Author/s:
Sikta, N;Gooley, S;Green, TE;Hoeper, O;Witkowski, T;Bennett, C;Francis, D;Reid, J;Mao, K;Awad, M;Roberts-Thomson, S;Bulluss, K;Clark, J;Scheffer, IE;Perucca, P;Bennett, MF;Bahlo, M;Berkovic, SF;Hildebrand, MS

Title:
Improving genetic diagnostic yield in familial and sporadic cerebral cavernous malformations: detection of copy number and deep Intronic variants

Date:
2025-08-01

Citation:
Sikta, N., Gooley, S., Green, T. E., Hoeper, O., Witkowski, T., Bennett, C., Francis, D., Reid, J., Mao, K., Awad, M., Roberts-Thomson, S., Bulluss, K., Clark, J., Scheffer, I. E., Perucca, P., Bennett, M. F., Bahlo, M., Berkovic, S. F. & Hildebrand, M. S. (2025). Improving genetic diagnostic yield in familial and sporadic cerebral cavernous malformations: detection of copy number and deep Intronic variants. Human Molecular Genetics, 34 (15), pp.1286-1293. https://doi.org/10.1093/hmg/ddaf077.

Persistent Link:
https://hdl.handle.net/11343/361822

Human Molecular Genetics, 2025, Vol. 34, 15, 1286–1293
https://doi.org/10.1093/hmg/ddaf077
Advance access publication date 22 May 2025
Original Article

Improving genetic diagnostic yield in familial and sporadic cerebral cavernous malformations: detection of copy number and deep Intronic variants

Neblina Sikta[1,†], Samuel Gooley[1,2,†], Timothy E. Green[1], Olivia Hoeper[1], Tom Witkowski[1], Caitlin Bennett[1], David Francis[3], Joshua Reid[1,4], Kevin Mao[1], Mohammed Awad[5], Samuel Roberts-Thomson[6], Kristian Bulluss[7], Jonathan Clark[8], Ingrid E. Scheffer[1,2,9], Piero Perucca[1,2], Mark F. Bennett[1,4,10], Melanie Bahlo[4,10], Samuel F. Berkovic[1,2,†], Michael S. Hildebrand[1,11,†,*]

[1]Epilepsy Research Centre, Department of Medicine, The University of Melbourne, Austin Health, 245 Burgundy St, Heidelberg, VIC 3084, Australia
[2]Bladin-Berkovic Comprehensive Epilepsy Program, Department of Neurology, Austin Health, 145 Studley Rd, Heidelberg, VIC 3084, Australia
[3]Victorian Clinical Genetics Services and Murdoch Children's Research Institute, The Royal Children's Hospital, 50 Flemington Rd, Parkville, VIC 3052, Australia
[4]Population Health and Immunity Division, The Walter and Eliza Hall Institute of Medical Research, 1G Royal Parade, Parkville, VIC 3052, Australia
[5]Department of Neurosurgery, The Royal Melbourne hospital, 300 Grattan St, Parkville, VIC 3052, Australia
[6]Anatomical Pathology, The Royal Melbourne hospital, 300 Grattan St, Parkville, VIC 3052, Australia
[7]Department of Neurosurgery, Austin Health, 145 Studley Rd, Heidelberg, VIC 3084, Australia
[8]Anatomical Pathology, Austin Health, Heidelberg, 145 Studley Rd. VIC 3084, Australia
[9]Department of Paediatrics, The University of Melbourne, The Royal Children's Hospital, 50 Flemington Rd, Florey Institute of Neuroscience and Mental Health, 30 Royal Parade, and Murdoch Children's Research Institute, 50 Flemington Rd, Parkville, VIC 3052, Australia
[10]Department of Medical Biology, University of Melbourne, Royal Parade, Parkville, VIC 3052, Australia
[11]Neuroscience Group, Murdoch Children's Research Institute, 50 Flemington Rd, Parkville, VIC 3052, Australia

*Corresponding author. Professor Michael S. Hildebrand, Epilepsy Research Centre, Level 1, Melbourne Brain Centre, 245 Burgundy Street, Heidelberg, Victoria 3084, Australia. E-mail: michael.hildebrand@unimelb.edu.au

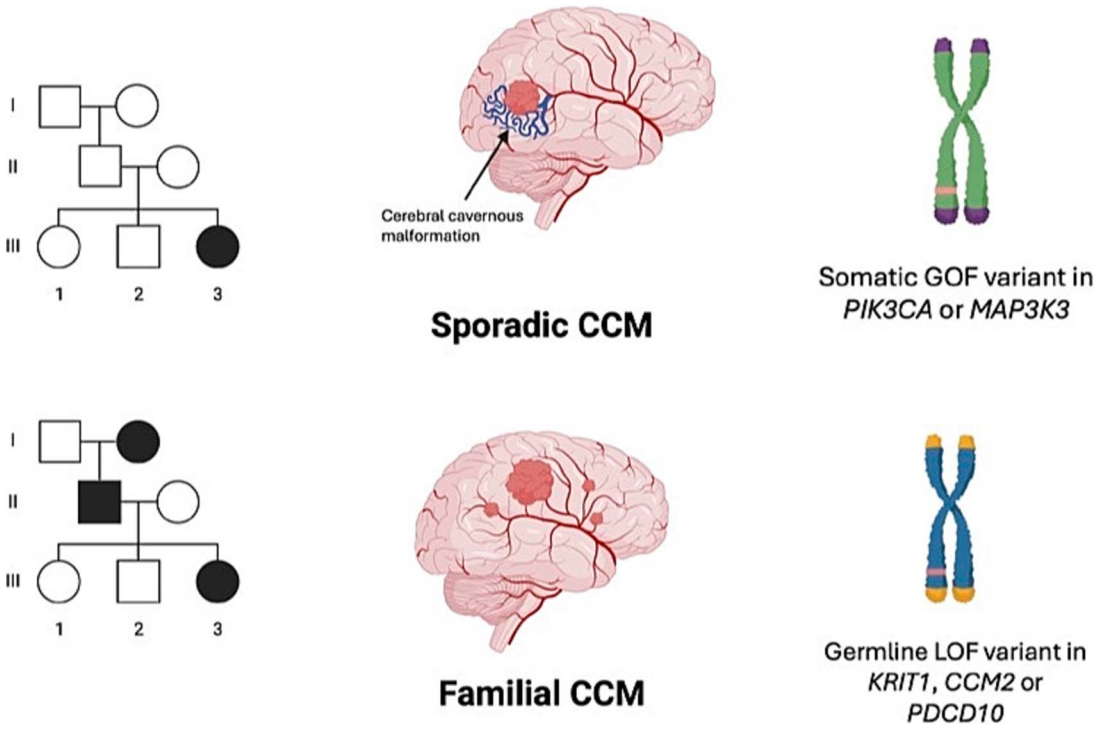
†Neblina Sikta, Samuel Gooley, Samuel F. Berkovic and Michael S. Hildebrand contributed equally to this work.

Abstract

Cerebral cavernous malformations (CCMs) are intracranial vascular lesions associated with risk of haemorrhages and seizures. While the majority are sporadic and often associated with somatic variants in *PIK3CA* and *MAP3K3*, around 20% are familial with germline variants in one of three CCM genes—*KRIT1/CCM1*, *CCM2* and *PDCD10/CCM3*. We performed comprehensive phenotyping and genetic analysis of nine multiplex families and ten sporadic individuals with CCM. In the familial cases, initial standard analyses had a low yield, we therefore searched for small copy number changes and deep intronic variants. Subsequently, pathogenic germline variants in *KRIT1/CCM1* or *CCM2* were identified in all 9 multiplex families. Single or multiple exon deletions or splice site variants in *KRIT1/CCM1* were found in 3/9 families. Where cavernous malformation tissue was available, second hit somatic *PIK3CA* variants were identified in 4/7 individuals. These 4 individuals were from separate families with germline *KRIT1/CCM1* variants. In 8/10 sporadic cases, we detected recurrent pathogenic somatic *PIK3CA, MAP3K3 or CCM2* variants. All familial cases had multiple CCMs, whereas the sporadic cases had a single lesion only, which was in the temporal lobe in 9/10 individuals. Our comprehensive approach interrogating deep intronic variants combined with detection of small copy number variants warrants implementation in standard clinical genetic testing pipelines to increase diagnostic yield. We also build on the established second hit germline and somatic variant mechanism in some CCM lesions. Genetic diagnosis has clinical implications such as reproductive counselling and provides potential eligibility for precision medicine therapies to treat rapidly growing CCMs.

Received: October 10, 2024. **Revised:** April 10, 2025. **Accepted:** May 1, 2025

Graphical Abstract



Keywords: Cerebral cavernous malformations; Somatic mosaicism; Deep intronic variant; Copy number variant; Germline variant

Introduction

Cerebral cavernous malformations (CCMs) are vascular anomalies characterised by clusters of thin-walled capillaries. The structural irregularities of these capillary clusters can alter the blood flow in the brain and disrupt normal brain function, leading to neurological manifestations such as haemorrhages and seizures. CCMs are estimated to occur in ~0.5% of the population and approximately 80% are single brain lesions in individuals with no family history of CCM (sporadic cases) [1–3]. In individuals with familial CCM, multiple vascular brain lesions occur in affected family members. Due to the occurrence of multiple lesions, familial CCMs are generally associated with more severe symptoms [4].

Familial CCM follows an autosomal dominant inheritance pattern and is linked to germline loss-of-function pathogenic variants in three primary CCM genes that encode proteins involved in capillary development: *KRIT1/CCM1* [HGNC: 6380, MIM: 604214], *CCM2* [HGNC: 1745, MIM: 607929], and *PDCD10/CCM3* [HGNC: 8767, MIM: 603285] [5–9]. In contrast, sporadic CCMs are commonly associated with somatic gain-of-function variants in *PIK3CA* [HGNC: 8977, MIM: 171834] or *MAP3K3* [HGNC: 6881, MIM: 602539] that encode proteins in the growth signalling mTOR and Ras/Raf/MAPK pathways, respectively. [10–12] Several studies of mosaicism in CCM tissue have identified second or even third 'hit' somatic pathogenic variants in a subset of both familial and sporadic cases indicating these benign vascular lesions can accumulate mutations in a similar fashion to malignant tumours [13–15]. In both the proposed second and third hit models, familial CCM lesions arise from a germline loss-of-function variant in any of the 3 CCM genes (*KRIT1*, *CCM2* or *PDCD10*), and can accumulate a mosaic loss-of-function variant in a CCM gene, and/or a gain-of-function *PIK3CA* or *MAP3K3* somatic pathogenic variant that promotes growth of the lesion.

While recent work has uncovered links between specific genotypes and certain clinical and radiological features of CCMs [10, 16], the observed clinical heterogeneity remains incompletely understood. Variability in disease severity and progression among affected individuals, even within families, underscores the need for comprehensive genetic characterization and phenotypic correlation. Here, we performed comprehensive phenotyping and genetic analysis of a cohort of patients with familial and sporadic CCM cases including interrogation of small indels and intronic variants to increase diagnostic yield and correlate with clinical manifestations.

Results

Phenotyping

Clinical data was reviewed for 16 familial (from nine families) and 10 sporadic cases (Tables 1 and 2). The familial cases all had multiple CCMs with most having more than 10 supra- and infratentorial lesions. The sporadic cases all had a single lesion only, located in the temporal lobe in 9/10 cases, and the cingulate gyrus in one. All the sporadic cases and 15/16 familial cases had focal epilepsy. The predilection for lesions in sporadic cases to be localised to the temporal lobe and the high frequency of epilepsy in our patients is likely due to ascertainment bias as our cohort was recruited through epilepsy programs. The mean age of seizure onset was similar in the familial cases (18 years; range 1–55 years) and the sporadic cases (21.7 years; range 10–40 years). Symptomatic intracerebral haemorrhages (ICH) occurred in 8/16 familial cases (including in 7/12 with a *KRIT1/CCM1* variant and in 1/4 with *CCM2* pathogenic variants) but did not occur in any of the sporadic cases. All 10 sporadic cases underwent epilepsy surgery with 7/10 achieving Engel class I outcomes (free of disabling

Table 1. Clinical and genetic summary of familial cases.

Family	ID/Sex	CCM burden	Seizure onset (yrs)	Seizure focus	Seizure types	ICH	Surgery	Germline variant	CCM tissue	Somatic variant (ddPCR)
A	IV.4 (F)	7 supra 5 infra	12	Multifocal	FAS, FIAS, FBTCS	0	Lesionectomy	*KRIT1* c.1524_1528delAAGAA	FFPE	*PIK3CA* p.E542K VAF 1.4%
B	II.4 (F)	3 supra	6	Temporal	FAS, FIAS, FBTCS	0	Lesionectomy	*KRIT1* c.136_139delAAA	FFPE	*PIK3CA* p.H1047R VAF 5.3%
C	III.2 (M)	Multiple (MRI report only)	1	Temporal + extra-temporal	FIAS	1	Right ATL	*KRIT1* insertion in exon 16/17	FFPE	-ve ddPCR
	III.1 (F)	7 supra 3 infra	1	Temporal	FAS, FIAS, FBTCS	1	Lesionectomies & Right ATL	*KRIT1* insertion in exon 16/17	FFPE	*PIK3CA* p.E545K VAF 2.8%
	II.1 (M)	>10 supra > 10 infra	4	Left frontal	FAS, FIAS, FBTCS	1	Lesionectomy	*KRIT1* insertion in exon 16/17	-	N/T
	I.2 (F)	Multiple (MRI report only)	53	Unable to localize	FIAS, FBTCS	0	N	*KRIT1* insertion in exon 16/17	-	N/T
D	III.8 (F)	>10 supra > 10 infra	30	Unable to localize	Focal motor, FBTCS	1	N	*CCM2* +ve c.205-2A > G	-	N/T
	II.1 (F)	Multiple (MRI report only)	29	Unable to localize	FAS, FIAS, FBTCS	0	N	*CCM2* +ve c.205-2A > G	-	N/T
E	II.1 (M)	3 supra	4	Unable to localize	FIAS	1	N	*KRIT1* c.1362_1363del p.Gln455ArgfsTer24	-	N/T
	I.2 (F)	Multiple (MRI report only)	No epilepsy	NA	NA	0	N	*KRIT1* c.1362_1363del p.Gln455ArgfsTer24	-	N/T
F	II.2 (F)	>10 supra > 10 infra	18	Unable to localize	FAS, FIAS	0	N	*CCM2* c.683_686del p.Phe228TrpfsTer63	-	N/T
	III.1 (M)	>10 supra 8 infra	9	Occipital	FAS, FIAS, FBTCS	0	Lesionectomies	*CCM2* c.683_686del p.Phe228TrpfsTer63	FFPE	-ve ddPCR
G	II.2 (M)	>10 supra > 10 infra	55	Unable to localize	FBTCS	4	Lesionectomy	*KRIT1* Heterozygous deletion exon 18	FFPE	-ve ddPCR
H	II.6 (M)	>10 supra 2 infra	19	Unable to localize	FBTCS	1	Lesionectomy	*KRIT1* Heterozygous deletion exon 16,17 & 18	FFPE	*PIK3CA p.*E545K VAF 2.9%
J	II.2 (M)	>10 supra > 10 infra	25	Unable to localize	FBTCS	1	N	*KRIT1* c.262 + 112A > G	-	N/T
	II.1 (M)	Multiple supra- and infra	15	Unable to localize	Focal	0	N	*KRIT1* c.262 + 112A > G	-	N/T

ATL: Anterior temporal lobectomy; **ddPCR**: droplet digital polymerase chain reaction; **FAS**: Focal aware seizure; **FIAS**: Focal impaired awareness seizure; **FBTCS**: Focal to bilateral tonic–clonic seizure; **FFPE**: Formalin-Fixed Paraffin-Embedded; **ICH**: Symptomatic intracerebral haemorrhage; **Infra**: Infratentorial; **N:** No; **N/T**: Not tested; **Supra**: Supratentorial; **VAF**: Variant allele fraction; **–ve:** Negative

Table 2. Clinical and genetic summary of sporadic CCM cases.

Case (Sex)	Location	CCM size (cm)	Seizure onset (yrs)	Seizure types	Age at surgery (yrs)	Surgical outcome (Engel)	Somatic variant/s	VAF	Detection method
1 (M)	Left cingulate	2	12	FIAS, FBTCS	23 & 24	4	*PIK3CA* p.H1047R	5.9%	ddPCR of FFPE
							CCM2 p.A147fs	5.9%	ddPCR of FFPE
2 (F)	Left mesial temporal	1.8	10	FIAS	21	4	*PIK3CA* p.H1047R	10%	ddPCR of FFPE
3 (M)	Left temporal pole	1.2	20	FIAS, FBTCS	41	1	*MAP3K3* p.I441M	8%	ddPCR of FFPE
							PIK3CA p.E545K	1.3%	ddPCR of FFPE
							CCM2 Del exons 6–10	13%	WES of FFPE, ddPCR
4 (F)	Right mesial temporal	1	40	FIAS	56	2	*MAP3K3* p.I441M	2.4%	ddPCR of FFPE
5 (F)	Right posterior temporal	1.5	29	FBTCS	33	1	*PIK3CA* p.C420R	2.0%	ddPCR of FFPE
6 (F)	left fusiform gyrus	1.5	31	FIAS, FBTCS	35	1	*PIK3CA* p.E542K	7.5%	ddPCR of FFPE
							MAP3K3 p.I441M	4.2%	ddPCR of FFPE
7 (F)	Right superior temporal gyrus	2.5	12	FAS, FIAS	29	1	*PIK3CA* p.E545K	2.3%	ddPCR of FFPE
8 (M)	Left mesial temporal	1.7	26	FAS	28	1	*MAP3K3* p.I441M	4.4%	ddPCR of FFPE
9 (M)	Left mesial temporal	1.5	18	FAS, FIAS	26	1	negative	NA	ddPCR of FFPE
10 (M)	Left anterior temporal	1.8	19	FAS, FIAS, FBTCS	27	1	negative	NA	ddPCR of FFPE

CCM size: maximum diameter in the axial plan on T1 weighted MRI or as measured by the pathologist after surgical resection; **ddPCR**: droplet digital polymerase chain reaction; **Engel:** Engel classification of outcome after epilepsy surgery (1: Free of disabling seizures, 2: Rare disabling seizures, 3: Worthwhile improvement, 4: No worthwhile improvement) [23]; **FAS**: Focal aware seizure; **FIAS**: Focal impaired awareness seizure; **FBTCS**: Focal to bilateral tonic–clonic seizure; **FFPE**: Formalin-Fixed Paraffin-Embedded; **VAF:** variant allele fraction.

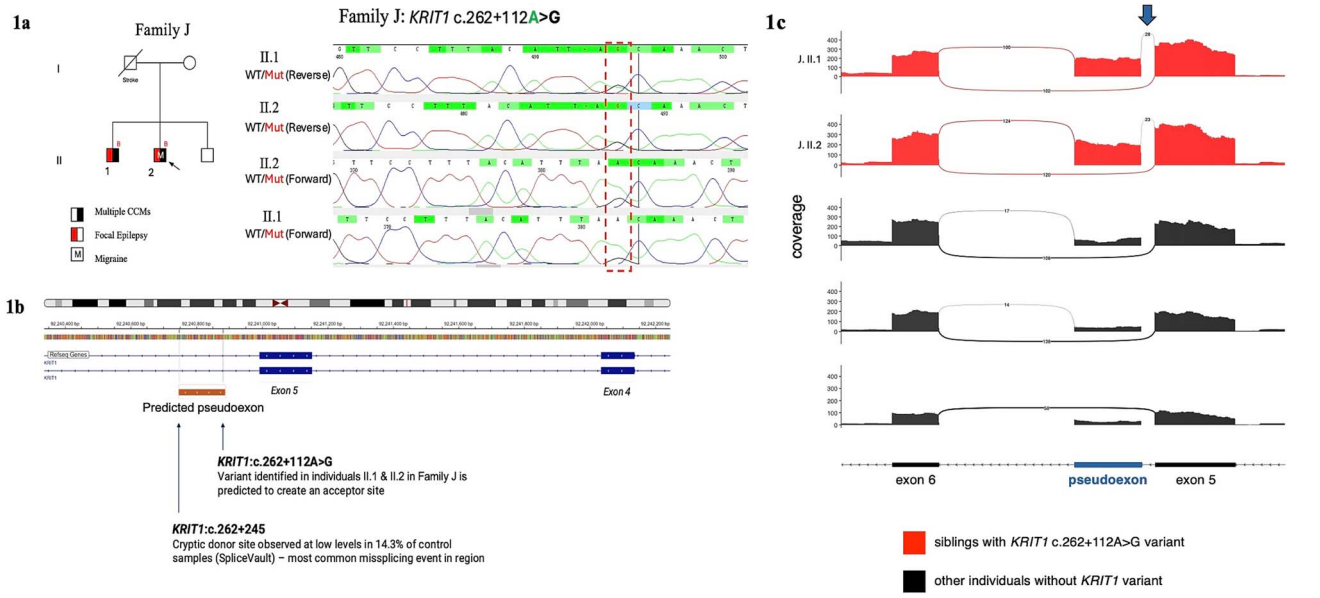


Figure 1. (a) Shows a heterozygous deep intronic variant in *KRIT1* (c.262 + 112A > G) in family J identified on whole exome sequencing and confirmed via sanger sequencing. (b) Shows the aberrant splicing event predicted to create an acceptor site 112 base pairs upstream of exon 5 by SpliceAI. (c) Shows a sashimi plot of RNA-sequencing reads. Splicing from the 3′ end of exon 5 to the 5′ end of the predicted pseudoexon (arrow) observed in both siblings with the variant (top 2 plots) but not in controls (bottom 3 plots).

seizures) and 3/10 achieving Engel class II-IV (mean follow up 11.9 years; range 2.5–29 years; Table 2). In the familial group, 2 individuals had surgery to resect haemorrhagic lesions and 6 had epilepsy surgery with 4/6 individuals achieving Engel class I and 2/6 Engel class II-IV (mean follow up 11.5 years; range 4–19 years).

Genetic analysis

Loss-of-function pathogenic germline variants in familial cases

We identified pathogenic germline variants in all nine families with CCM. In four families, pathogenic single nucleotide or small indel variants were detected on standard clinical or research testing (Table 1). These comprised a *KRIT1* frameshift variant (c.1524_1528delAAGAA) in family A that was confirmed on research exome sequencing, an inframe deletion in *KRIT1* (c.136_139delAAA) in family B, an insertion across exon 16/17 in family C (Fig. 2b), and a *CCM2* splice site variant (c.205-2A > G) in family D (Table 1).

Of the remaining five families, three (E, F and G) were negative on prior clinical testing. The negative results from the previous clinical testing in Family E was due to the proband having an epilepsy gene panel with 302 epilepsy genes that did not include *KRIT1*. In Family F, the proband had clinical testing of *KRIT1*, *CCM2* & *PCDC10* with both sequencing and Multiplex Ligation-dependent Probe Amplification but neither method detected any variants. Upon re-analysis of available sequencing data from Family E, we identified a recurrent frameshift deletion variant in *KRIT1* (c.1362_1363del). In Family F, we subsequently performed whole genome sequencing on both affected individuals and detected a recurrent frameshift deletion in *CCM2* (c.683_686del) (Table 1 and Supplementary Fig. 1). Both variants have previously been reported as pathogenic variants in familial CCM [17, 18]. The proband in Family G previously had clinical testing of *KRIT1* with denaturing high-performance liquid chromatography (dHPLC), which would not have been able to detect a deletion of an entire exon.

We performed research exome sequencing on Families G, H and J. No pathogenic variants were identified on initial analysis of CCM genes in Families G and H. However, subsequent interrogation for potential copy number variants in the sequencing data from these families revealed decreased read depth in two regions of *KRIT1*. A heterozygous deletion of *KRIT1* exon 18 in family G (Supplementary Fig. 2) and *KRIT1* exons 15, 16 and 17 in family H (Supplementary Fig. 3) were independently confirmed by chromosomal microarray and droplet digital PCR. In family J, a heterozygous deep intronic variant in *KRIT1* (c.262 + 112A > G) was identified, which was predicted to create an acceptor site 112 base pairs upstream of exon 5 by SpliceAI, which would lead to aberrant splicing and the introduction of a pseudoexon. Subsequent RNA sequencing of blood-derived RNA from the two affected individuals in this family confirmed the expression of transcripts containing the predicted pseudoexon. (Table 1 and Fig. 1a–c).

Recurrent gain-of-function somatic pathogenic variants in familial cases

Droplet digital PCR (ddPCR) analysis of resected CCM tissue available from seven individuals in six families with CCM revealed the presence of recurrent somatic variants in *PIK3CA* in four individuals from unrelated families. These included: p.E545K in two (family H:II.6 and family C:III.1, Table 1 and Fig. 2a), the p.H1047R in individual II.4 from family B, and p.E542K in individual IV.4 from family A; all were at low variant allele fractions of < 6% (Table 1).

Recurrent gain-of-function somatic pathogenic variants in sporadic cases

We identified recurrent somatic pathogenic gain-of-function variants in *MAP3K3* or *PIK3CA* in 8/10 sporadic cases tested using droplet digital PCR (ddPCR). The *MAP3K3* (p.I441M) variant was observed in 4/10 sporadic cases (3, 4, 6 and 8; Table 2) with VAFs ranging from 2.4% to 8%. Three *PIK3CA* variants (p.H1047R, p.E545K and p.E542K) were detected in 5 sporadic cases (1, 2, 3, 6 and 7; Table 2). Among the 8 sporadic cases with at least one gain-of-function variant, three individuals harboured multiple somatic variants in different genes, including cases 1 and 6 with two somatic variants and case 3 with three somatic variants (Table 2).

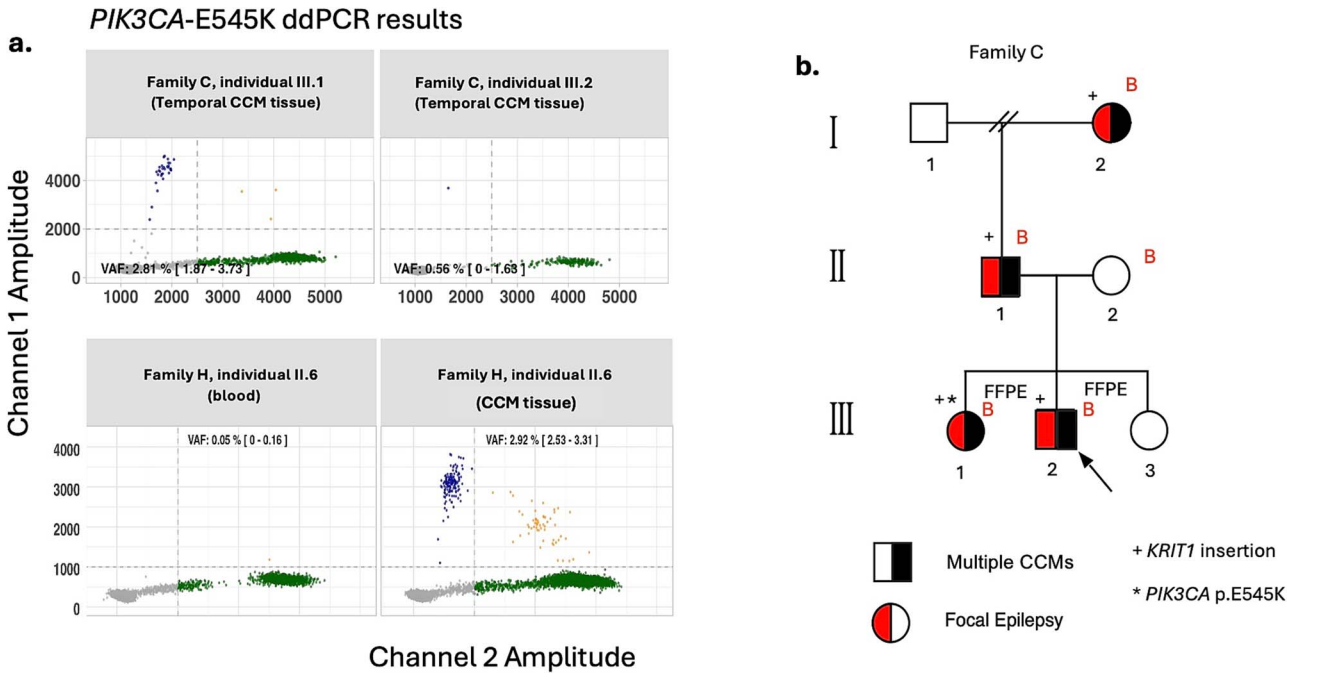


Figure 2. (a). Droplet-digital PCR results identified a recurrent *PIK3CA* (p.E545K) gain-of-function variant identified at a variant allele frequency (VAF) ~ 3% in the resected CCM tissue from individual III.1 from family C, but the variant was not detected in the resected CCM tissue from her brother, III.2 (top row). The variant was also detected in CCM tissue from individual II.6 from family H at a VAF ~ 3% but was not detected on testing of their peripheral blood sample (bottom row). (b). Pedigree showing inherited germline *KRIT1* variant present only in affected members. The recurrent gain-of-function somatic variant in *PIK3CA* (p.E545K) was only present in one of the affected siblings (III.1).

Discussion

There have been significant advances in understanding the genetic aetiology of CCMs over the past decade. [11–13, 19, 20] Our study highlights the importance of interrogating deep intronic regions for splice site variants and searching for small copy number variants in sequencing data to increase the diagnostic yield. In our cohort, one third of families initially genetically unsolved were subsequently found to have a causative deep intronic or copy number variant upon re-analysis. Standard testing methods lack the sensitivity to detect all types of causative variants. For example, families G and H had heterozygous single or multiple exon deletions of *KRIT1* that would not have been detected on a standard clinical gene panel or clinical exome optimised for single nucleotide variant detection. In family J, we identified a deep intronic variant 112 base pairs upstream from exon 5 of *KRIT1* that would have been outside of the region interrogated on a clinical gene panel or clinical exome.

Our study also builds on the recently observed phenomenon of second-hit somatic variants in CCM lesions of familial cases. Somatic gain-of-function pathogenic *PIK3CA* variants were detected in addition to a germline *KRIT1* variant in 4/7 familial cases where CCM tissue was available. We also identified three sporadic cases with at least two somatic variants. Two studies have shown that *KRIT1*, *CCM2* or *PDCD10* loss-of-function variants and the recurrent *MAP3K3* gain-of-function variant all increase Ras/Raf/MAPK pathway signalling [10, 19, 21], whereas recurrent *PIK3CA* variants increase PI3K–mTOR pathway signalling. [13] Hyperactivation of either, or both, of these signalling pathways likely drives overgrowth of CCMs contributing to clinical sequelae. A better understanding of these molecular mechanisms could lead to improved treatment options, for example the drug inhibitor, alpelisib, that specifically targets *PIK3CA* or other inhibitors that target the Ras (e.g. trametinib) or mTOR (e.g. sirolimus) signalling pathways could be repurposed to treat drug-resistant epilepsy due to CCMs.

Clinical phenotyping showed the sporadic cases all had a single CCM, whereas the familial cases all had multiple CCMs, consistent with the typical pattern seen in individuals with germline *KRIT1*, *CCM2* or *PDCD10* variants. All except one individual had epilepsy, reflecting their recruitment through epilepsy programs. Interestingly, the familial and sporadic cases had a similar mean age of seizure onset and comparable outcomes from epilepsy surgery. Symptomatic ICH occurred in 50% of familial cases but did not occur in any of the sporadic cases despite follow up for an average of 11.9 years (range 2.5–29 years).

There were several limitations of this study. First, the lack of resected CCM tissue from all familial cases to search for second hit variants. Second, screening for only recurrent somatic variants in *PIK3CA* and *MAP3K3* in CCM tissue of familial and sporadic cases precluded discovery of somatic variants in novel genes. Overall, our findings highlight the phenotypic and genetic heterogeneity of familial and sporadic CCMs and support more comprehensive clinical testing approaches incorporating deep coverage to discover somatic variants and interrogation of intronic regions to increase genetic diagnostic yield.

Materials and methods

Cohort and sample collection

Our CCM cohort consists of nine families with 16 affected individuals and ten sporadic cases. We collected blood and/or saliva samples from all family members and sporadic individuals. Lesional tissue was available for seven familial and all ten sporadic cases who underwent CCM resection through an epilepsy surgery program. Written informed consent (Austin HREC H2007/02961) was obtained from all cases.

Genetic testing

Two familial CCM cases underwent clinical gene panel testing (family A: IV.4 and family E: II.1). The individual II.1 from family E underwent commercial clinical genetic testing via Invitae using a gene panel consisting of 302 epilepsy genes. The individual IV.4 from family A had a cerebral vascular malformations gene panel that included *KRIT1*, *CCM2*, *PDCD10* and 14 other vascular malformations genes (*ACTA2*, *ACVRL1*, *ANGPYL6*, *CBL*, *COL3A1*, *ENG*, *GUCY1A3*, *PDGFRB*, *RASA1*, *RNF213*, *SAMHD1*, *SCL2A10*, *SMAD4* & *YY1AP1*) at an average read depth ~ 50x. Affected individuals from family C, individual II.4 from family B, III.8 from family D, II.2 from family F, and II.2 from family G underwent research-based genetic testing for SNPs and small indels via denaturing High-Performance Liquid Chromatography (dHPLC) and Single-Strand Conformation Polymorphism techniques. dHPLC was only performed to screen variants in *KRIT1*. Additionally, individual III.8 from family D was screened for pathogenic variants via targeted sequencing of *CCM2* and *PDCD10* (*CCM3*). Finally, individual III.1 from family F underwent screening for copy number variants in the 3 CCM genes (*KRIT1*, *CCM2* and *PDCD10*) via Multiplex Ligation-dependent Probe Amplification.

High depth sequencing

Next-Generation Sequencing (NGS) data was obtained from DNA derived from blood samples of individuals who did not have previous positive findings on clinical testing. Whole exome sequencing (WES) data was obtained from seven samples (family A: IV.4, family D: II.1, family E: II.1, family F: III.1, family G: II.1, family H: II.6 and family J II.2 from Table 1), achieving a target coverage of 50x. NGS data was processed following GATK Best Practices [13]. Paired reads were aligned to the hg38 reference genome using BWA-MEMv0.7.17-r1188. GATK (v4.1.9) was used to mark duplicates from the mapped reads and for base quality score recalibration. GATK HaplotypeCaller was used to call germline SNPs and indels and GATK Mutect2 v4.2.5.0 was used to search for mosaic variants from the aligned data. Germline and mosaic variants were annotated using VEP [14] and filtered for variants in CCM genes (*KRIT1*, *CCM2*, *PCDC10*, *MAP3K3* and *PIK3CA*) that are ultra-rare (allele count < 5 in gnomAD v.4.0.0). We searched for loss-of-function variants based on the following criteria: variants resulting in frameshift, splice acceptor or donor site alterations, stop or start codon losses or gains, with a CADD Phred score of ≥ 20 or a SpliceAI score > 0.5. Variants in the canonical transcripts of CCM-related genes *KRIT1* (NM_194454.3), *CCM2* (NM_031443.4), *PDCD10* (NM_007217.4), *PIK3CA* (NM_006218.5), and *MAP3K3* (NM_002401.5) were analysed to identify those associated with cerebral cavernous malformations of the affected individuals.

Sanger sequencing was used to validate variants identified via NGS data and in affected individuals without any available sequencing data (family E: I.2, family F: II.2 and family J: II.1 from Table 1).

Coverage analysis

In familial cases with no pathogenic small variants (SNVs/indels) identified from sequencing data, we assessed read depth across the 3 *CCM* genes to identify potential copy number variants (CNVs). We calculated read depth across the exons of *KRIT1*, *CCM2*, and *PDCD10* for our cases and compared it against the read depth across these exons in control samples from individuals without a diagnosis of CCM sequenced in the same batch as our cases. Coverage across these exons was visually compared, identifying CCM individuals with regions of comparatively lower depth indicative of exonic deletions. Chromosomal microarray analysis (CMA) and droplet-digital PCR (ddPCR) was used to confirm heterozygous exon deletions in these regions.

Chromosomal microarray analysis (CMA)

CMA was conducted on cases from family G (Table 1). Blood-derived DNA was processed using the Illumina (San Diego, California) Infinium workflow on the Global Screening Array (GSA)-24v1 or v2 (650 k SNP probes) BeadChips, as previously described [22]. The mean probe spacing was 6 kb and was validated for the detection of deletions and duplications to 200 kb and/or 20 probes. The minimum 20 probe threshold for CNV calling allowed for detection of CNVs to 1 kb in size. Sensitivity for mosaicism was validated to detect CNVs (whole chromosome, segmental chromosome (>10 Mb) and smaller CNVs (<10 Mb)) at 5%–10% mean variant allele fraction or greater. Copy number analysis was performed using the Illumina Karyostudio v1.4 or BioDiscovery NxClinical v6.0.

Droplet digital PCR (ddPCR)

Droplet-digital PCR (ddPCR) was used to validate small copy number variants detected via CMA and coverage analysis. We also ran assays to confirm specific variants using droplet-digital PCR (ddPCR) assays on DNA derived from formalin-fixed paraffin-embedded (FFPE) lesion tissues of 10 sporadic cases and 4 familial cases. TaqMan SNP Genotyping Assays from Thermo Fisher Scientific Inc were used to detect and quantify recurrent variants (Supplementary Table 1). Droplet generation, PCR cycling, and droplet reading were performed according to the manufacturer's recommendations (Bio-Rad).

RNA sequencing

RNA was extracted from 2.5 ml whole-blood collected in PAXgene RNA tubes using the PAXgene blood RNA kit as per the manufacturer's instructions (PreAnalytix PAXgene Blood RNA Kit Handbook). Libraries for two siblings with the *KRIT1* c.262 + 112A > G intronic variant predicted to create to a pseudoexon, and other individuals without the variant were prepared from blood-derived RNA using the Illumina Stranded Total RNA Prep with Ribo-Zero Plus kit and sequenced on the NovaSeq platform. Paired-end, 151 base pair reads were processed with the nf-core/rnaseq pipeline (v3.15.0) [24] using STAR (v2.7.10a) [25] to align reads to the GRCh38 reference genome (Ensembl release 113). Sashimi plots were generated using ggsashimi [26] to visualize RNA-sequencing reads that support the presence of the predicted pseudoexon comparing the MANE Select *KRIT1* transcript (ENST00000394505) with the predicted pseudoexon (chr7:92240748–92 240 880) added manually.

Supplementary data

Supplementary data is available at *HMG Journal* online.

Conflict of interest statement: Nothing to report.

Funding

We thank all patients and families for their participation in the study. This study was funded by the National Health and Medical Research Council Australia and Medical Research Future Fund Australia. It was supported by the Victorian State Government Operational Infrastructure Support and Australian Government

National Health and Medical Research Council (NHMRC) independent research Institute Infrastructure Support Scheme (IRIISS). P.P. was supported by an NHMRC Investigator Grant (2017651), The University of Melbourne, Monash University, the Austin Medical Research Foundation, and the Norman Beischer Medical Research Foundation. I.E.S was supported by an NHMRC Investigator Grant (1172897) and Synergy Grant (2010562). M.B. was supported by an NHMRC Investigator Grant (1195236). S.F.B. was supported by an Investigator Grant (1196637). S.F.B. and M.S.H. were supported by an NHMRC Project Grant (1129054). M.S.H. was supported by an NHMRC Ideas Grant (2012287). I.E.S., M.F.B., S.F.B., and M.S.H were supported by an MRFF Genomics Health Futures Mission Grant (2007707).

References

1. Rigamonti D, Hadley MN, Drayer BP. *et al.* Cerebral cavernous malformations. *N Engl J Med* 1988;**319**:343–347.
2. Al-Holou WN, O'Lynnger TM, Pandey AS. *et al.* Natural history and imaging prevalence of cavernous malformations in children and young adults: clinical article. *J Neurosurg: Pediatr* 2012;**9**:198–205.
3. Labauge P, Denier C, Bergametti F. *et al.* Genetics of cavernous angiomas. *Lancet Neurol* 2007;**6**:237–244.
4. Petersen TA, Morrison LA, Schrader RM. *et al.* Familial versus sporadic cavernous malformations: differences in developmental venous anomaly association and lesion phenotype. *Am J Neuroradiol* 2010;**31**:377–382.
5. Couteulx SL, Jung HH, Labauge P. *et al.* Truncating mutations in CCM1, encoding KRIT1, cause hereditary cavernous angiomas. *Nat Genet* 1999;**23**:189–193.
6. Sahoo T, Johnson EW, Thomas JW. *et al.* Mutations in the gene encoding KRIT1, a Krev-1/rap1a binding protein, cause cerebral cavernous malformations (CCM1). *Hum Mol Genet* 1999;**8**:2325–2333.
7. Liquori CL, Berg MJ, Siegel AM. *et al.* Mutations in a gene encoding a novel protein containing a Phosphotyrosine-binding domain cause type 2 cerebral cavernous malformations. *Am J Hum Genet* 2003;**73**:1459–1464.
8. Denier C, Goutagny S, Labauge P. *et al.* Mutations within the MGC4607 gene cause cerebral cavernous malformations. *Am J Hum Genet* 2004;**74**:326–337.
9. Bergametti F, Denier C, Labauge P. *et al.* Mutations within the programmed cell death 10 gene cause cerebral cavernous malformations. *Am J Hum Genet* 2005;**76**:42–51.
10. Weng J, Yang Y, Song D. *et al.* Somatic MAP3K3 mutation defines a subclass of cerebral cavernous malformation. *Am J Hum Genet* 2021;**108**:942–950.
11. Peyre M, Miyagishima D, Bielle F. *et al.* Somatic PIK3CA mutations in sporadic cerebral cavernous malformations. *N Engl J Med* 2021;**385**:996–1004.
12. Hong T, Xiao X, Ren J. *et al.* Somatic MAP3K3 and PIK3CA mutations in sporadic cerebral and spinal cord cavernous malformations. *Brain* 2021;**144**:2648–2658.
13. Ren AA, Snellings DA, Su YS. *et al.* PIK3CA and CCM mutations fuel cavernomas through a cancer-like mechanism. *Nature* 2021;**594**:271–276.
14. Akers AL, Johnson E, Steinberg GK. *et al.* Biallelic somatic and germline mutations in cerebral cavernous malformations (CCMs): evidence for a two-hit mechanism of CCM pathogenesis. *Hum Mol Genet* 2009;**18**:919–930.
15. Gault J, Shenkar R, Recksiek P. *et al.* Biallelic somatic and germ line CCM1 truncating mutations in a cerebral cavernous malformation lesion. *Stroke* 2005;**36**:872–874.
16. Wang J, Tang J, Yang Y. *et al.* Genotype-phenotype correlations in multiple lesions of familial cerebral cavernous malformations concerning phosphatidylinositol 3-kinase catalytic subunit alpha mutations. *Clin Transl Med* 2024;**14**:e1610.
17. Mondéjar R, Solano F, Rubio R. *et al.* Mutation prevalence of cerebral cavernous malformation genes in Spanish patients. *PLoS One* 2014;**9**:e86286.
18. Hirota K, Akagawa H, Kikuchi A. *et al.* KRIT1 mutations in three Japanese pedigrees with hereditary cavernous malformation. *Hum Genome Var* 2016;**3**:16032.
19. Zhou Z, Tang AT, Wong WY. *et al.* Cerebral cavernous malformations arise from endothelial gain of MEKK3–KLF2/4 signalling. *Nature* 2016;**532**:122–126.
20. Riant F, Cecillon M, Saugier-Veber P. *et al.* CCM molecular screening in a diagnosis context: novel unclassified variants leading to abnormal splicing and importance of large deletions. *Neurogenetics* 2013;**14**:133–141.
21. Zhou Z, Rawnsley DR, Goddard LM. *et al.* The cerebral cavernous malformation pathway controls cardiac development via regulation of endocardial MEKK3 Signaling and KLF expression. *Dev Cell* 2015;**32**:168–180.
22. Francis DI, Stark Z, Scheffer IE. *et al.* Comparing saliva and blood for the detection of mosaic genomic abnormalities that cause syndromic intellectual disability. *Eur J Hum Genet* 2023;**31**:521–525.
23. Engel J Jr, Van Ness PC, Rasmussen TB. *et al.* Outcome with respect to epileptic seizures. In: Engel J. Jr. (ed.), *Surgical treatment of the epilepsies*, 2nd edn. New York: Raven Press, 1993, 609–621.
24. Ewels PA, Peltzer A, Fillinger S. *et al.* The nf-core framework for community-curated bioinformatics pipelines. *Nat Biotechnol* 2020;**38**:276–278.
25. Dobin A, Davis CA, Schlesinger F. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 2013;**29**:15–21.
26. Garrido-Martín D, Palumbo E, Guigó R. *et al.* Ggsashimi: sashimi plot revised for browser- and annotation-independent splicing visualization. *PLoS Comput Biol* 2018;**14**:e1006360.

Human Molecular Genetics, 2025, Vol. 34, 15, 1286–1293
https://doi.org/10.1093/hmg/ddaf077
Original Article