Targeted Exonic Sequencing Identifies Novel Causes of Disease in a Cerebral Small Vessel Disease Cohort

Paul Dunn

Queensland University of Technology

Neven Maksemous

Queensland University of Technology

Robert Smith

Queensland University of Technology

Heidi Sutherland

Queensland University of Technology

Larisa Haupt

Queensland University of Technology

Lyn Griffiths (✉ lyn.griffiths@qut.edu.au)

Queensland University of Technology

Research Article

Keywords: Cerebral small vessel disease, CADASIL, Collagen Type IV, Targeted exonic sequencing

Posted Date: February 25th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1333758/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Cerebral small vessel diseases (CSVDs) are a set of conditions that affect the small blood vessels in the brain and can cause severe neurological pathologies such as stroke and vascular dementia. The most common monogenic CSVD is cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy (CADASIL) which is caused by mutations in *NOTCH3*. However, only 15-20% of CADASIL cases referred for genetic testing have pathogenic mutations in *NOTCH3*. Investigation of other monogenic forms of CSVD caused by *HTRA1*, *COL4A1*, *COL4A2*, *CTSA*, *TREX1* and *GLA* were rarely requested as follow-up to CADASIL diagnostic testing. We performed whole exome sequencing for 50 individuals suspected of having CADASIL, but did not exhibit a disease-causing mutation in *NOTCH3*, and applied targeted analysis of all monogenic forms of CSVD including *NOTCH3*, *HTRA1*, *COL4A1*, *COL4A2*, *CTSA*, *TREX1* and *GLA* to look for rare, protein-altering mutations. We identified three mutations affecting the Collagen type IV genes in three individuals likely to be causative of CSVD. This suggests that screening for all monogenic forms of CSVD when one monogenic form is clinically suspected may improve diagnosis in clinically suspected monogenic CSVD. However, despite these findings, the majority of *NOTCH3* negative CSVD cases did not have candidate mutations in known CSVD genes, suggesting that additional genetic factors contributing to the disease are yet to be identified.

Introduction:

Cerebral Autosomal Dominant Arteriopathy with Subcortical Infarcts and Leukoencephalopathy (CADASIL) is a cerebral small vessel disease (CSVD) which is the most common inherited monogenic cause of stroke and vascular dementia in the world[1]. It has an estimated prevalence of 2-4 per 100,000, although this is thought to be under-representative of incidence[2, 3]. The clinical manifestations of CADASIL can be heterogeneous and include subcortical ischaemic events (60-80%); cognitive impairment including dementia and apathy (40-60%); migraine with/without aura (20-40%); mood disturbances including severe depression and manic episodes (~20%); motor disturbances such as gait disturbances (90%), urinary incontinence (80-90%) and pseudobulbar palsy (50%); and other neurological manifestations such as epilepsy/seizures (~10%)[4, 5].

Mutations in *NOTCH3* were originally identified to be cysteine altering in the epidermal growth factor like repeats (EGFR), primarily within exons 3 and 4, however with the advent of next generation sequencing (NGS) technology, there has been an increase in the number and type of mutations found to cause the disease[6]. Further work has also identified that the severity of CADASIL may also be related to the position of the change in the EGFR of CADASIL, where mutations in EGFR 1 - 6 are thought to be causative of a more severe phenotype and mutations in EGFR 7 - 24 thought to be associated with milder phenotypes[7, 8]. Despite our current understanding of the clinical presentation of CADASIL as well as the strong monogenic causal link, genetic testing of clinically referred individuals by the Genomics Research Centre (GRC) has identified mutations in 15-22% of patients, with up to 85% of patients unable to receive a definitive diagnosis[9, 10].

One potential reason for this relatively low diagnostic yield may be phenotypic overlap with other CSVDs as well as other complex disorders such as familial hemiplegic migraine, motor neuron disease, multiple sclerosis and Alzheimer's disease (AD)[11]. The overlapping symptoms of these disorders with CADASIL has been well documented. In particular, six genes have been established to cause five CSVDs with overlapping clinical presentations to CADASIL[12]. The first of these genes is High-Temperature Requirement Serine Peptidase 1 (*HTRA1*) where mutations are known to cause the recessive form of CADASIL, cerebral autosomal recessive arteriopathy with subcortical infarcts and leukoencephalopathy (CARASIL)[13]. The second is Cathepsin-A (*CTSA*), causing Cathepsin-A related arteriopathy with strokes and leukoencephalopathy (CARASAL) where symptoms include migraine, hypertension, sinusitis, asthma and depression with a family history of stroke and cognitive decline, as well as observed white matter hyper-intensities seen via magnetic resonance imaging (MRI)[14].

In addition, *COL4A1* and *COL4A2* cause brain small vessel disease with or without ocular anomalies where clinical features include neurological problems such as hemiparesis, porencephaly, vascular lesions in the brain and peripheral organs, transient ischaemic attacks, adult onset haemorrhagic stroke, periventricular brain abnormalities, white matter hyperintensities, leukoencephalopathy (including cerebellar hypoplasia, cerebral atrophy and vascular changes) and ophthalmological signs such as small/tilted discs, myopia or amblyopia[15, 16]. *TREX1* causes retinal vasculopathy with cerebral leukoencephalopathy where major features of the disease can include retinopathy, cognitive impairment, psychiatric disease, seizures and white matter disease. Finally, *GLA* which encodes for Galactosidase Alpha, is known to cause Fabry disease, cerebrovascular manifestations of which include early stroke/transient ischaemic attacks, white matter lesions, hemiparesis, vertigo or dizziness and other complications associated with vascular disease (diplopia, hemiataxia, memory loss, hearing loss, etc.) [17, 18].

Previous work utilising diagnostic genetic testing of CADASIL patients found evidence to suggest that some CADASIL referred patients with no pathogenic *NOTCH3* mutations had mutations in other monogenic forms of CSVD which caused a CADASIL-like phenotypes[9]. Due to the overlapping phenotypic presentation of these diseases and CADASIL, we sought to investigate the likelihood of other genetic causes of CSVD in a cohort of *NOTCH3* negative patients with a clinical referral for CADASIL.

Methods:

Patient Cohort:

Blood samples were chosen from CADASIL referred patients (n=50) with no pathogenic *NOTCH3* mutations. All patients had approved diagnostic testing for CADASIL with their doctors and ethical approval for this study was obtained through the QUT HREC along with appropriate consents for the patient cohort (Approval Number 1800000611). Diagnostic testing was performed using the Genomics Research Centre (GRC) custom 5-gene panel where only the notch receptor 3 (*NOTCH3*) gene was analysed. Further screening of for calcium voltage-gated channel subunit alpha 1A (*CACNA1A*), ATPase

Na⁺/K⁺ Transporting subunit Alpha 2 (*ATP1A2*), sodium voltage-gated channel alpha subunit 1 (*SCN1A*), and Potassium two pore domain channel subfamily K member 18 (*KCNK18*)[10] which have all previously been associated with other diseases such as familial hemiplegic migraine and migraine with/without aura, episodic ataxia type 2, spinocerebellar ataxia type 6 and epilepsy was also performed to rule out any conditions with overlapping symptoms to CADASIL.

Whole Exome Sequencing:

We used an input of 80ng of DNA used for whole exome sequencing (WES) using the Ion AmpliSeq Exome RDY-kits (Carlsbad, Ca., USA) for library preparation, according to manufactures' instructions (MAN0010084). Completed libraries from samples were quantified using an Invitrogen Qubit 3 Fluorometer (Venlo, Netherlands) and combined at equimolar concentration of 100pMol prior to sequencing. Template preparation, enrichment and chip loading was performed using the Ion P1 Hi-Q Chef Kit (Cat. Number A30011) and 540 Chips on the Thermo Fisher Scientific Ion Chef (Carlsbad, Ca., USA) targeted at 200bp lengths. Sequencing was performed using the Ion Proton and Ion S5+ platforms with sequencing alignment (Hg19) and variant calling was completed via the Ion Torrent software (Carlsbad, Ca., USA).

Analysis and variant Validation

An in-house developed analysis pipeline (VCF-DART) was utilised for targeted analysis of non-amyloidogenic monogenic CSVD genes *NOTCH3*, *HTRA1*, *COL4A1*, *COL4A2*, *CTSA*, *TREX1* and *GLA* (Table 1) [19]. Variants were filtered based on potential functional impacted including non-synonymous, frameshifts, stop-gain or losses, or in canonical splice sites, and with a minor allele frequency (MAF) <0.001. Furthermore, *in silico* pathogenicity tools including Sorting Intolerant From Tolerant (SIFT), Polymorphism Phenotyping v2, Mutation Taster and PredictSNP2 were also used to determine likely pathogenicity [20-23]. Variants were then stratified as potentially disease causing based on the American College of Medical Genetics (ACMG) Guidelines for determining pathogenicity [24]. Suspected disease-causing mutations were validated by Sanger sequencing using BigDye Terminator v3.1 Cycle Sequencing Kits (Applied Biosystems) and analysed using the ABI3500 Genetic Analyser (Life Technologies/ThermoFisher Scientific).

Results:

WES was performed on 50 diagnostic samples referred for CADASIL gene testing that had been screened for disease causing mutations by the GRC custom 5-gene NGS panel (*NOTCH3*, *CACNA1A*, *ATP1A2*, *SCN1A* and *KCNK18*), but reported negative for pathogenic mutations in those genes. Targeted analysis of WES data for *NOTCH3* and the other CSVD genes (*COL4A1*, *COL4A2*, *HTRA1*, *TREX1*, *CTSA* and *GLA*) identified 10 rare mutations in the exonic regions of *COL4A1*, *COL4A2*, *NOTCH3*, *CTSA* and *GLA* (Table 3). There were no rare exonic mutations identified in *HTRA1* or *TREX1* in our cohort which suggests that there may be other genes involved.

Through stratification via the ACMG guidelines for determining pathogenicity, we have identified two likely pathogenic mutations in DGR320 (*COL4A1* – c.3574G>C p.Gly1192Arg) and DGR343 (*COL4A2* – c.290G>A p.Gly97Glu) (Table 3). A third mutation was also identified as a candidate causal mutation in DGR028 (*COL4A2* c.4802C>T p.Ala1601Val) due to the predominantly damaging/disease causing effect that was predicted, although it could only be classified as a variant of unknown significance (VoUS) by the ACMG guidelines.

Clinical information available for the 3 participants is shown in Table 2. Both DGR028 and DGR320 were identified via imaging suggesting small vessel disease, including white matter hyperintensities in DGR320. DGR028 also showed progressive cognitive decline, as well as mood disturbances and a history of psychiatric illnesses. DGR320 had a strong family history of stroke and white matter hyperintensities which appears to follow a dominant inheritance pattern with the symptomology also observed in the proband's sibling and child (Figure 1). DGR343 had no clinical notes provided or made available on request.

A novel mutation was identified in *COL4A1* c.3574G>C p.Gly1192Arg seen in DGR320. Mutations previously reported in *COL4A2*, c.290G>A p.Gly97Glu (rs749501904) and c.4802C>T p.Ala160Val (rs754118201) were present in DGR343 and DGR028, respectively. Familial analysis was completed to include the mother of DGR320 where the same *COL4A1* mutation was identified (Figure 1). Clinical investigation of the mother of the proband showed no obvious clinical signs and symptoms associated with CSVD. Clinical testing of the sibling of the proband identified identical MRI findings, however genetic testing was declined with clinical information of the sibling suffering a stroke before the age of 55.

Another 5 samples had 8 rare (MAF <0.001) exonic mutations that were identified across *COL4A1*, *NOTCH3*, *CTSA* and *GLA*, however, these were considered as unlikely to be causative of the CSVD pathology. From these mutations, those identified in DGR332 (*GLA* c.48T>G p.Leu16Leu), DGR342 (*NOTCH3* c.4914A.G p.Glu1638Glu), DGR351 (*CTSA* c.939T>A p.Thr313Thr) and DGR360 (*NOTCH3* c.120C>G p.Ala40Ala) were synonymous and considered as likely benign according to the ACMG guidelines. There was also two missense mutations identified in *cis* in DGR360 in *NOTCH3* c.154G>A p.Gly52Arg and c.133G>C p.Asp45His which were classified as VoUS. These changes were not considered as candidate mutations for CSVD in this study as there were multiple lines of computation evidence which supported a benign/tolerated effect and the amino acid change was not typical for CADASIL. Despite this, these changes in *NOTCH3* should be functionally assessed as it may give evidence of non-Cysteine altering mutations causing CADASIL in some individuals.

Discussion:

Phenotypes for monogenic forms of CSVD are known to be similar and often at a clinical examination it is difficult to identify the genetic basis for disease. Whilst more common forms of CSVD can be brought upon by lifestyle and environmental factors, when a monogenic form of disease is suspected, using NGS approaches (WES, WGS or targeted gene panels) should be employed. An example of this strategy

working was through targeted WES analysis of CSVD patients screened for *NOTCH3* mutations. Three candidate causal mutations in 2 genes (*COL4A1* and *COL4A2*) were identified in three patients related to monogenic CSVD. Mutations in these genes have previously been identified as causing syndromes with overlapping phenotypic traits to CADASIL including brain small vessel disease with or without ocular anomalies (BSVD1) (MIM # 175780), microangiopathy and leukoencephalopathy, pontine, autosomal dominant (PADMAL) (MIM # 618564) and brain small vessel disease 2 (BSVD2) (MIM # 618564). Shared features of these monogenic conditions with CADASIL include incidences of ischaemic stroke and intracerebral haemorrhages, cognitive impairment/dementia, lacunar infarcts, white matter hyperintensities and/or cerebral microbleeds[12, 25].

A heterozygous p.Ala1601Val (rs754118201) in *COL4A2* identified in DGR028 was identified as the most likely genetic cause for symptoms. However, based on ACMG guidelines it was only determined to be a VoUS. This was in part due to the *COL4A2* mutation affecting the C-terminal tandem repeat region of the *COL4A2* protein rather than the characteristic glycine-altering mutations that disrupt the triple-helix structure of the protein previously associated with pathogenesis[15]. Despite this finding, the C-terminal p.Ala1601Val change may still be implicated in the disease as other mutations in the NC1 domain of the protein, such as p. Ala1690Thr, have been identified to impair *COL4A2* secretion and therefore may cause decreased *COL4A1/COL4A2* heterotrimeric formation in the extracellular matrix[26].

DGR343 had a clinical suspicion of CADASIL with no additional clinical notes available, but based on criteria for suspected CADASIL diagnosis we assumed that the DGR343 had at least one of the following clinical symptoms including stroke-like episodes with neurological deficits and/or dementia/cognitive decline and/or a mood disorder and/or migraine as well as being <55 years old at presentation[27]. A likely-pathogenic mutation was identified in *COL4A2* (rs749501904) due to it being quite rare in gnomAD, with MAF of 7.1×10^{-5} and an allele count of 20/280896, it being the classical Glycine-altering mutation affecting the Alpha-chain of the protein and multiple levels of *in silico* prediction tools identifying it as likely damaging/deleterious[24, 25]. However, despite our interpretation it has previously been classified as having uncertain significance in ClinVar (VCV000521905.1).

DGR320 was identified to have a novel characteristic Glycine-altering mutation in *COL4A1* affecting the triple-helix domain of the Collagen type IV chain. This type of mutation has previously been linked to characteristic *COL4A1* and *COL4A2* related small vessel diseases and a recent recommendation from the European Academy of Neurology indicates that all Gly-altering mutations in this region should be denoted as pathogenic[25]. DGR320 also presented with a strong family history of white matter hyperintensities as well as other neurological and neurovascular symptoms such as migraine, vision issues and stroke identified in a sibling. Based on this being a novel mutation, that follows the same classical Glycine-change seen in *COL4A1* related SVD, a family history that indicated that the symptoms match a monogenic form of CSVD and multiple *in silico* tools identifying it as damaging/deleterious with no tools predicting a tolerated effect, it was classified as a likely pathogenic mutation.

However, further investigation of this mutation in the mother of the proband identified the same mutation in *COL4A1*. This was despite the fact the proband's mother had no detectable clinical symptoms associated with SVD, including previous known ischaemic/haemorrhagic events, or signs of cognitive decline. Unfortunately, imaging via MRI was not conducted, and it remains unclear whether she exhibited the same white matter hyperintensities seen in the extended family of the proband. This mutation may also cause variable phenotype, as this is a phenomena that has been examined related to *COL4A1* and *COL4A2* mutations[28]. Studies have shown that phenotypic heterogeneity has been noted in both broad intra- and interfamilial variation with some evidence suggesting that there are reduced penetrance mechanisms[29]. One example found heterozygous *COL4A1/COL4A2* mutations that cause severe phenotypes in affected infants, while parents with the same mutation show no clinical signs of disease. As such, it has been hypothesised that *COL4A1/2* mutations may serve as risk factors for CSVD phenotypes and additional risk factors may be required for disease to be seen [29-31]. In line with this observation, other functional mutations identified in DGR320 may hold clues into some of the clinical features identified.

The Genomic Research Centre has previously found that targeted gene investigations of NGS data is an effective method for identifying causal mutations of monogenic conditions[10, 32, 33]. Furthermore, other labs have successfully used these approaches when investigating individuals that have some of the major symptoms of CSVD including lacunar stroke, cognitive impairment and leukodystrophy[34, 35]. Interestingly, these studies identified *HTRA1* mutations as the next most common cause of these conditions, where comparatively our findings failed to identify any causative mutations in this gene[34, 35]. This in turn may be due to the clinical suspicion of CADASIL where only individuals with a perceived autosomal dominant inherited cause of CSVD was suspected.

Whilst we were able to identify candidate mutations within three patients, we were not able to functionally assess these mutations to further elucidate their pathogenesis. Furthermore, only one family member from one of the probands (DGR320) consented for further investigation via segregation analysis. In addition, this study was not able to identify mutations in the remaining 47 probands. It is possible that there are mutations in other genes not investigated in detail, including possible mutations in the mitochondrial genome causative of mitochondrial encephalopathy with lactic acidosis and stroke-like episodes (MELAS), in some of these individuals that was not detected[36-38]. This is a limitation in using Ion AmpliSeq exome RDY sequencing which does not amplify the mitochondrial genome which limited our scope of investigation. This study was also limited by the number of samples and the lack of more detailed clinical information available for some patients, requiring the assumption of sufficient symptomology being used as a premise for clinical CADASIL testing. This decreased our ability to identify if other genes were indirectly affecting the clinical phenotypes present within the study cohort.

Despite these limitations, we were able to identify monogenic causes of CSVD in a small cohort. Further work should focus on the identification of novel genetic causes of CSVD in conjunction with functional and segregation studies which can then be implemented with diagnostic testing to aid in better treatments for individuals. This work highlights a need for better diagnostic testing, moving away from

single gene targeted testing and to further implement multi-gene targeted whole exome or whole genome sequencing for diagnostic testing of CSVD in the future.

Conclusion:

Overall, three probands were identified to have mutations in *COL4A1* and *COL4A2* causing related CSVD conditions in a *NOTCH3* negative cohort. While we were able to identify clinically relevant mutations in these three individuals, we were not able to identify CSVD mutations in the remaining 47 probands. Further investigation of the exome sequencing data of these individuals may yield other potential novel causes of CSVD. These results show that genetic investigation of patients with a clinical indication for CADASIL should also include the CSVD genes, *COL4A1* and *COL4A2*, and that there may be other novel causes of these diseases that are yet to be discovered.

Declarations:

Funding:

This work was supported through the National Health and Medical Research Council of Australia (GNT1168601) Dora Lush Biomedical Postgraduate Research scholarship.

Conflicts of Interest:

The authors declare that there are no conflicts of interest

Availability of Data:

All data relevant for this study is included within this manuscript or has been submitted to relevant databases. This includes genetic mutation information which has been submitted to ClinVar and dbSNP. Due to the medical nature and ethical agreements for use of samples, any further information may be made available on request.

Code availability:

Not applicable

Author's contributions:

Work was conceived by PJD, LRG and LMH. Laboratory methods to generated whole exome sequencing was completed by PJD and NM and previous diagnostic testing and analysis completed by NM and RAS. Analytical approaches included input from HS and PJD. Writing and editing of the work was completed by PJD, NM, RAS, HGS, LMH and LRG.

Ethics Approval:

Ethical approval was completed through the Queensland University of Technology (QUT) human research ethics committee (HREC) (Approval Number 1800000611)

Consent to Participate:

Informed consent was obtained in line with the ethics approval through the QUT HREC (Approval Number 1800000611) and is in line with the 1964 Declaration of Helsinki and its later amendments

Consent for Publication:

Consent for publication was obtained in line with the ethics approval through the QUT HREC (Approval Number 1800000611)

Acknowledgements:

We would like to acknowledge support for this study through the NHMRC Dora Lush Biomedical Postgraduate Research scholarship. Furthermore, the team at the Genomics Research Centre for all their work and many discussions. Finally, the patients and clinicians who helped make it possible to complete this work.

References:

1. Joutel, A., et al., *Notch3 mutations in CADASIL, a hereditary adult-onset condition causing stroke and dementia*. Nature, 1996. **383**(6602): p. 707-10.
2. Razvi, S.S., et al., *The prevalence of cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy (CADASIL) in the west of Scotland*. J Neurol Neurosurg Psychiatry, 2005. **76**(5): p. 739-41.
3. Choi, J.C., *Cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy: a genetic cause of cerebral small vessel disease*. J Clin Neurol, 2010. **6**(1): p. 1-9.
4. Choudhary, S., et al., *Cerebral Autosomal Dominant Arteriopathy with Subcortical Infarcts and Leukoencephalopathy (CADASIL)*. J Clin Aesthet Dermatol, 2013. **6**(3): p. 29-33.
5. Di Donato, I., et al., *Cerebral Autosomal Dominant Arteriopathy with Subcortical Infarcts and Leukoencephalopathy (CADASIL) as a model of small vessel disease: update on clinical, diagnostic, and management aspects*. BMC Med, 2017. **15**(1): p. 41.
6. Joutel, A., et al., *Notch3 mutations in cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy (CADASIL), a mendelian condition causing stroke and vascular dementia*. Ann N Y Acad Sci, 1997. **826**: p. 213-7.
7. Rutten, J.W., et al., *The effect of NOTCH3 pathogenic variant position on CADASIL disease severity: NOTCH3 EGFr 1–6 pathogenic variant are associated with a more severe phenotype and lower survival compared with EGFr 7–34 pathogenic variant*. Genetics in Medicine, 2019. **21**(3): p. 676-682.

8. Rutten, J.W., et al., *Archetypal NOTCH3 mutations frequent in public exome: implications for CADASIL*. Ann Clin Transl Neurol, 2016. **3**(11): p. 844-853.
9. Dunn, P.J., et al., *Investigating diagnostic sequencing techniques for CADASIL diagnosis*, in *Human Genomics*. 2020. p. 2.
10. Maksemous, N., et al., *Targeted next generation sequencing identifies novel NOTCH3 gene mutations in CADASIL diagnostics patients*. Hum Genomics, 2016. **10**(1): p. 38.
11. Saleem, S., et al., *Periventricular Hyperintensities Mimicking Multiple Sclerosis*. Cureus, 2019. **11**(8): p. e5326-e5326.
12. Choi, J.C., *Genetics of cerebral small vessel disease*. J Stroke, 2015. **17**(1): p. 7-16.
13. Hara, K., et al., *Association of HTRA1 mutations and familial ischemic cerebral small-vessel disease*. N Engl J Med, 2009. **360**(17): p. 1729-39.
14. Hwang, Y.T., et al., *Brainstem phenotype of cathepsin A-related arteriopathy with strokes and leukoencephalopathy*. Neurology: Genetics, 2017. **3**(4): p. e165.
15. Verbeek, E., et al., *COL4A2 mutation associated with familial porencephaly and small-vessel disease*. Eur J Hum Genet, 2012. **20**(8): p. 844-51.
16. Rannikmäe, K., et al., *Common variation in COL4A1/COL4A2 is associated with sporadic cerebral small vessel disease*. Neurology, 2015. **84**(9): p. 918-926.
17. Desnick, R.J., et al., *Fabry disease, an under-recognized multisystemic disorder: expert recommendations for diagnosis, management, and enzyme replacement therapy*. Ann Intern Med, 2003. **138**(4): p. 338-46.
18. Mitsias, P. and S.R. Levine, *Cerebrovascular complications of Fabry's disease*. Ann Neurol, 1996. **40**(1): p. 8-17.
19. Benton, M.C., et al., *Variant Call Format-Diagnostic Annotation and Reporting Tool: A Customizable Analysis Pipeline for Identification of Clinically Relevant Genetic Variants in Next-Generation Sequencing Data*. J Mol Diagn, 2019. **21**(6): p. 951-960.
20. Bendl, J., et al., *PredictSNP2: A Unified Platform for Accurately Evaluating SNP Effects by Exploiting the Different Characteristics of Variants in Distinct Genomic Regions*. PLoS Comput Biol, 2016. **12**(5): p. e1004962.
21. Sim, N.-L., et al., *SIFT web server: predicting effects of amino acid substitutions on proteins*. Nucleic acids research, 2012. **40**(Web Server issue): p. W452-W457.
22. Adzhubei, I., D.M. Jordan, and S.R. Sunyaev, *Predicting functional effect of human missense mutations using PolyPhen-2*. Current protocols in human genetics, 2013. **Chapter 7**: p. Unit7.20-Unit7.20.
23. Schwarz, J.M., et al., *MutationTaster evaluates disease-causing potential of sequence alterations*. Nature Methods, 2010. **7**: p. 575.
24. Richards, 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**(5): p. 405-24.
25. Mancuso, M., et al., *Monogenic cerebral small-vessel diseases: diagnosis and therapy. Consensus recommendations of the European Academy of Neurology*. Eur J Neurol, 2020. **27**(6): p. 909-927.
 26. Jeanne, M., et al., *COL4A2 mutations impair COL4A1 and COL4A2 secretion and cause hemorrhagic stroke*. American journal of human genetics, 2012. **90**(1): p. 91-101.
 27. Mizuta, I., et al., *New diagnostic criteria for cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy in Japan*. J Neurol Sci, 2017. **381**: p. 62-67.
 28. Kuo, D.S., C. Labelle-Dumais, and D.B. Gould, *COL4A1 and COL4A2 mutations and disease: insights into pathogenic mechanisms and potential therapeutic targets*. Hum Mol Genet, 2012. **21**(R1): p. R97-110.
 29. Meuwissen, M.E.C., et al., *The expanding phenotype of COL4A1 and COL4A2 mutations: clinical data on 13 newly identified families and a review of the literature*. Genetics in Medicine, 2015. **17**(11): p. 843-853.
 30. Gould, D.B., et al., *Col4a1 mutation causes endoplasmic reticulum stress and genetically modifiable ocular dysgenesis*. Human Molecular Genetics, 2007. **16**(7): p. 798-807.
 31. Labelle-Dumais, C., et al., *COL4A1 Mutations Cause Ocular Dysgenesis, Neuronal Localization Defects, and Myopathy in Mice and Walker-Warburg Syndrome in Humans*. PLOS Genetics, 2011. **7**(5): p. e1002062.
 32. Dunn, P.J., et al., *Tiered analysis of whole-exome sequencing for epilepsy diagnosis*. Molecular Genetics and Genomics, 2020. **295**(3): p. 751-763.
 33. Sutherland, H.G., et al., *Comprehensive Exonic Sequencing of Hemiplegic Migraine-Related Genes in a Cohort of Suspected Proband Identifies Known and Potential Pathogenic Variants*. Cells, 2020. **9**(11).
 34. Chen, Z., et al., *High Diagnostic Utility Incorporating a Targeted Neurodegeneration Gene Panel With MRI Brain Diagnostic Algorithms in Patients With Young-Onset Cognitive Impairment With Leukodystrophy*. Frontiers in Neurology, 2021. **12**(50).
 35. Tan, R.Y.Y., et al., *How common are single gene mutations as a cause for lacunar stroke? A targeted gene panel study*. Neurology, 2019. **93**(22): p. e2007-e2020.
 36. Ikeda, T., et al., *Mitochondrial DNA 3243A>T mutation in a patient with MELAS syndrome*. Human Genome Variation, 2018. **5**(1): p. 25.
 37. El-Hattab, A.W., et al., *MELAS syndrome: Clinical manifestations, pathogenesis, and treatment options*. Molecular Genetics and Metabolism, 2015. **116**(1): p. 4-12.
 38. Sproule, D.M. and P. Kaufmann, *Mitochondrial Encephalopathy, Lactic Acidosis, and Stroke-like Episodes*. Annals of the New York Academy of Sciences, 2008. **1142**(1): p. 133-158.
 39. Joutel, A., et al., *A human homolog of bacterial acetolactate synthase genes maps within the CADASIL critical region*. Genomics, 1996. **38**(2): p. 192-8.

40. Joutel, A., et al., *Strong clustering and stereotyped nature of Notch3 mutations in CADASIL patients*. Lancet, 1997. **350**(9090): p. 1511-5.
41. Haffner, C., R. Malik, and M. Dichgans, *Genetic factors in cerebral small vessel disease and their impact on stroke and dementia*. J Cereb Blood Flow Metab, 2016. **36**(1): p. 158-71.
42. Nishimoto, Y., et al., *A novel mutation in the *HTRA1* gene causes CARASIL without alopecia*. Neurology, 2011. **76**(15): p. 1353-1355.
43. Tikka, S., et al., *CADASIL and CARASIL*. Brain Pathol, 2014. **24**(5): p. 525-44.
44. Haffner, C. and H.V. Vinters, *CADASIL, CARASIL, CARASAL: The linguistic subtleties of cerebral small vessel disease*. Neurology, 2016. **87**(17): p. 1752-1753.
45. Verdura, E., et al., *Heterozygous HTRA1 mutations are associated with autosomal dominant cerebral small vessel disease*. Brain, 2015. **138**(8): p. 2347-2358.
46. Kint, J.A., *Fabry's disease: alpha-galactosidase deficiency*. Science, 1970. **167**(3922): p. 1268-9.
47. Ishii, S., H. Sakuraba, and Y. Suzuki, *Point mutations in the upstream region of the alpha-galactosidase A gene exon 6 in an atypical variant of Fabry disease*. Hum Genet, 1992. **89**(1): p. 29-32.
48. Plancher, J.M., et al., *Case of Small Vessel Disease Associated with COL4A1 Mutations following Trauma*. Case Reports in Neurology, 2015. **7**(2): p. 142-147.
49. Chowdhury, D., et al., *The Exonuclease TREX1 Is in the SET Complex and Acts in Concert with NM23-H1 to Degrade DNA during Granzyme A-Mediated Cell Death*. Molecular Cell, 2006. **23**(1): p. 133-142.
50. Richards, A., et al., *C-terminal truncations in human 3'-5' DNA exonuclease TREX1 cause autosomal dominant retinal vasculopathy with cerebral leukodystrophy*. Nat Genet, 2007. **39**(9): p. 1068-70.
51. DiFrancesco, J.C., et al., *TREX1 C-terminal frameshift mutations in the systemic variant of retinal vasculopathy with cerebral leukodystrophy*. Neurol Sci, 2015. **36**(2): p. 323-30.
52. Bugiani, M., et al., *Cathepsin A-related arteriopathy with strokes and leukoencephalopathy (CARASAL)*. Neurology, 2016. **87**(17): p. 1777-1786.

Tables:

Table 1 - Tier 1 gene list with associated OMIM listed conditions

Genes	Condition	MIM#	References
<i>NOTCH3</i>	Cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy (CADASIL)	125310	[1, 39-41]
<i>HTRA1</i>	Cerebral autosomal recessive arteriopathy with subcortical infarcts and leukoencephalopathy (CARASIL)	600142,	[13, 42-44]
	Cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy, type 2 (CADASIL, type 2)	616779	[45]
<i>GLA</i>	Fabry Disease	301500	[12, 41, 46, 47]
<i>COL4A1</i>	[Haemorrhage, intracerebral, susceptibility to]	614519,	[12, 16, 28, 48]
	Brain Small vessel disease with or without ocular anomalies	175780,	
	Microangiopathy and leukoencephalopathy, pontine, autosomal dominant (PADMAL)	618564	
<i>COL4A2</i>	[Haemorrhage, intracerebral, susceptibility to]	614519,	[15, 16, 28]
	Brain Small vessel disease 2	614483	
<i>TREX1</i>	Retinal Vasculopathy with Cerebral Leukoencephalopathy	192315	[49-51]
<i>CTSA</i>	Cathepsin-A Related Arteriopathy with Strokes and Leukoencephalopathy	N/A	[44, 52]

Table 2 - Clinical information for participants with the Tier 1 mutations detected. ^a CADASIL phenotype of stroke/cognitive decline/abnormal MRI were assumed as they were referred by a clinician for NOTCH3 testing originally. PTSD – post traumatic stress disorder, GAD – generalized anxiety disorder, MDD – major depressive disorder, OCD – obsessive compulsive disorder.

Sample	Sex	Age	Clinical Indication
DGR028	M	64	Progressive cognitive decline over 10 years (worst past 6months). Hx Migraines. Background of PTSD, GAD, MDD, OCD. Declining function at home. Query CADASIL. MRI: Small vessel disease
DGR320	F	49	<p>No clinical notes provided with initial testing- Patient provided: white matter lesions + blurry vision on/off. Sister and daughter (10 years old) also have white matter lesions. Sister has had previous strokes. Son has migraines w/aura.</p> <p>Clinical consultation letter included right optic disc swelling. MRI identified multiple areas of white matter lesions, including optic nerve - suspicious of multiple sclerosis. MRI remained unchanged over time and there are no lesions on spinal cord. CADASIL skin Bx was negative. Sister's MRI scan identical indicating a hereditary condition. PDGFB/R mutations screening was also recommended.</p>
DGR332	F	20	Recurrent hemiparesis. Tests requested: CADASIL genotype, haemophilia screen
DGR336	M	51	FHx CVAs (early). Meiparetic migraine. MRI - white matter changes around corpus callosum. Skin Bx EM -ve
DGR342	F	65	Query central nervous system demyelination/stroke. 18/12 memory loss. MRI features present of WMH. Few vascular risks. Patient suffers from headaches
DGR343	F	39	No clinical notes ^a
DGR351	F	64	No clinical notes ^a
DGR360	M	60	Extensive MRI brain ischemic changes Querying CADASIL gene positive

Table 3 - Mutations detected with in silico pathogenicity scores and population frequencies included for rare exonic mutations (MAF < 0.001) from the monogenic CSVD targeted analysis. Green highlighted cells indicate the a predicted benign/tolerated effect, whilst pink highlighted cells indicate a damaging/deleterious effect. Blue highlighted cells indicate and unknown predicted effect for that in silico tool.

Sample	DGR320	DGR336	DGR343	DGR028	DGR342	DGR360	DGR360	DGR360	DGR351	DGR332
CHROM	13	13	13	13	19	19	19	19	20	X
POS	110823062	110835579	111077190	111160489	15281342	15308354	15308375	15308388	44523477	100662844
Gene	COL4A1	COL4A1	COL4A2	COL4A2	NOTCH3	NOTCH3	NOTCH3	NOTCH3	CTSA	GLA
Mutation	ENST00000375820.4:c.3574G>C	ENST00000375820.4:c.1942C>G	ENST00000360467.5:c.290G>A	ENST00000360467.5:c.4802C>T	ENST00000263388.2:c.4914A>G	ENST00000263388.2:c.154G>A	ENST00000263388.2:c.133G>C	ENST00000263388.2:c.120C>G	ENST00000372484.3:c.939T>A	ENST00000218516.3:c.48T>G
Protein	ENSP00000364979.4:p.Gly1192Arg	ENSP00000364979.4:p.Pro648Ala	ENSP00000353654.5:p.Gly97Glu	ENSP00000353654.5:p.Ala1601Val	ENSP00000263388.1:p.Glu1638%3D	ENSP00000263388.1:p.Gly52Arg	ENSP00000263388.1:p.Asp45His	ENSP00000263388.1:p.Ala40%3D	ENSP00000361562.3:p.Thr313%3D	ENSP00000218516.3:p.Leu16%3D
SIFT	0	0.17	0	0.06	N/A	0	0.03	N/A	N/A	N/A
PolyPhen2	1	0.9	1	0.996	N/A	0.998	0.219	N/A	N/A	N/A
MutationTaster	Disease causing - aa changed, protein maybe affected, splice site	Disease causing - aa changed, protein maybe affected, splice site	Disease causing - aa changed, protein maybe affected	Disease causing - aa changed, protein maybe affected, splice site	Synonymous polymorphism (non-disease causing)	Disease causing - aa changed, protein maybe affected, splice site	Disease causing - aa changed, protein maybe affected, splice site	Disease causing - splice site changes (synonymous mutation)	Synonymous polymorphism (non-disease causing)	Synonymous polymorphism (non-disease causing)
PredictSNP2	D-87	B-63	B-63	D-87	B-96	D-82	B-63	B-96	B-96	B-93
CADD	D-84	B-61	D-63	B-68	B-91	B-56	B-56	B-79	B-83	B-92
DANN	D-70	B-84	B-76	D-77	B-97	D-77	D-60	B-90	B-90	B-96
FATHMM	D-80	D-83	D-82	D-53	B-96	B-75	B-63	B-73	B-87	D-60
FunSeq2	?	D-62	D-61	D-61	B-93	D-62	D-62	B-93	B-93	B-93
GWAVA	D-51	D-51	D-52	D-50	B-60	B-56	B-53	B-60	B-59	?
dbSNP	N/A	rs1413099763	rs749501904	rs754118201	rs149222385	rs148166997	rs142031490	rs146904189	rs143206710	rs201449986
ClinVar Accession	N/A	N/A	VCV000521905.2	N/A	N/A	VCV000447788.2	VCV000447782.2	VCV000256119.2	N/A	VCV000042455.8
ClinVar_Clinical Significance	N/A	N/A	uncertain_significance	N/A	N/A	uncertain_significance	uncertain_significance	conflicting_interpretations_of_pathogenicity	N/A	uncertain_significance&likely_benign&conflicting_interpretations_of_pathogenicity
gnomAD MAF	N/A	4.01E-06	6.41E-05	1.60E-05	2.50E-05	1.22E-05	8.18E-06	2.47E-05	3.98E-06	0.0001963
ACMG Pathogenic Criteria	PS1, PM1, PP2, PP3, PP4	PP3, PP4	PS1, PP3, PP4	PP3, PP4	N/A	PP3, PP4	PP3, PP4	N/A	N/A	N/A
ACMG Benign Criteria	N/A	BP4	BP4	BP4	BP4, BP7	BP4	BP4	BP4, BP7	BP4, BP7	BP4, BP7
ACMG Rating	Likely Pathogenic	Uncertain Significance	Likely Pathogenic	Uncertain Significance	Likely Benign	Uncertain Significance	Uncertain Significance	Likely Benign	Likely Benign	Likely Benign
Candidate Mutation	Yes	No	Yes	Yes	No	No	No	No	No	No

Figures

Figure 1

DGR320 Pedigree showing CSVD related symptoms created using progeny pedigree drawing software (<https://pedigree.progenygenetics.com/>). Only the proband (DGR320) and the proband's mother have consented to genetic testing for this study.