



Monogenic Causes in Familial Stroke Across Intracerebral Hemorrhage and Ischemic Stroke Subtypes Identified by Whole-Exome Sequencing

Li-Hsin Chang¹ · Nai-Fang Chi^{2,3} · Chun-Yu Chen² · Yung-Shuan Lin² · Shao-Lun Hsu² · Jui-Yao Tsai² · Hui-Chi Huang² · Chun-Jen Lin^{2,3} · Chih-Ping Chung^{2,3} · Chien-Yi Tung⁵ · Chung-Jiuan Jeng^{4,6} · Yi-Chung Lee^{2,3,4} · Yo-Tsen Liu^{1,2,3,4} · I-Hui Lee^{1,2,3,4}

Received: 9 September 2022 / Accepted: 20 December 2022 / Published online: 29 December 2022
© The Author(s) 2022

Abstract

Whole exome sequencing (WES) has been used to detect rare causative variants in neurological diseases. However, the efficacy of WES in genetic diagnosis of clinically heterogeneous familial stroke remains inconclusive. We prospectively searched for disease-causing variants in unrelated probands with defined familial stroke by candidate gene/hotspot screening and/or WES, depending on stroke subtypes and neuroimaging features at a referral center. The clinical significance of each variant was determined according to the American College of Medical Genetics guidelines. Among 161 probands (mean age at onset 53.2 ± 13.7 years; male 63.4%), 33 participants (20.5%) had been identified with 19 pathogenic/likely pathogenic variants (PVs; WES applied $152/161 = 94.4\%$). Across subtypes, the highest hit rate (HR) was intracerebral hemorrhage (ICH, $7/18 = 38.9\%$), particularly with the etiological subtype of structural vasculopathy ($4/4 = 100\%$, PVs in *ENG*, *KRIT1*, *PKD1*, *RNF213*); followed by ischemic small vessel disease (SVD, $15/48 = 31.3\%$; PVs in *NOTCH3*, *HTRA1*, *HBB*). In contrast, large artery atherosclerosis (LAA, $4/44 = 9.1\%$) and cardioembolism ($0/11 = 0\%$) had the lowest HR. *NOTCH3* was the most common causative gene ($16/161 = 9.9\%$), presenting with multiple subtypes of SVD ($n = 13$), ICH ($n = 2$), or LAA ($n = 1$). Importantly, we disclosed two previously unreported PVs, *KRIT1* p.E379* in a familial cerebral cavernous malformation, and *F2* p.F382L in a familial cerebral venous sinus thrombosis. The contribution of monogenic etiologies was particularly high in familial ICH and SVD subtypes in our Taiwanese cohort. Utilizing subtype-guided hotspot screening and/or subsequent WES, we unraveled monogenic causes in 20.5% familial stroke probands, including 1.2% novel PVs. Genetic diagnosis may enable early diagnosis, management and lifestyle modification.

Yo-Tsen Liu and I.-Hui Lee have contributed equally to this work.

✉ Yo-Tsen Liu
ytliu2@vghtpe.gov.tw

✉ I-Hui Lee
ihlee@vghtpe.gov.tw

¹ Institute of Brain Science, National Yang Ming Chiao Tung University, Taipei, Taiwan

² Department of Neurology, Neurological Institute, Taipei Veterans General Hospital, No. 201, Sec. 2, Shipai Rd., Beitou District, 11217 Taipei City, Taiwan

³ School of Medicine, National Yang Ming Chiao Tung University, Taipei, Taiwan

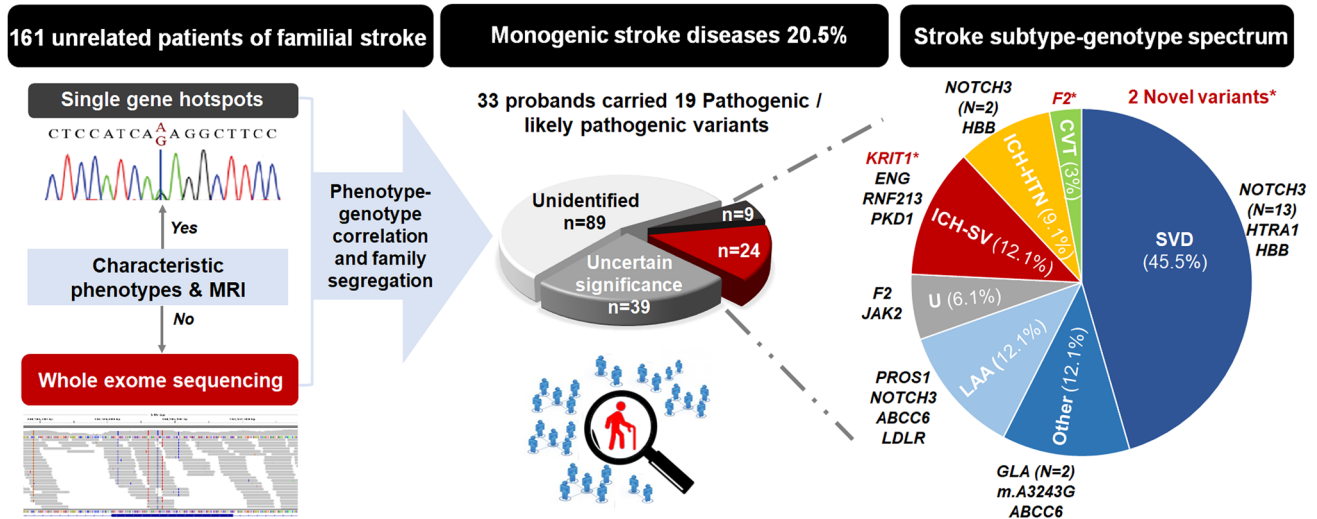
⁴ Brain Research Center, National Yang Ming Chiao Tung University, Taipei, Taiwan

⁵ Cancer Progression Research Center, National Yang Ming Chiao Tung University, Taipei, Taiwan

⁶ Institute of Anatomy and Cell Biology, School of Medicine, National Yang Ming Chiao Tung University, Taipei, Taiwan

Graphical Abstract

Among 161 familial stroke probands, 33 (20.5%) had been identified pathogenic or likely pathogenic monogenic variants related to stroke. The positive hit rate among all subtypes was high in intracerebral hemorrhage (ICH) and ischemic small vessel disease (SVD). Notably, two previously unreported variants, *KRIT1* p.E379* in a familial cerebral cavernous malformation and *F2* p.F382L in familial cerebral venous sinus thrombosis, were disclosed. *CVT* cerebral venous thrombosis; *HTN* Hypertensive subtype; *LAA* large artery atherosclerosis; *SV* structural vasculopathy; *U* Undetermined.



Keywords Stroke · Family history · WES · Monogenic · Taiwan

Introduction

Stroke is a leading cause of death and disability, which is attributed to complex genetic and environmental risk factors (Starby et al. 2014). Familial aggregation is a risk factor that increases stroke susceptibility (Chung et al. 2016; Ilinca et al. 2016). Whole-exome sequencing (WES) has become an effective tool to detect rare monogenic disorders (Rabbani et al. 2014). Lingren and his colleagues reported possibly disease-causing variants in six out of 22 familial young stroke patients (under 55 years old) by WES, revealing a higher hit rate (27.3%) than targeted gene panels (Ilinca et al. 2020). However, only a few WES studies are available addressing limited and specific cerebrovascular diseases (Santoro et al. 2018; Carrera et al. 2016; Mukawa et al. 2017; Sauvigny et al. 2020). The strategy and efficacy to apply clinical WES in familial stroke are still elusive. The main challenges come from the high heterogeneity in both phenotypes and genotypes of familial stroke patients, as well as inconsistent variant interpretation from WES data across laboratories.

There is a lack of comprehensive WES studies in Asian stroke patients. Compared to Caucasians, East Asians are more susceptible to small vessel disease (SVD), intracerebral hemorrhage (ICH), and intracranial large artery

atherosclerosis (LAA) (Toyoda et al. 2015; Wei et al. 2021; Gurdasani et al. 2019; Kim and Kim 2014). Genetic variations may play an essential role in the different distribution of stroke subtypes among ethnic populations. In East Asians, cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy (CADASIL) caused by *NOTCH3* mutations (Lee et al. 2019), CADASIL2 and CARASIL (recessive) caused by *HTRA1* mutations (Liao et al. 2015), and moyamoya disease (MD) caused by *RNF213* mutations, are particularly prevalent (Bang et al. 2020). To delineate the monogenic causes and the genotype–phenotype correlations of familial stroke in Asians, here we investigated the frequency of likely disease-causing variants across a full spectrum of stroke subtypes in a Taiwanese cohort of familial stroke patients aged 18–79 using prevalent gene/hotspot screening and/or WES, depending on stroke subtypes and neuroimaging features. The efficacy of this integrative diagnostic pipeline and the therapeutic implications for the unraveled monogenic stroke disorders were addressed.

Materials and Methods

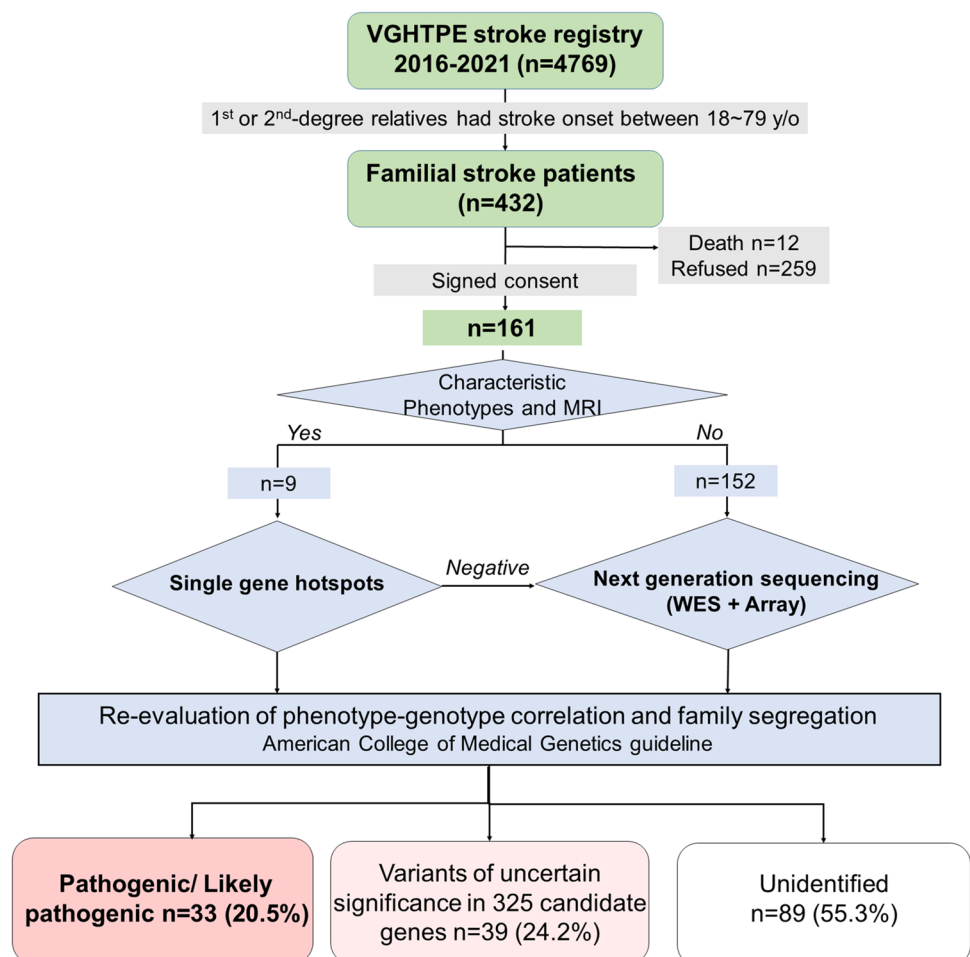
Patient Enrollment and Study Design

This study was performed with the approval of the Taipei Veterans General Hospital Institutional Review Board (2016-09-010C, 2020-02-007B) in accordance with the tenets of the Declaration of Helsinki. We prospectively screened 432 patients with familial ischemic or hemorrhagic stroke between September 2016 and July 2021 from 4769 inpatients consecutively admitted to Taipei Veterans General Hospital, a cohort from one of the largest tertiary referral centers in Taiwan. A positive family history was defined as a patient and at least one of the first or second-degree relatives having their first stroke between 18 and 79 years old. The inclusion criteria and diagnostic algorithm of this study are summarized in Fig. 1. The diagnosis of stroke was established by brain MRI/MRA and corresponding neurological deficits. The subtype classification was determined according to the classification of TOAST (Trial of Org 10172 Acute Stroke Treatment)

for ischemic stroke (IS) and SMASH-U for intracerebral hemorrhage (ICH) (Adams et al. 1993; Meretoja et al. 2012). Cerebral venous sinus thrombosis-induced stroke was separately subtyped.

Totally, 161 probands fulfilled and provided consent to participate. Patients with characteristic clinical and neuroimaging features were first screened by conventional Sanger sequencing of specific genetic hotspots, including *NOTCH3* p.R544C for CADASIL, m.3243A>G for mitochondrial encephalopathy, lactic acidosis, and stroke-like episodes (MELAS), and the *GLA* gene for Fabry disease. The remaining undiagnosed probands were submitted to an integrative approach which combined WES and a supplementary single nucleotide polymorphism (SNP) array customized for the genomic variants of Taiwanese population, including intronic and mitochondrial variants (Axiom Genome-Wide Taiwan Biobank Array Plate, TWB 2.0 National Center for Genome Medicine, Academia Sinica, Taiwan) (Wei et al. 2021). Rare variants are defined by the minor allele frequency (MAF) equal to or less than 0.1% in the global population among gnomAD database. The MAF in local control population was obtained from Taiwan

Fig. 1 The inclusion criteria and diagnostic algorithm of familial stroke. We prospectively screened 432 patients with familial history of stroke in 4769 consecutive inpatients admitted to Taipei Veterans General Hospital (VGHTPE stroke registry) between 2016 and 2021. Totally 161 probands fulfilled the criteria and participated. Nine probands were diagnosed by sequencing specific gene hotspots and the other 152 probands were sent for whole-exome sequencing (WES) and a customized array. Thirty-three probands were identified carried pathogenic/likely pathogenic variants and 39 carried variants of uncertain significance among 325 stroke-related genes. The other 89 probands were categorized as unidentified, including 42 probands with no variants identified after variant filtering and other 47 with incompatible variants with phenotype or inheritance mode. MRI: magnetic resonance imaging



Biobank whole-genome sequencing dataset of 1517 normal subjects. The clinical significance of the genetic variant was determined according to the American College of Medical Genetics (ACMG) guideline (Kalia et al. 2017). Detailed clinical data including the medical records, intensive interviewing, laboratory, and imaging studies were reviewed by two neurologists to confirm the phenotype-genotype correlation and segregation. The level of genetic diagnosis was categorized into (1) the definite genetic diagnosis with a pathogenic/likely pathogenic variant (PV) (i.e., monogenic stroke), (2) variants of uncertain significance (VUS) in the 325 stroke-associated genes (Supplementary Table S1), and (3) unidentified (Fig. 1).

Brain Magnetic Resonance Imaging (MRI)

All probands received MRI within 7 days after stroke onset in a 1.5 Tesla MRI scanner (Signa HDxt, GE Healthcare, Milwaukee, WI). A regular stroke MRI protocol consisted of T1-weighted imaging, T2-weighted and Fluid Attenuated Inversion Recovery (FLAIR) imaging, Diffusion-Weighted Imaging (DWI), Apparent Diffusion Coefficient (ADC), and time-of-flight (TOF) MR Angiography (MRA). For probands with SVD or ICH, T2*-weighted Susceptibility-Weighted Angiography (SWAN) was additionally applied. For probands with suspected dissection, a high-resolution contrast-enhanced vessel wall imaging was followed for intramural hematoma, intimal flap or double lumen in a 3.0 Tesla MRI scanner (Discovery MR750, GE Healthcare).

Whole-Exome Sequencing (WES)

Genomic DNA was extracted from peripheral blood with a Gentra Puregene Blood Kit (QIAGEN, Hilden, Germany) according to the manufacturer's instructions. Thirty-eight samples were enriched by NimbleGen SaqCap v3 (Roche NimbleGen, Inc., California, USA) and processed by the Genome Research Center of National Yang Ming Chiao Tung University, while the other 114 samples were enriched by SureSelect V6 (Agilent Technologies, California, USA) and processed by the Genomics Biotechnology Co., Ltd. All sequencing was performed on the Illumina NovaSeq 6000 platform.

Bioinformatic Pipelines of Variant Annotation

The quality of raw reads was checked by FastQC (v0.11.4) (Andrews 2010). We used BWA (v0.7.12) (Li and Durbin 2009) and Samtools (v1.9) (Li et al. 2009) for alignment to the GRCh38 reference genome by default parameters. Marking duplications, base quality score recalibration, and variant calling were performed using GATK (v4.1.4.1) (McKenna et al. 2010; Van der Auwera et al. 2013). Variants with read

depths less than 30 were excluded from data interpretation. After variant calling, multiple reference databases, including ClinVar (retrieved 2021-07), dbNSFP (v4.1a), dbSNP (v154), and Online Mendelian Inheritance in Man (OMIM, retrieved 2021-07), were applied for further annotation (McLaren et al. 2016).

Variant Filtration and Prioritization of Candidate Causal Variants

We first searched for stroke-related pathogenic and likely pathogenic variants (PVs) in the ClinVar database (<https://www.ncbi.nlm.nih.gov/clinvar/>). The stroke-related phenotypes were determined according to the OMIM database (<https://www.omim.org/>). If no reported or phenotype-compatible PV was found, we then examined novel rare variants which were predicted to cause structural and functional impacts on the coding products in 325 stroke-related genes (Supplementary Table S1) by searching the terms of stroke, ischemic stroke, hemorrhagic stroke, stroke-like episode, cerebral hemorrhage, lacunar stroke, and thromboembolic stroke in OMIM and PubMed (Ilinca et al. 2019; Hamosh et al. 2005; Dichgans et al. 2019). The selected variants must have met all of the following inclusion criteria: (1) the global MAF $\leq 0.1\%$ in public genome database; (2) not reported as a benign or likely benign variants in ClinVar or SNP database; (3) relatively evolutionarily conserved (GREP + + > 3); (4) computational predictions assessed by at least two scoring systems (SIFT, Polyphen2, MutationTaster, computer-aided design and drafting score > 20) or Ensembl rare exome variant ensemble learner score ≥ 0.6 (Ioannidis et al. 2016); (5) considered phenotype-genotype correlation. If the selected variant fulfilled pathogenic or likely pathogenic according to the ACMG guideline (PVs, judged with InterVar's assistance) (Kalia et al. 2017; Li and Wang 2017), Sanger sequencing validation and intrafamilial segregation were performed in the proband and available family members.

Statistical Analysis

The statistical analysis was conducted using SPSS (Version 20.0 for Windows, Armonk, NY: IBM Corp. USA). To compare the characteristics between probands with and without monogenic PVs, two-sample independent t-tests were used for age at onset, initial severity of National Institute of Health Stroke Scale (NIHSS) scores, modified Rankin Scale (mRS) scores at baseline and three months after stroke, and risk factors. Risk factor numbers (0–7) were derived by summing the risk factors present in a stroke patient, including hypertension, diabetes, hyperlipidemia, atrial fibrillation, cigarette smoking, excessive alcohol consumption, and obesity (body mass index ≥ 30). Categorical variables, including

sex, stroke subtypes, and obesity, were compared by the chi-square test or Fisher's exact test. The hit rate of monogenic stroke based on stroke subtyping was defined as the patient number with monogenic causes divided by the total patient number of that specific subtype. Statistical significance was set at $p < 0.05$.

Results

Patient Characteristics of the Familial Stroke Caused by Monogenic Etiologies

Nine out of 161 patients (9/161 = 5.6%) having characteristic phenotypes and brain MRI/MRA were diagnosed via conventional Sanger sequencing of specific gene hotspots, including six with *NOTCH3*-related CADASIL (all carried p.R544C), two with *GLA*-related Fabry disease (p.Y365* and IVS4+919G > A), and one with m.3243A > G-related MELAS. For the remaining 152 probands who had a family history but the candidate genes were unclear (152/161 = 94.4%), WES revealed pathogenic/likely pathogenic variants (PVs) in 24 individuals (24/152 = 15.8%). The overall yield rate to detect a PV for relevant phenotype by the combination of WES and conventional sequencing was high, 20.5% (33/161). The clinical diseases and the genotypes of the 33 probands with ascertained monogenic stroke are summarized in Table 1. The detailed phenotypic and genotypic spectrum was described as follows. Another 23 family members ascertained with monogenic strokes shared the corresponding phenotypes and genotypes with individual probands (Supplementary Fig. S1-S2).

The demographic characteristics of 161 probands are shown in Table 2. The mean age at onset was 53.2 ± 13.7 years old and 63.4% were male. The most common ischemic stroke subtype was SVD (29.8%) and LAA (27.3%). ICH accounted for 11.3%. The stroke subtype distribution was close to previous publications for Asians (Kim and Kim 2014). Notably, the subtype distribution was different between the patients with and without ascertained monogenic stroke. The patients ascertained with monogenic stroke ($n = 33$) appeared to have more SVD (45.5% vs 25.8%, $p = 0.046$) and ICH (21.2% vs 8.6%, $p = 0.04$) as compared to those unassigned probands. In contrast, LAA and cardioembolism were the less common subtypes in the monogenic group. Additionally, the patients with monogenic stroke had fewer vascular risk factors (the proportion with 0–1 risk factor: 42.4% vs 19.5%, $p = 0.01$). Apart from stroke subtypes and risk factors, we did not find significant differences in the other clinical features between the two groups, including age at onset, sex, the number of strokes affected members per family cluster, the initial severity of National Institute of

Health Stroke Scale (NIHSS) and the modified Rankin Scale (mRS) outcome at three months after stroke.

Phenotypic and Genotypic Spectrum and the Hit Rate of Monogenic Causes

Among the 33 probands with ascertained monogenic stroke (Table 1), CADASIL was the most common hereditary stroke disease ($n = 16$, 48.5%). The other disorders included pseudoxanthoma elasticum ($n = 2$), coagulation factor II disorder ($n = 2$), beta-thalassemia ($n = 2$), Fabry disease ($n = 2$), MELAS ($n = 1$ for each disease below), protein S deficiency, hereditary hemorrhagic telangiectasia (HHT), polycystic kidney disease (PKD), CADASIL2, moyamoya disease (MD), cerebral cavernous malformation (CCM), familial hyperlipidemia, and primary thrombocythemia.

The hit rates (HR) of detecting causative genes for a specific subtype are shown in Fig. 2. The greatest diagnostic yield was in ICH patients (7/18, 38.9%). Specifically, all four patients with structural vasculopathy subtype had a monogenic etiology (4/4, 100%): *KRIT1* for CCM, *RNF213* for MD, *ENG* for HHT, and *PKD1* for PKD. For ischemic stroke (IS), the SVD subtype had the highest HR (15/48, 31.3%), followed by other determined causes (4/21, 19%), whereas the LAA (4/44, 9.1%) and the cardioembolic (0%) subtypes had the lowest HR. *NOTCH3* mutations presented with various subtypes, including SVD ($n = 13$), hypertensive ICH ($n = 2$), and LAA ($n = 1$).

Intracerebral Hemorrhage (ICH)

The representative monogenic cases identified using WES are shown in Fig. 3. We disclosed a novel stop-gain variant, *KRIT1* p.E379* in a 63-year-old man with multiple CCMs presented as recurrent ICH and cavernoma-related epilepsy (Fig. 3a). His epilepsy was firstly diagnosed at 27 years old and poorly controlled by medication since then. After his first ICH event in 2019, multiple CCMs had been found by MRI and the largest (4.8 cm) cavernoma at the left temporal region was removed by craniotomy. The histopathology showed aggregated dilated vascular space with thinning wall and thrombus formation, which was compatible with cavernoma. His younger brother, who carried the same variant, also had recurrent ICH due to multiple CCMs and microbleeds in the bilateral cerebral and cerebellar hemispheres. The segregation analysis of *KRIT1* p.E379* was consistent with the disease and only in two affected family members. This novel variant is absent in global population database gnomAD and Taiwan Biobank database. The stop-gain variant is presumed to cause loss of function of the gene at vascular endothelial tight junctions (Merello et al. 2016). It is predicted to be deleterious by multiple lines of bioinformatic evidence (CADD phred score 40, GERP+ + NS 5.66). Therefore, this variant

Table 1 Genotypes and phenotypes of the 33 unrelated probands diagnosed as monogenic familial stroke, including two carrying novel variants^a

Diseases	Stroke sub-type	Gene	Variants	Inheritance	ClinVar	ACMG (Classification)	SNP ID
Ischemic Stroke							
CADASIL	SVD (<i>n</i> = 12, 1 Hom), LAA (<i>n</i> = 1), SVD (<i>n</i> = 1)	<i>NOTCH3</i>	p.R544C	AD	Pathogenic	Pathogenic	rs201118034
		<i>NOTCH3</i>	p.R1231C	AD	Conflicting	Likely pathogenic (PM1 + PP2 + PP3 + PP4 + PP5)	rs201680145
CADASIL2	SVD (<i>n</i> = 1)	<i>HTRA1</i>	p.G276A	AD	Likely pathogenic	Likely pathogenic	rs1554952277
Fabry disease	Other (<i>n</i> = 2)	<i>GLA</i>	p.Y365*	XLR	Pathogenic	Pathogenic	rs104894849
		<i>GLA</i>	IVS4+919G>A	XLR	Pathogenic	Pathogenic	rs199473684
Familial hyperlipidemia	LAA (<i>n</i> = 1)	<i>LDLR</i>	p.D90N	AD	Conflicting	Likely pathogenic (PM1 + PM5 + PP2 + PP3)	rs749038326
MELAS	Other (<i>n</i> = 1)	<i>mtDNA</i>	m.3243A>G	Maternal	Pathogenic	Pathogenic	rs199474657
Protein S deficiency	LAA (<i>n</i> = 1)	<i>PROS1</i>	p.Y592*	AD	Pathogenic	Pathogenic	rs199469503
Myeloproliferative neoplasia	Undetermined (<i>n</i> = 1)	<i>JAK2</i>	p.V617F	AD	Pathogenic	Pathogenic	rs77375493
Pseudoxanthoma elasticum	LAA (<i>n</i> = 1, Het)	<i>ABCC6</i>	p.R1235W	AD	Likely pathogenic	Likely pathogenic	rs63750402
	Other (<i>n</i> = 1, Hom)	<i>ABCC6</i>	p.E709K	AR	Variation of uncertain significance	Likely pathogenic (PM1 + PM2 + PP3 + PP4)	rs114303883
Thalassemia	SVD (<i>n</i> = 1)	<i>HBB</i>	p.F42fs	AD	Pathogenic	Pathogenic	rs80356821
Coagulation factor II disorder	Undetermined (<i>n</i> = 1)	<i>F2</i>	p.R596Q	AD	Pathogenic	Pathogenic	rs387907201
Intracerebral Hemorrhage							
CADASIL	HTN (<i>n</i> = 2)	<i>NOTCH3</i>	p.R544C	AD	Pathogenic	Pathogenic	rs201118034
Cerebral cavernous malformation ^a	SV (<i>n</i> = 1)	<i>KRITI</i>	p.E379*	AD	Not reported	Pathogenic (PVS1 + PM2 + PP3 + PP4)	Novel
Hereditary hemorrhagic telangiectasia	SV (<i>n</i> = 1)	<i>ENG</i>	p.Y258Lfs*76	AD	Pathogenic	Pathogenic	rs1588581902
Moyamoya disease	SV (<i>n</i> = 1)	<i>RNF213</i>	p.R4810K	AD	Likely pathogenic	Likely pathogenic	rs112735431
PKD	SV (<i>n</i> = 1)	<i>PKD1</i>	c.11014-10C>A	AD	Pathogenic	Pathogenic	rs555703777
Thalassemia	HTN (<i>n</i> = 1)	<i>HBB</i>	p.K18*	AD	Pathogenic	Pathogenic	rs33986703
Cerebral venous sinus thrombosis							
Coagulation factor II disorder ^a	Hemorrhagic infarction (<i>n</i> = 1)	<i>F2</i>	p.F382L	AD	Not reported	Likely pathogenic (PM1 + PM2 + PP1 + PP2 + PP4)	Novel

ACMG American College of Medical Genetics, AD autosomal dominant, AR autosomal recessive, CADASIL Cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy, ClinVar <https://www.ncbi.nlm.nih.gov/clinvar/>, Het Heterozygosity, Hom Homozygosity, HTN hypertensive subtype, ICH Intracerebral hemorrhage, LAA Large artery atherosclerosis, MELAS Mitochondrial encephalopathy, lactic acidosis, and stroke-like episodes, PKD Polycystic kidney disease, PM Pathogenic moderate, PS Pathogenic supporting, PVS Pathogenic very strong, SV Structural vasculopathy, SVD Small vessel disease, VUS Variants of uncertain significance, XLR x-link recessive

^aNovel variant identified

is considered pathogenic according to ACMG guideline criteria (PVS1 + PM2 + PP3 + PP4, see Table 1). Based on the genetic diagnosis, we provided genetic counseling, close

follow-up of brain MRI, and strict blood pressure control for the patients and affected family members.

Table 2 Demographics of ascertained monogenic versus unassigned familial stroke patients

Clinical features	Total	Monogenic stroke	Unassigned	p-value
Patient number	161	33 (20.5%)	128 (79.5%)	N/A
Age at onset (y)	53.2 ± 13.7	51.5 ± 14.5	53.7 ± 13.5	0.43
Age ≤ 55, n (%)	94 (58.4)	19 (57.6)	75 (58.6)	0.91
Sex, Men (%)	63.4	60.6	63.6	0.66
Stroke affected members per family cluster	3.1 ± 1.4	3.0 ± 1.0	3.1 ± 1.5	0.5
Infarction subtype (%)	142 (88.2%)	25 (75.8%)	117(90.7%)	0.08
Small vessel disease	48 (29.8%)	15 (45.5%)	33 (25.8%)	0.046*
Large artery atherosclerosis	44 (27.3%)	4 (12.1%)	40 (31.2%)	0.047*
Undetermined	18 (11.2%)	2 (6.1%)	16 (12.5%)	0.54
Other determined	21 (13%)	4 (12.1%)	17 (13.3%)	0.99
Cardiac embolism	11 (6.8%)	0 (0%)	11 (8.6%)	N/A
Intracerebral hemorrhage	18 (11.3%)	7 (21.2%)	11 (8.6%)	0.04*
Cerebral venous thrombosis	1 (0.6%)	1 (3%)	0 (0%)	N/A
Risk factor numbers (0–7)	2.2 ± 1.2	1.8 ± 1.4	2.4 ± 1.1	0.03*
Risk factors ≤ 1	39 (24%)	14 (42.4%)	25 (19.5%)	0.01*
Initial severity NIHSS	6.2 ± 7.4	7.3 ± 8.3	6 ± 7.2	0.43
3-month favorable outcome	108 (67.1%)	24 (72.7%)	84 (65.6%)	0.57

All continuous variables were presented as mean ± standard deviation

NIHSS National Institute of Health Stroke Scale; mRS Modified Rankin Scale at 3 months. Favorable outcome as mRS 0–2. Risk factors included hypertension, diabetes mellitus, hyperlipidemia, atrial fibrillation, cigarette smoking, excessive alcohol consumption, and obesity (body mass index ≥ 30).

*p < 0.05 considered significant for monogenic stroke vs. non-monogenic case

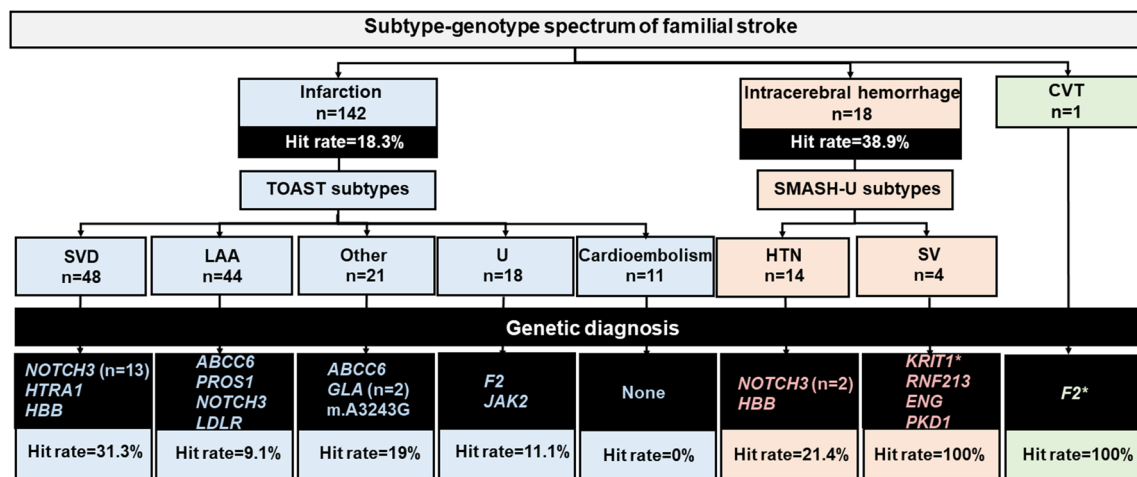


Fig. 2 Phenotypic and genotypic spectrum and the hit rate across stroke subtypes. CVT Cerebral venous sinus thrombosis, HTN hypertensive subtype, LAA large artery atherosclerosis, SMASH-U Structural lesion, Medication, Amyloid angiopathy, Systemic/other dis-

ease, Hypertension, Undetermined, SV structural vasculopathy, SVD small vessel disease, TOAST Trial of Org 10172 in Acute Stroke Treatment. *Novel variant

A 35-year-old woman with no known risk factor presented by acute intraventricular hemorrhage and obstructive hydrocephalus, and the brain MRA documented Suzuki stage IV with occlusive changes in the bilateral distal internal carotid arteries and tenuous anterior and middle cerebral arteries (Fig. 3b). Her mother expired due

to a hemorrhagic stroke in her late 40 s. She was identified to carry a heterozygous *RNF213* p.R4810K variant, a hotspot variant of Moyamoya disease (MD) in the East Asia population (Lee et al. 2015; Ishigami et al. 2022). We provided education about the condition and transferred to neurosurgery department for extracranial-intracranial

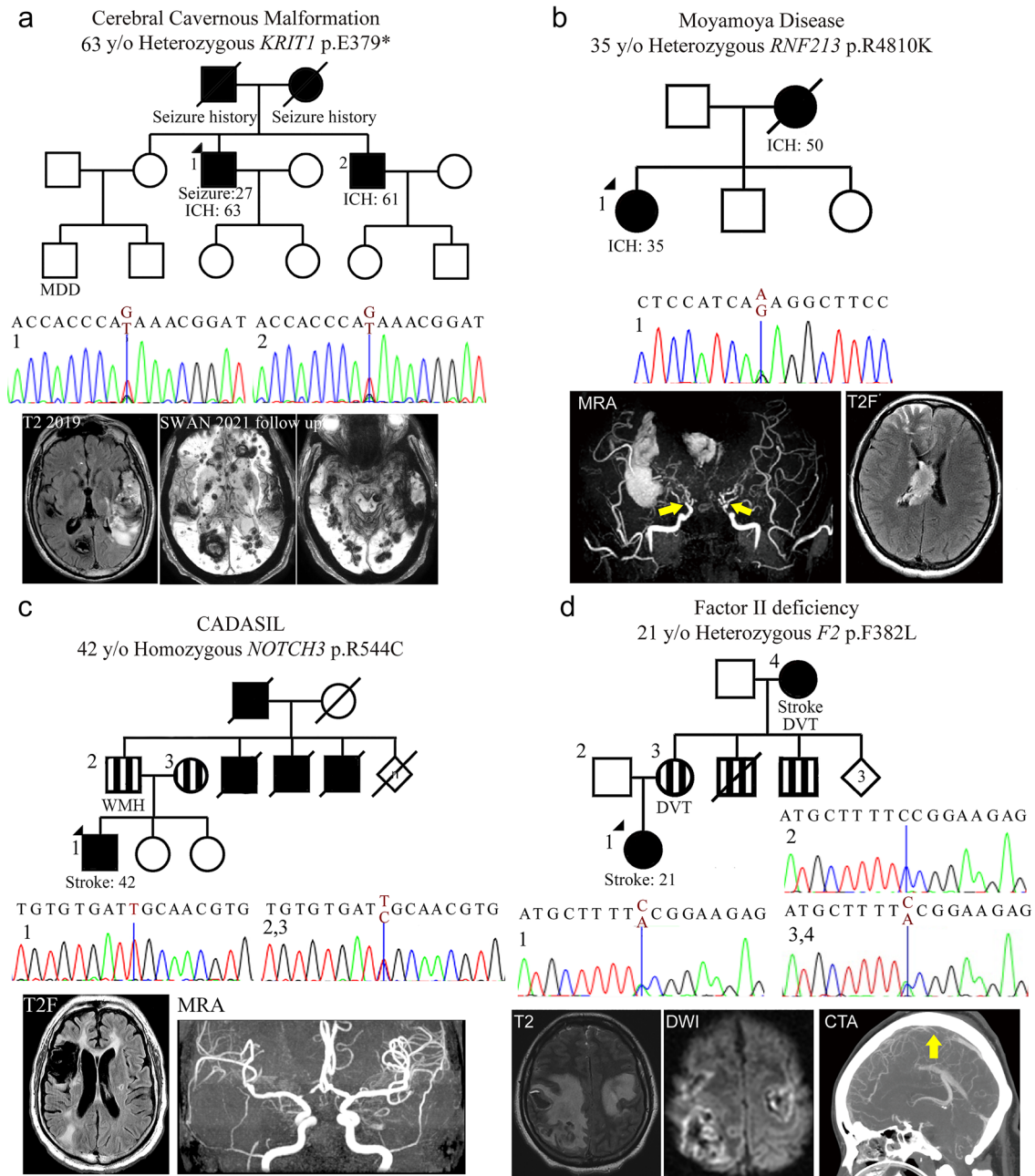


Fig. 3 Representative pedigrees of monogenic familial strokes. **a** A 63-year-old man with numerous cerebral cavernous malformations carried a novel stop-gain variant *KRIT1* p.E379*, which was segregated in terms of genotype and phenotype with his younger brother. **b** A 35-year-old woman with intraventricular hemorrhage due to moyamoya disease carried a heterozygous *RNF213* p.R4810K variant. **c** A 42-year-old man with cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy (CADASIL) with recurrent infarctions and severe white matter hyperintensities carried

homozygous *NOTCH3* p.R544C variants, which was segregated with his parents. **d** A 21-year-old woman with superior sagittal venous sinus thrombosis and bilateral hemorrhagic infarctions carried a novel likely pathogenic variant *F2* p.F382L, which was segregated within her family. *AO* age of onset, *CTA* Computed tomography angiography, *DVT* deep vein thrombosis, *DWI* diffuse weighted image, *MRA* magnetic resonance angiography, *SWAN* Susceptibility weighted angiography, *WMH* white matter hyperintensity. Yellow arrowheads indicate the occlusive vessel

arterial bypass according to treatment recommendations (Fujimura et al. 2022).

WES detected an intronic pathogenic variant *PKD1* c.11014-10C>A in a 38-year-old man who suffered from

subarachnoid hemorrhage by a ruptured anterior communicating artery aneurysm and PKD. His father had PKD and died from a ruptured brain aneurysm. All his three siblings had been diagnosed as PKD and shared the same *PKD1*

variant. Computed tomography (CT) angiography was suggested for these siblings to detect asymptomatic brain aneurysms. Also, a 23-year-old woman who had arteriovenous fistula presented by ICH and seizure had been identified *ENG* p.Y258Lfs*76. The proband and both of her two siblings had spontaneous and recurrent nosebleeds since childhood, which was compatible with hereditary hemorrhagic telangiectasia (HHT) phenotype. Education was also provided for these two probands.

Another two probands carried the CADASIL hotspot PV *NOTCH3* p.R544C and suffered from dementia prior to the ICH event. *HBB* p.K18* variant was detected in a 63-year-old man who suffered from hypertensive thalamic ICH and thalassemia (low hemoglobin).

Ischemic Small Vessel Disease (SVD)

In this subtype, twelve probands were identified as carrying the hotspot PV *NOTCH3* p.R544C and one as carrying the *NOTCH3* p.R1321C. A representative proband, carrying the homozygous *NOTCH3* p.R544C variants (Fig. 3c), had his first stroke at 42 years old and later had early-onset dementia (Mini-Mental State Examination = 21) at the age of 46. His brain MRI/MRA revealed large tissue loss in the right anterior frontal region, bilateral severe white matter hyperintensities and multisegmental mild narrowing in the M1 segment of right middle cerebral artery. Also, he had post-stroke seizure under Frisium and Depakine. Both of his parents were confirmed carrying heterozygous *NOTCH3* p.R544C with moderate white matter hyperintensities. Regular MRI follow-up for brain microbleeds and careful antithrombotic adjustment were suggested. The other pedigrees of *NOTCH3*-related stroke were presented in Supplementary Fig. S1. Another 51-year-old woman with CADASIL syndrome of stroke and mild dementia carried heterozygous *HTRA1* p.G276A, which was reported to be likely pathogenic to cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy type 2 (CADASIL2) (Lee et al. 2018). Furthermore, a 52-year-old woman who carried *HBB* p.F42fs variant and who suffered from a lacunar stroke in the motor area with thalassemia.

Cerebral Venous Sinus Thrombosis (CVT)

We discovered a novel variant *F2* p.F382L in a 21-year-old woman suffering from acute consciousness change and generalized tonic–clonic seizures due to cerebral venous sinus thrombosis and hemorrhagic infarction at the bilateral frontoparietal cortices (Fig. 3d). Her mother and maternal grandmother, who also carried *F2* p.F382L, both suffered from recurrent episodes of deep vein thrombosis and cerebral venous thrombosis. One maternal uncle died from pulmonary embolism at his 30 s and another suffer from stroke

and deep vein thrombosis at his 50 s. Decreased coagulation factor II activities were observed in the proband, her mother, and grandmother (44.7, 67.1, and 68.5%, respectively; reference: 79–131%). The segregation analysis was consistent with the disease in the aforementioned three affected family members and one unaffected family member. This novel variant is absent in global population database gnomAD and Taiwan Biobank database. However, it is unknown how this mutation leads to low factor II activities and prothrombotic and prohemorrhagic tendency. The variant location is highly evolutionary conserved (GERP++NS = 5.23, phast-Cons = 1) for encoding factor IIa thrombin and fibrin interaction exosite I domain (Huntington 2005). Missense variants in the *F2* gene have very low rate as benign variants (ClinVar *F2* missense variants: 21/38 pathogenic or likely pathogenic, while 2/38 benign or likely benign). Taken together, this variant is considered likely pathogenic according to ACMG guideline criteria (PM1 + PM2 + PP1 + PP2 + PP4, see Table 1). The establishment of genetic etiology introduced timely non-vitamin-K oral anticoagulant therapy with a specific factor X inhibitor (Yaghi et al. 2022), which successfully prevented the recurrence of thromboembolic events and achieved good functional outcomes in this family for years of followed-ups till now.

Other Determined Diseases and Undetermined Ischemic Stroke

We found a 43-year-old man proband carrying a rare homozygous *ABCC6* p.E709K variant had infarcts at the bilateral corona radiata, bilateral basal ganglia, and pons because of vertebrobasilar artery dissection, which was phenotypic compatible with *ABCC6* related pseudoxanthoma elasticum. Pseudoxanthoma elasticum is an inherited elastic fiber disorder usually caused by *ABCC6* variants in an autosomal recessive inheritance pattern, but heterozygous carriers may express partial phenotypes (Miksch et al. 2005; Plomp et al. 2004). Besides, a rare pathogenic variant, *F2* p.R596Q, known for thrombophilia and antithrombin resistance (Takagi et al. 2016), was identified in a 41-year-old man presented by embolic stroke in the right parietal cortex. The proband's factor II activity remained mildly low during follow-ups. Lastly, a well-known pathogenic variant *JAK2* p.V617F for the myeloproliferative disorder was identified in a 72-year-old woman with thrombocytopenia and an undetermined subtype of stroke.

Large Artery Atherosclerosis (LAA)

We found a 59-year-old man proband who had basilar artery atherosclerotic infarction as well as peripheral artery occlusive disease, carrying the heterozygous *ABCC6* p.R1235W variant, which is known to be pathogenic for

pseudoxanthoma elasticum. Also, *LDLR* p.D90N, a PV to familial hyperlipidemia, was detected in a 55-year-old man suffered from recurrent stroke with the third event of an acute vertebral artery occlusion infarction over the bilateral cerebellum and left medulla oblongata. Furthermore, WES identified *PROS1* p.Y592* mutation in a 72-year-old man who had protein S deficiency and episodes of large artery atherosclerosis-related infarction and deep vein thrombosis. The same variant was also detected in his son who suffered from deep vein thrombosis and decreased protein S activity as well. Lastly, a 62-year-old woman suffered from right posterior cerebral artery in-situ thrombosis infarction was identified carrying the hotspot pathogenic variant *NOTCH3* p.R544C.

Stroke-Associated Variants of Uncertain Significance (VUS)

Apart from pathogenic or likely pathogenic findings, 41 rare VUS (Table 3) in 325 stroke-related candidate genes (Supplementary Table S1) were detected in 39 probands, including eight novel variants: *ABCA1* p.H1223P, *ACTA2* p.N298S, *COL5A1* p.P930R, *FBN1* p.Q514E, *FBN1* p.K1052T, *GP1BA* p.P437*, *RNF213* p.N1631Efs*34, and *TSC1* p.P397del. Although the evidence of pathogenicity is not enough due to the unavailability of familial segregation or the paucity of their functional impacts currently, the rare allele frequency in our population, in silico prediction and the corresponded clinical manifestations indicate that these variants may still be potentially associated with the pathogenesis of stroke. We would like to report them here and expect more evidence to elucidate their clinical significance in future. Finally, we identified no findings after variant filtering in 42 probands and the other 47 had incompatible variants with phenotype or inheritance mode, they were categorized as unidentified ($n = 89$).

Discussion

We proposed a stroke subtype-guided WES in combination with prevalent gene/hotspot screening and achieved an overall diagnostic yield 20.5% of monogenic causes in 33 out of 161 unrelated probands with well-defined familial stroke. This study included all IS and ICH subtypes, and thus outlined a more comprehensive genomic landscape of familial stroke. The identified PVs were distributed in a wide range of genes and rare diseases. For example, CADASIL syndrome could be caused by *NOTCH3* or *HTRA1* mutations, both account for a significant number of SVD and/or ICH patients in our population (Lee et al. 2018). Notably, we identified two novel PVs, *KRITI* p.E379* in a family of cerebral cavernous malformation-related ICH, and *F2*

p.F382L in a family of cerebral venous thrombosis-related stroke, respectively. Our findings demonstrated the power and necessity of WES in genetic diagnosis of monogenic stroke, a disease entity with high heterogeneity and complex etiologies.

Compared with the previous next-generation sequencing (NGS) studies in stroke patients, our diagnostic yield was relatively high as summarized in Table 4, although these study designs, methods and key findings were substantially different. Overall, most NGS studies were limited by small sizes, restricted screened genes, or narrow stroke subtypes (Tan et al. 2019; Lee et al. 2019; Bersano et al. 2016; Ilinca et al. 2020; Mishra et al. 2019; Auer et al. 2015). According to our results, the subtypes of familial ICH and SVD stroke, determined by the clinical features and characteristic brain MRI features, can provide the most informative clue to possible underlying monogenic etiologies. The highest HR was in the patients with the ICH (38.9%), especially for structural vasculopathy (100%), followed by SVD (31.3%). In parallel, when a patient with familial stroke has very few (0–1) conventional vascular risk factors, an underlying monogenic etiology should be aggressively searched. Otherwise, there was no difference in age at onset, sex, family history, the initial stroke severity or stroke outcome between the patients with and without a definite genetic diagnosis.

NOTCH3 is the most prevalent disease-causing gene in our cohort of familial stroke. The existence of certain *NOTCH3* variants is also an important risk factor for sporadic stroke in the Taiwan population (Lee et al. 2019). Because *NOTCH3* p.R544C is the leading hotspot in both familial stroke (15/161, 9.3%, in this study) and sporadic stroke patients in Taiwan (Wei et al. 2021), suspicious patients with CADASIL syndrome could be screened for *NOTCH3* p.R544C by Sanger sequencing. However, if the hotspot screening turns negative, it would be tedious and time-consuming to screen the whole *NOTCH3* gene by conventional sequencing, or search for other monogenic cerebrovascular diseases. In this study, we adopted Taiwan Biobank (TWB) 2.0 SNP Array in the genetic diagnosis pipeline, which includes *NOTCH3* p.R544C and other pathogenic intronic and mitochondrial variants not covered by WES, like m.3243 A > G for MELAS (Wei et al. 2021). Our integrative NGS approach combining WES and TWB 2.0 Array may be an efficient alternative to conventional sequencing for non-CADASIL stroke patients. The establishment of genetic diagnosis is important for personalized antiplatelet agents (Chung et al. 2020) and early risk factor modification for the carriers of these disease-associated variants (Lee et al. 2019).

Genetic diagnosis of familial CCM is an emerging issue. Familial CCM, with the hallmark of multiple ICH of structural vasculopathy subtype, has significantly higher risks of rebleeding and epilepsy than the sporadic CCM cases.

Table 3 Variants of uncertain significance in 325 stroke-associated genes detected in this study

Gene	Disease	Variants	SNP ID	ACMG
<i>ABCA1</i>	Familial HDL deficiency	c.A3668C: p.H1223P	Novel	PM2
		c.G5140A: p.A1714T	rs758640488	PM2, PP3
<i>ACTA2</i>	Moyamoya disease	c.A893G:p.N298S	Novel	PM2,PP2,PP3,PP4
<i>ACVRL1</i>	Hereditary Hemorrhagic Telangiectasia	c.G1444A:p.A482T	rs777374619	PM1,PM2,PP3
<i>ANGPTL6</i>	Familial intracranial aneurysms	c.1003dup: p.Y335Lfs*10	rs754920128	PM2
<i>APP</i>	Cerebral amyloid angiopathy	c.A143G:p.Q48R	rs1387322785	PM1, PM2, PP3
<i>COL1A1</i>	Ehlers-Danlos syndrome	c.G344A:p.G115E	rs767924851	PP2, PP3, BP1
<i>COL4A2</i>	PADMAL	c.G2954T:p.G985V	rs769851485	PP3
		c.C4514T:p.P1505L	rs777086508	PM1,PM2,PP3,PP4
<i>COL5A1</i>	Ehlers-Danlos syndrome	c.C2789G:p.P930R	Novel	PM2, PP3
<i>COL5A2</i>		c.C382T:p.R128C	rs1247714244	PM2, PP3
		c.C3038T:p.A1013V	rs372220538	PP3
<i>F2</i>	Coagulation factors deficiency	c.C1366T:p.R456W	rs139552841	PM1,PM2,PP2
<i>F5</i>		c.C4949T:p.A1650V	rs753691316	PM1,PM2,PP3
<i>F11</i>		c.C1107A: p.Y369*	rs773905328	PVS1, PP4, PP5
<i>FBN1</i>	Marfan syndrome	c.G763A:p.G255R	rs781113082	PM1,PM2,PP3,PP4
		c.C1540G:p.Q514E	Novel	PM2
		c.A3155C:p.K1052T	Novel	PM1, PM2
<i>FLNC</i>	Cardiomyopathy	c.C1205T:p.T402I	rs374757755	PM2, PP2, PP3
<i>GPIBA</i>	Bernard-Soulier syndrome	c.1275_1278del: p.T427Qfs*44	rs763460861	PVS1, PM2
		c.A1311*: p.P437*	Novel	PM2
<i>JAK2</i>	Chronic myeloproliferative diseases	c.C1141T:p.H381Y	rs774526479	PM1,PP3,PP4
<i>KCNA5</i>	Familial atrial fibrillation	c.213_245dup: p.D72_P82dup	rs144879674	PM4, PP4
<i>LDLR</i>	Familial hypercholesterolemia,	c.G409A: p.G 137S	rs730882082	PM2, PP3, PP5
<i>MYBPC3</i>	Cardiomyopathy	c.G104A: p.R35Q	rs397515885	PM2, PP2, BP1
<i>MYH6</i>	Cardiomyopathy	c.G5696A:p.R1899H	rs61731171	PP2, PP3
<i>MYLK</i>	Aortic aneurysm / arterial dissection	c.G3379C:E1127Q	rs777401981	PM2, PP2, PP3
		c.G4528A: p.D1510N	rs1483753716	PM2, PP2, PP3
<i>NBEAL2</i>	Gray platelet syndrome (from same proband)	c.G211A: p.A71T	Novel	PM2, PP3, PP4
		c.C7351T: p.R2451W	rs748364740	
		c.C7486T: p.R2496W	rs754133402	
<i>NOTCH3</i>	CADASIL	c.G391C:p.G131R	rs767150916	PM1, PM2, PP3
<i>PKD1</i>	Polycystic kidney disease	c.C6070T: p.R2024C	rs199943712	PP3, BP1
<i>RNF213</i>	Moyamoya disease	c.4890dup:p.N1631Efs*34	Novel	PM2
		c.T11668C:p.S3890P	rs1203492091	PM2, BP4
		c.G12913T: p.V4305L	rs1399368898	PM2
		c.C13726T:p.P4576S	rs776390324	PP3
<i>SCN2B</i>	Familial atrial fibrillation	c.G71-1A: p.?	rs758825110	PP3
<i>TGFBR1</i>	Loeys-Dietz syndrome	c.A491G:p.N164S	rs767785290	PM2, PP4
<i>TREX1</i>	RVCL	c.294dup:p.C99Mfs*3	rs760594164	PM1
<i>TSC1</i>	Tuberous sclerosis	c.1188_1190del: p.P397del	Novel	PM2, PM4, PP4, BP3

ACMG American College of Medical Genetics, *CADASIL* cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy, *HDL* High-density lipoprotein, *OMIM* Mendelian Inheritance in Man, *PADMAL* pontine autosomal dominant microangiopathy and leukoencephalopathy, *PM* Moderate evidence of pathogenicity, *PP* Supporting evidence of pathogenicity, *PVS* Very strong evidence of pathogenicity, *RVCL* Retinal Vasculopathy with Cerebral Leukoencephalopathy

The Angioma Alliance Scientific Advisory Board Clinical Experts Panel has recommended that diagnostic genetic testing should be performed when there is a positive family history or multiple CCMs (Akers et al. 2017). The three

responsible genes, *CCM1* (*KRIT1*), *CCM2* (*MGC4607*) and *CCM3* (*PDCD10*), are associated with variable risk of rebleeding or refractory epilepsy (Fischer et al. 2013; Yadla et al. 2010; Riant et al. 2010). *CCM3* variations have been

Table 4 Comparison of the diagnostic yield of monogenic stroke diseases among previously relevant genetic studies and our study*

Study	Study design and inclusion criteria (n)	Nation	Sequencing method	Diagnostic yield of monogenic stroke diseases
Chang et al.*	Familial stroke including ischemic and intracerebral hemorrhage, onset 18–79 y/o (n = 161)	Taiwan	WES (VUS in 325 genes) + Targeted array	Overall 20.5%; ICH: 35.3%; SVD: 31.3%; O: 22.7%; U: 11.1%, LAA: 9.1%
Ilinca et al 2020	Familial stroke, onset ≤ 55 y/o (n = 22)	Sweden, Finland	WES (VUS in 254 genes)	Overall 27.3%; ICH: 1/1; CE: 1/2; SVD: 3/11; U: 1/4
Mishra et al 2019	Case–control design, cohort of noninstitutionalized stroke ≤ 80 y/o (n = 1469)	France	WES (VUS in 5 genes)	SVD: 0.8%
Tan et al 2019	MRI confirmed small vessel disease stroke, onset ≤ 70 y/o (n = 950)	U.K	Targeted array (15 genes)	1.5%
Lee et al 2019	Consecutive ischemic stroke patients (n = 800) versus 7038 healthy controls	Taiwan	Single-gene sequencing of <i>NOTCH3</i>	6.5% in SVD; 0.9% in healthy controls
Bersano et al 2016	Suspected monogenic diseases, fulfilled at least one of the following criteria: 1. Unknown cause, < 3 conventional vascular risk factors 2. Stroke onset ≤ 55 y/o 3. Positive familial history, 4. ≥ 2 associated clinical features (n = 227)	Italy	Sanger sequencing of the hotspots for 5 diseases (CADASIL, Fabry disease, MELAS, CAA, Marfan syndrome)	Overall 7% CADASIL: 9/103 Fabry: 1/33 CAA: 1/70 MELAS: 2/16 Marfan: 1/5

CADASIL Cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy, *CAA* Cerebral amyloid angiopathy, *ICH* intracerebral hemorrhage, *LAA* Large artery atherosclerosis, *MELAS* Mitochondrial encephalopathy, lactic acidosis, and stroke-like episodes, *NHLBI ESP* The National Heart, Lung, and Blood Institute, Exome Sequencing Project, *SVD* Small vessel diseases, *WES* Whole-exome sequencing

reported to be associated with a higher risk of early-onset ICH and recurrent bleeding (Choquet et al. 2015; Riant et al. 2013; Shenkar et al. 2015). Therefore, the genetic etiology is essential for prediction of the risk and severity of CCM-related epilepsy or ICH and interconnected to decision of genetic testing and individualized treatment plans.

There are limitations of this study. It should be very cautious in determining the clinical significance of the identified rare variants. First, it is particularly challenging when facing the novel variants or VUS with limited clinical evidence. Accumulating genomic data in large population may help to clarify the significance of some undetermined variants (VUS in 39 out of 161, 24.2%) in future. Second, the function of the novel variant *F2 p.F382L* is unknown. Further studies of the impact on thrombin-fibrin interaction and thrombin antithrombin complex (Miyawaki et al. 2012) are warranted to elucidate the pathophysiology. Third, the scope of this study primarily focused on known genes related to stroke. Given that trio study was unavailable for most of our mid-age-onset and late-onset stroke probands, there are probable novel causative genes concealed in the unidentified cases (89 out of 161, 55.3%). Fourth, severe patients could not sign the

consent, e.g., unconscious or unstable vital signs, thus there was selection bias by the severity of stroke. Lastly, from our small, single center cohort, LAA and cardioembolic stroke subtypes took much smaller proportions in the patients with a definite genetic diagnosis, suggesting that these subtypes are more likely secondary to acquired and aging factors rather than monogenic causes. Further studies are warranted.

In summary, monogenic diseases cause a significant number of familial strokes, particularly in patients with the ICH or SVD subtypes. This study supported the subtype-based WES is effective in delineating rare monogenic etiologies of familial stroke. Well-characterized phenotypes, including the clinical and neuroimaging manifestations, are the cornerstone to guide the diagnostic approach and enhance the genetic yield rate. We also proposed the potential of SNP array to be an alternative to screen ethnic gene hotspots. The integrative diagnostic pipeline could target the most prevalent monogenic causes and benefit patients and their families by introducing therapeutic interventions and therefore improve the personalized stroke medicine. Accumulating knowledge of the genotype–phenotype correlations would strengthen our understanding of the genomic landscape of stroke.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10571-022-01315-3>.

Acknowledgements We acknowledge the technical services provided by the Genomics Center for Clinical and Biotechnological Applications of the Cancer Progression Research Center (National Yang Ming Chiao Tung University), the National Core Facility for Biopharmaceuticals, Ministry of Science and Technology. We also thank the Clinical Research Core Laboratory of the Taipei Veterans General Hospital for providing experimental space and facilities. We thank the National Center for Genome Medicine at Academia Sinica for technical support in TWB 2.0 genotyping.

Author Contributions IHL and YTL contributed to conceptualization and funding acquisition. LHC contributed to data curation, formal analysis, investigation, and visualization. ; NFC, CYC, YSL, SLH, JYT, HCH, CJL, CPC, YCL, YTL, and IHL contributed to resources; CYT contributed to software; LHC, IHL, and YTL contributed to writing – original draft. NFC, CJL, CPC, YCL, CJJ, IHL, and YTL contributed to writing – review & editing.

Funding This study was supported by the Ministry of Science and Technology (MOST), Taiwan; Taipei Veterans General Hospital (TVGH); the Brain Research Center (BRC), National Yang Ming Chiao Tung University through the Featured Areas Research Center Program within the framework of the Higher Education Sprout Project by the Ministry of Education (MOE), Taiwan. IHL is grateful to MOST (110-2634-F-008-007-, 110-2320-B-075-002, 111-2320-B-075-010) and TVGH (V110C-124, V111C-213, V112C-036) for financial support. YTL is grateful to MOST (MOST 108-2321-B-010-012-MY2, 109-2314-B-075-055, 110-2628-B-075-014-) and TVGH (V110C-080, V111C-077). Both IHL and YTL are grateful to BRC/MOE (109BRC-B408, 110BRC-B408).

Data Availability The datasets analyzed during the current study are not publicly available due to IRB restriction. The corresponding author is willing to apply for IRB permission on reasonable request.

Declarations

Competing interests The authors declare no competing interests.

Ethical Approval This study was performed with the approval of the Taipei Veteran General Hospital Institutional Review Board (2016-09-010C, 2020-02-007B).

Consent for Publication Not applicable.

Research Involving Human and Animal Rights This study was performed in accordance with the tenets of the Declaration of Helsinki.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Adams HP Jr, Bendixen BH, Kappelle LJ, Biller J, Love BB, Gordon DL, Marsh EE 3rd (1993) 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* 24(1):35–41. <https://doi.org/10.1161/01.str.24.1.35>
- Akers A, Al-Shahi Salman R, Issam AA, Dahlem K, Flemming K, Hart B et al (2017) Synopsis of guidelines for the clinical management of cerebral cavernous malformations: consensus recommendations based on systematic literature review by the angioma alliance scientific advisory board clinical experts panel. *Neurosurgery* 80(5):665–680. <https://doi.org/10.1093/neuros/nyx091>
- Andrews S (2010) FastQC: A Quality Control Tool for High Throughput Sequence Data.
- Auer PL, Nalls M, Meschia JF, Worrall BB, Longstreth WT Jr, Seshadri S et al (2015) Rare and coding region genetic variants associated with risk of ischemic stroke: the NHLBI exome sequence project. *JAMA Neurol* 72(7):781–788. <https://doi.org/10.1001/jamaneurol.2015.0582>
- Bang OY, Chung J-W, Kim DH, Won H-H, Yeon JY, Ki C-S et al (2020) Moyamoya disease and spectrums of RNF213 vasculopathy. *Transl Stroke Res* 11(4):580–589. <https://doi.org/10.1007/s12975-019-00743-6>
- Bersano A, Markus HS, Quaglini S, Arbustini E, Lanfranconi S, Micieli G et al (2016) Clinical pregenetic screening for stroke monogenic diseases: results from Lombardia GENS registry. *Stroke* 47(7):1702–1709. <https://doi.org/10.1161/strokeaha.115.012281>
- Carrera C, Jimenez-Conde J, Derdak S, Rabionet K, Vives-Bauza C, Soriano-Tarrega C et al (2016) Whole exome sequencing analysis reveals TRPV3 as a risk factor for cardioembolic stroke. *Thromb Haemost* 116(6):1165–1171. <https://doi.org/10.1160/th16-02-0113>
- Choquet H, Pawlikowska L, Lawton MT, Kim H (2015) Genetics of cerebral cavernous malformations: current status and future prospects. *J Neurosurg Sci* 59(3):211–220
- Chung J-W, Kim BJ, Han M-K, Kang K, Park J-M, Park S-S et al (2016) Family history and risk of recurrent stroke. *Stroke* 47:1990–1996. <https://doi.org/10.1161/strokeaha.116.013148>
- Chung CP, Chen JW, Chang FC, Li WC, Lee YC, Chen LF, Liao YC (2020) Cerebral microbleed burdens in specific brain regions are associated with disease severity of cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy. *J Am Heart Assoc* 9(13):e016233. <https://doi.org/10.1161/jaha.120.016233>
- Dichgans M, Pulit SL, Rosand J (2019) Stroke genetics: discovery, biology, and clinical applications. *Lancet Neurol* 18(6):587–599. [https://doi.org/10.1016/S1474-4422\(19\)30043-2](https://doi.org/10.1016/S1474-4422(19)30043-2)
- Fischer A, Zalvide J, Faurobert E, Albiges-Rizo C, Tournier-Lasserre E (2013) Cerebral cavernous malformations: from CCM genes to endothelial cell homeostasis. *Trends Mol Med* 19(5):302–308. <https://doi.org/10.1016/j.molmed.2013.02.004>
- Fujimura M, Tominaga T, Kuroda S, Takahashi JC, Endo H, Ogawara K, Miyamoto S (2022) 2021 Japanese guidelines for the management of moyamoya disease: guidelines from the research committee on moyamoya disease and japan stroke society. *Neurol Med Chir (tokyo)* 62(4):165–170. <https://doi.org/10.2176/jns-nmc.2021-0382>
- Gurdasani D, Barroso I, Zeggini E, Sandhu MS (2019) Genomics of disease risk in globally diverse populations. *Nat Rev Genet* 20(9):520–535. <https://doi.org/10.1038/s41576-019-0144-0>
- Hamosh A, Scott AF, Amberger JS, Bocchini CA, McKusick VA (2005) Online Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders. *Nucleic Acids*

- Res 33(Database Issue):D514–D517. <https://doi.org/10.1093/nar/gki033>
- Huntington JA (2005) Molecular recognition mechanisms of thrombin. *J Thromb Haemost* 3(8):1861–1872. <https://doi.org/10.1111/j.1538-7836.2005.01363.x>
- Ilinca A, Kristoffersson U, Soller M, Lindgren AG (2016) Familial aggregation of stroke amongst young patients in Lund stroke register. *Eur J Neurol* 23(2):401–407. <https://doi.org/10.1111/ene.12881>
- Ilinca A, Samuelsson S, Piccinelli P, Soller M, Kristoffersson U, Lindgren AG (2019) A stroke gene panel for whole-exome sequencing. *Eur J Hum Genet* 27(2):317–324. <https://doi.org/10.1038/s41431-018-0274-4>
- Ilinca A, Martinez-Majander N, Samuelsson S, Piccinelli P, Truvé K, Cole J et al (2020) Whole-exome sequencing in 22 young ischemic stroke patients with familial clustering of stroke. *Stroke* 51(4):1056–1063. <https://doi.org/10.1161/STROKEAHA.119.027474>
- Ioannidis NM, Rothstein JH, Pejaver V, Middha S, McDonnell SK, Baheti S et al (2016) REVEL: an ensemble method for predicting the pathogenicity of rare missense variants. *Am J Hum Genet* 99(4):877–885. <https://doi.org/10.1016/j.ajhg.2016.08.016>
- Ishigami D, Miyawaki S, Imai H, Shimizu M, Hongo H, Dofuku S et al (2022) RNF213 p.Arg4810Lys heterozygosity in moyamoya disease indicates early onset and bilateral cerebrovascular events. *Transl Stroke Res* 13(3):410–419. <https://doi.org/10.1007/s12975-021-00956-8>
- Kalia SS, Adelman K, Bale SJ, Chung WK, Eng C, Evans JP et al (2017) Recommendations for reporting of secondary findings in clinical exome and genome sequencing, 2016 update (ACMG SF v2.0): a policy statement of the American College of Medical Genetics and Genomics. *Genet Med* 19(2):249–255. <https://doi.org/10.1038/gim.2016.190>
- Kim BJ, Kim JS (2014) Ischemic stroke subtype classification: an asian viewpoint. *J Stroke* 16(1):8–17. <https://doi.org/10.5853/jos.2014.16.1.8>
- Lee MJ, Chen YF, Fan PC, Wang KC, Wang K, Wang J, Kuo MF (2015) Mutation genotypes of RNF213 gene from moyamoya patients in Taiwan. *J Neurol Sci* 353(1–2):161–165. <https://doi.org/10.1016/j.jns.2015.04.019>
- Lee YC, Chung CP, Chao NC, Fuh JL, Chang FC, Soong BW, Liao YC (2018) Characterization of heterozygous HTRA1 mutations in Taiwanese patients with cerebral small vessel disease. *Stroke* 49(7):1593–1601. <https://doi.org/10.1161/STROKEAHA.118.021283>
- Lee Y-C, Chung C-P, Chang M-H, Wang S-J, Liao Y-C (2019) NOTCH3 cysteine-altering variant is an important risk factor for stroke in the Taiwanese population. *Neurology*. <https://doi.org/10.1212/WNL.0000000000008700>
- Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25(14):1754–1760. <https://doi.org/10.1093/bioinformatics/btp324>
- Li Q, Wang K (2017) InterVar: clinical interpretation of genetic variants by the 2015 ACMG-AMP guidelines. *Am J Human Genet* 100(2):267–280. <https://doi.org/10.1016/j.ajhg.2017.01.004>
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N et al (2009) The sequence alignment/Map format and SAMtools. *Bioinformatics* 25(16):2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>
- Liao Y-C, Hsiao C-T, Fuh J-L, Chern C-M, Lee W-J, Guo Y-C et al (2015) Characterization of CADASIL among the Han Chinese in Taiwan: distinct genotypic and phenotypic profiles. *PLoS ONE* 10(8):e0136501. <https://doi.org/10.1371/journal.pone.0136501>
- McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernysky A et al (2010) The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 20(9):1297–1303. <https://doi.org/10.1101/gr.107524.110>
- McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GRS, Thormann A, Flicek P, Cunningham F (2016) The ensembl variant effect predictor. *Genome Biol* 17(1):122. <https://doi.org/10.1186/s13059-016-0974-4>
- Merello E, Pavanello M, Consales A, Mascelli S, Raso A, Accogli A, Cama A, Valeria C, De Marco P (2016) Genetic screening of pediatric cavernous malformations. *J Mol Neurosci* 60(2):232–238. <https://doi.org/10.1007/s12031-016-0806-8>
- Meretoja A, Strbian D, Putaala J, Curtze S, Haapaniemi E, Mustanoja S et al (2012) SMASH-U: a proposal for etiologic classification of intracerebral hemorrhage. *Stroke* 43(10):2592–2597. <https://doi.org/10.1161/STROKEAHA.112.661603>
- Miksch S, Lumsden A, Guenther UP, Foerzler D, Christen-Zäch S, Daugherty C et al (2005) Molecular genetics of pseudoxanthoma elasticum: type and frequency of mutations in ABCC6. *Hum Mutat* 26(3):235–248. <https://doi.org/10.1002/humu.20206>
- Mishra A, Chauhan G, Violleau MH, Vojinovic D, Jian X, Bis JC et al (2019) Association of variants in HTRA1 and NOTCH3 with MRI-defined extremes of cerebral small vessel disease in older subjects. *Brain* 142(4):1009–1023. <https://doi.org/10.1093/brain/awz024>
- Miyawaki Y, Suzuki A, Fujita J, Maki A, Okuyama E, Murata M et al (2012) Thrombosis from a prothrombin mutation conveying antithrombin resistance. *N Engl J Med* 366(25):2390–2396. <https://doi.org/10.1056/NEJMoa1201994>
- Mukawa M, Nariai T, Onda H, Yoneyama T, Aihara Y, Hirota K et al (2017) Exome sequencing identified CCER2 as a novel candidate gene for moyamoya disease. *J Stroke Cerebrovasc Dis* 26(1):150–161. <https://doi.org/10.1016/j.jstrokecerebrovasdis.2016.09.003>
- Plomp AS, Hu X, de Jong PTVM, Bergen AAB (2004) Does autosomal dominant pseudoxanthoma elasticum exist? *Am J Med Genet Part A* 126A(4):403–412. <https://doi.org/10.1002/ajmg.a.20632>
- Rabbani B, Tekin M, Mahdih N (2014) The promise of whole-exome sequencing in medical genetics. *J Hum Genet* 59(1):5–15. <https://doi.org/10.1038/jhg.2013.114>
- Riant F, Bergametti F, Ayrignac X, Boulday G, Tournier-Lasserre E (2010) Recent insights into cerebral cavernous malformations: the molecular genetics of CCM. *FEBS J* 277(5):1070–1075. <https://doi.org/10.1111/j.1742-4658.2009.07535.x>
- Riant F, Bergametti F, Fournier HD, Chapon F, Michalak-Provost S, Cecillon M et al (2013) CCM3 mutations are associated with early-onset cerebral hemorrhage and multiple meningiomas. *Mol Syndromol* 4(4):165–172. <https://doi.org/10.1159/000350042>
- Santoro C, Giugliano T, Kraemer M, Torella A, Schwitala JC, Cirillo M et al (2018) Whole exome sequencing identifies MRVII as a susceptibility gene for moyamoya syndrome in neurofibromatosis type 1. *PLoS ONE* 13(7):e0200446. <https://doi.org/10.1371/journal.pone.0200446>
- Sauvigny T, Alawi M, Krause L, Renner S, Spohn M, Busch A et al (2020) Exome sequencing in 38 patients with intracranial aneurysms and subarachnoid hemorrhage. *J Neurol* 267(9):2533–2545. <https://doi.org/10.1007/s00415-020-09865-6>
- Shenkar R, Shi C, Rebeiz T, Stockton RA, McDonald DA, Mikati AG et al (2015) Exceptional aggressiveness of cerebral cavernous malformation disease associated with PDCD10 mutations. *Genet Med* 17(3):188–196. <https://doi.org/10.1038/gim.2014.97>
- Starby H, Delavaran H, Andberg G, Lovkvist H, Norrving B, Lindgren A (2014) Multiplicity of risk factors in ischemic stroke patients: relations to age, sex, and subtype—a study of 2,505 patients from the lund stroke register. *Neuroepidemiology* 42(3):161–168. <https://doi.org/10.1159/000357150>
- Takagi Y, Murata M, Kozuka T, Nakata Y, Hasebe R, Tamura S et al (2016) Missense mutations in the gene encoding prothrombin

- corresponding to Arg596 cause antithrombin resistance and thrombomodulin resistance. *Thromb Haemost* 116(6):1022–1031. <https://doi.org/10.1160/th16-03-0223>
- Tan RYY, Traylor M, Megy K, Duarte D, Deevi SVV, Shamardina O et al (2019) How common are single gene mutations as a cause for lacunar stroke? *Neurology* 93(22):e2007. <https://doi.org/10.1212/WNL.0000000000008544>
- Toyoda K, Koga M, Hayakawa M, Yamagami H (2015) Acute reperfusion therapy and stroke care in Asia after successful endovascular trials. *Stroke* 46(6):1474–1481. <https://doi.org/10.1161/STROKEAHA.115.008781>
- Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, Del Angel G, Levy-Moonshine A et al (2013) From FastQ data to high confidence variant calls: the genome analysis toolkit best practices pipeline. *Curr Protoc Bioinformatics* 43(1):11–10. <https://doi.org/10.1002/0471250953.bi1110s43>
- Wei C-Y, Yang J-H, Yeh E-C, Tsai M-F, Kao H-J, Lo C-Z et al (2021) Genetic profiles of 103,106 individuals in the Taiwan Biobank provide insights into the health and history of Han Chinese. *Npj Genomic Med* 6(1):10. <https://doi.org/10.1038/s41525-021-00178-9>
- Yadla S, Jabbour PM, Shenkar R, Shi C, Campbell PG, Awad IA (2010) Cerebral cavernous malformations as a disease of vascular permeability: from bench to bedside with caution. *Neurosurg Focus* 29(3):E4. <https://doi.org/10.3171/2010.5.FOCUS10121>
- Yaghi S, Shu L, Bakradze E, Omran SS, Giles JA, Amar JY et al (2022) Direct oral anticoagulants versus warfarin in the treatment of cerebral venous thrombosis (ACTION-CVT): a multicenter international study. *Stroke* 53(3):728–738. <https://doi.org/10.1161/STROKEAHA.121.037541>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.