *SCN1A* mutations in children with Dravet syndrome. Hum Mutat 2015;36:861–872.