De Novo Germline and Somatic Variants Convergently Promote Endothelial-to-Mesenchymal Transition in Simplex Brain Arteriovenous Malformation

Hao Li,* Yoonhee Nam,* Ran Huo,* Weilun Fu,* Blaobin Jiang©, Qluxia Zhou, Dong Song©, Yingxi Yang, Yuming Jiao, Jiancong Weng©, Zihan Yan, Lin Di, Jie Li©, Jie Wang, Hongyuan Xu, Shuo Wang, Jizong Zhao, Zilong Wen, Jiguang Wang©, Yong Cao©

RATIONALE: Brain arteriovenous malformations (bAVMs) are abnormal entanglement of blood vessels in brain, with direct connections from arteries to veins, lacking functional capillary bed. Although several somatic mutations were reported, the molecular mechanism and genetic disposition of bAVM remain poorly understood.

OBJECTIVE: We aim to identify transcriptional anomalies and critical functional pathways in bAVM lesions and explore their association with key de novo germline and somatic variants in bAVM patients.

METHODS AND RESULTS: We established a comprehensive bAVM dataset from 269 patients, by performing single-cell sequencing of 17 bAVM lesions, whole-exome sequencing of germline DNA from 60 case-unaffected-parental trios, and genomic/transcriptomic sequencing of 231 bAVM lesions. We found abnormal expression of endothelial and mesenchymal markers in bAVM at both bulk and single-cell level, which was validated by flow cytometric analysis and immunofluorescence staining, suggesting an involvement of endothelial-to-mesenchymal transition (EndMT) process in AVM (arteriovenous-malformation-like). Using data from the 60 trios, we identified nonsynonymous de novo germline mutations affecting 46 genes, including EXPH5 (detected in 2 independent cases), and vessel-related genes, such as EPAS1 and ENG. Interestingly, knockdown of epas1 in zebrafish embryo showed AVM-like phenotype exclusively in brain. Subsequent computational and experimental analyses demonstrated that expression of genes affected by de novo germline mutations was enriched in vascular cell types and was involved in EndMT-relevant behaviors including cell migration, angiogenesis, and cell marker transition. Moreover, we detected somatic KRAS mutations in 129 of 179 (72%) cases and showed that KRAS mutations were associated with bleeding as the first symptom (P=0.0072). Following experimental studies demonstrated that KRAS mutations independently regulated EndMT features, consolidating the involvement of EndMT in this disease. Lastly, we showed that lovastatin reversed EndMT features in vitro and ex vivo.

CONCLUSIONS: Our results suggest the convergent role of de novo germline mutations and somatic mutations in regulating EndMT in bAVM and provided a potential therapeutic option.

GRAPHIC ABSTRACT: An online graphic abstract is available for this article.

Key Words: computational biology | germ-line mutation | high-throughput nucleotide sequencing | intracranial arteriovenous malformations

In This Issue, see p 821 | Meet the First Author, see p 822
Novelty and Significance

What Is Known?
• Germline mutations such as ENG, SMAD4, and ACVRL1 are associated with hereditary vascular malformation syndromes.
• Somatic mutations in KRAS and BRAF genes have been identified in endothelial cells of brain arteriovenous malformations (bAVMs) from patients with no family history.
• Activating somatic KRAS mutations in the endothelium leads to AVM-like (arteriovenous-malformation-like) phenotypes in mouse and zebrafish models.

What New Information Does This Article Contribute?
• Whole-exome sequencing of case-unaffected-parental bAVM trios revealed de novo germline mutations in ENG, JUP, EXPH5, and EPAS1, and subsequent in vivo study showed that epas1-morpholino induced brain-specific vascular malformation in zebrafish.
• Comprehensive multiomics and experimental studies demonstrated that de novo germline and somatic variants convergently promote endothelial-to-mesenchymal transition in simplex bAVM, and lovastatin treatment reverses the features of endothelial-to-mesenchymal transition in vitro and ex vivo.
• Somatic KRAS mutation frequently occurs in bAVMs with bleeding as the first symptom.

Nonstandard Abbreviations and Acronyms

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AVM</td>
<td>arteriovenous malformation</td>
</tr>
<tr>
<td>bAVM</td>
<td>brain arteriovenous malformation</td>
</tr>
<tr>
<td>DNM</td>
<td>de novo germline mutation</td>
</tr>
<tr>
<td>EC</td>
<td>endothelial cell</td>
</tr>
<tr>
<td>EndMT</td>
<td>endothelial-to-mesenchymal transition</td>
</tr>
<tr>
<td>FDR</td>
<td>false discovery rate</td>
</tr>
<tr>
<td>HUVEC</td>
<td>human umbilical vein endothelial cell</td>
</tr>
<tr>
<td>NES</td>
<td>normalized enrichment score</td>
</tr>
<tr>
<td>RNA-seq</td>
<td>RNA sequencing</td>
</tr>
<tr>
<td>SA</td>
<td>superficial temporal artery</td>
</tr>
<tr>
<td>SNP</td>
<td>single-nucleotide polymorphism</td>
</tr>
<tr>
<td>VAF</td>
<td>variant allele frequency</td>
</tr>
<tr>
<td>WES</td>
<td>whole-exome sequencing</td>
</tr>
<tr>
<td>WT</td>
<td>wild type</td>
</tr>
</tbody>
</table>

Brain arteriovenous malformations (bAVMs) are abnormal entanglement of blood vessels in brain, responsible for 3% to 4% of intracranial hemorrhage in children. The disease lesion of AVM (arteriovenous-malformation-like) is called nidus, which generates direct connections from the arteries to the veins lacking a functional capillary bed. About 95% of bAVM cases have no known family history, and their pathogenesis is unclear. Recent studies found somatic KRAS mutations in bAVMs with frequencies ranging from 28.5% to 76.2% across different studies. In these studies, the enrolled number of AVM cases was typically small, limiting the discovery of this mutation’s clinical associations. Subsequent functional studies proposed that bAVMs develop as a consequence of somatic mutations in endothelial cells (ECs), and in vivo experiments demonstrated that endothelial-specific gain of these mutations is sufficient to induce bAVM-like phenomenon in mice and zebrafish models. While the driving role of KRAS in bAVM has been well investigated, its clinical implications and molecular mechanisms are still not fully understood. Moreover, genetic factors of KRAS-WT (wild type) cases remain unexplored.

Anomalies of ECs, including differentiation, polarization, and mesenchymal behaviors, were observed in bAVM lesions, and manipulated genetic mutation in ECs caused bAVM-like phenotype in animal models. However, critical genetic determination factors and key transcripational pathways of the AVM genesis process are still unclear. Great efforts have been devoted to study...
the disease lesion of bAVMs through RNA sequencing (RNA-seq) that implied the association of KRAS/MAPK (mitogen-activated protein kinase) signaling, inflammation, and impaired vascular wall integrity with bAVM lesions. Nonetheless, these approaches were limited to characterizing one cell type or averaged gene expressions of the whole tissue composed of various types of cells, lacking the resolution to decouple the complexity of this disease. In contrast, single-cell technology that allows detailed analysis of the heterogeneity and molecular changes of individual cells may lead to a better understanding of the disease mechanism.

In this study, we aim to establish a comprehensive dataset of bAVM using multiomics sequencing. On one hand, enlightened by the study of de novo germline mutations (DNMs) in other simplex human genetic disorders and considering that bAVM patients are typically young, we hypothesized that DNMs might also contribute to the pathogenesis of simplex bAVMs. Thus, we identify the anomalies of bAVM ECs and investigate the underlying genetic causes from both somatic and germline alterations. On the other hand, we perform both bulk and single-cell sequencing to characterize the transcriptional program of AVM ECs, based on which we report the key process driving AVM genesis.

**METHODS**

**Data Availability**

The raw data of this study are available from the Genome Sequence Archive with accession number HRA01250 upon reasonable request. For the software and research materials, please see the Major Resources Table in the Data Supplement. Details of Materials and Methods are described in the Data Supplement.

**RESULTS**

**Patients**

A total of 269 simplex bAVM patients from January 2016 to May 2020 were included in this study (Table I in the Data Supplement). Of those, 239 had fresh-frozen or fresh-resected nidus tissue and 154 had matched blood. We performed bulk RNA-seq on the bAVM nidus of 66 patients and single-cell RNA-seq on 17 patients to profile transcriptomic pattern of cells in bAVM lesions. Whole-exome sequencing (WES) was performed on germline DNA of 60 patients together with the matched unaffected parental blood for the identification of DNMs. In addition, 219 nidus tissues were used for WES, amplicon sequencing, and droplet digital polymerase chain reaction to identify somatic mutations (Figure 1A through 1C). Flow cytometry, immunofluorescence staining, electron microscopy, and patient-derived EC culture were also carried out to characterize the nidus. For control purpose, non-nidus vessels were obtained from 8 bAVMs during standard procedure of craniotomies and bAVM resection. These vessels included superficial temporal arteries (SAs), scalp SAs, feeding arteries, and draining veins.

As patient-level control, we analyzed brain vessels from 11 patients undergoing temporal lobe resection for non-AVM epilepsy and SAs from 4 non-AVM nonepilepsy patients during standard neurosurgical procedures like pterional craniotomy for aneurysm clipping. Sample details are provided in Table I in the Data Supplement.

**Endothelial-to-Mesenchymal Transition in BAVMs**

We performed bulk RNA-seq on nidus from 66 bAVM cases and control vessels from 7 non-bAVM epilepsy patients. Gene set enrichment analysis found significantly lower expression of brain endothelial markers in bAVMs than control (Figure 2A). Notably, ICAM2, FLT1, and KDR were among the most downexpressed genes (Figure 2B). To explore bAVM ECs at a single-cell level, we performed single-cell RNA-seq on nidus from 17 patients. A total of 13,328 high-quality single cells were clustered and visualized (Figure 2C; Figure IA and IB in the Data Supplement). We detected 3 distinct EC clusters and 1 mural/fibroblast cluster. Unexpected expression of mesenchymal markers was observed in EC1-EC3, in contrast to ECs from non-AVM human brain (Figure 2D; Figure IC and ID in the Data Supplement). Other than the ECs, we also investigated the abnormal subclusters in mural and fibroblast cells. Particularly, we merged AVM EC and mural/fibroblast clusters (EC1-EC3 and mural/fibroblast in Figure 2) with EC (control-EC) and pericyte cells (control-pericytes) from the control into one uniform manifold approximation and projection (Figure IIA in the Data Supplement). Control-EC cluster majorly overlapped with AVM-EC1, with a little overlap with the AVM-EC2 cluster. Control-pericyte cluster overlapped with mural/fibroblast cluster from AVM lesions. However, within the mural/fibroblast cluster, we identified a subcluster of cells that did not overlap with control-pericyte (Figure IIB in the Data Supplement). We termed this subcluster AVM-specific mesenchymal cells. Interestingly, the AVM-specific mesenchymal cell cluster simultaneously expressed CLDN5 (EC marker), MYL9 (smooth muscle cell marker), PDGFRB, and RGS5 (pericyte markers), implying that the cluster may represent a novel group of cells carrying expression features of both endothelial and mesenchymal cells (Figure IIC in the Data Supplement). Further pathway enrichment analysis showed that expression level of the genes associated with blood vessel diameter maintenance, regulation of tube diameter, and muscle contraction pathways were significantly higher in the AVM-specific mesenchymal cell cluster (Figure IID and IIE in the Data Supplement). This interesting cluster provides new evidence of endothelial-to-mesenchymal transition (EndMT) in bAVM. Collectively,
we reasoned that bAVM might undergo the process of EndMT.\textsuperscript{19,20}

To verify the EndMT hypothesis, we utilized fluorescence-activated cell sorting to purify cells using CD31 (cluster of differentiation 31; an endothelial marker) and \(\alpha\)-SMA (alpha-smooth muscle actin; a commonly used mesenchymal marker), in independent bAVM cases compared with non-AVM epilepsy control. Interestingly, we found that \(\alpha\)-SMA\(^+\) CD31\(^+\) CD45\(^-\) cells constituted on average 24.8\% (SEM, \(\pm\) 6.1; \(n = 9\)) among all CD31\(^+\) CD45\(^-\) (endothelial) cells in bAVM nidus, significantly higher than the control (mean\(\pm\)SEM, 3.8\(\pm\)1.3; \(n = 4\); Figure 2E; Figure III in the Data Supplement; Mann-Whitney nonparametric test, \(P = 0.0028\)).

Considering that the artery/vein ratio might be a confounding factor introducing bias to the current results, we next performed multicolor immunofluorescent staining to examine marker genes at protein level including \(\alpha\)-SMA, SLUG, (SNAI2, snail family transcriptional repressor 2), TAGLN (Transgelin), S100A4, and MYL9 in bAVM endothelium in comparison to SAs. The mesenchymal markers were remarkably stained in the endothelium lining the lumen of bAVM lesions but were seldomly expressed in control endothelium, further demonstrating the occurrence of EndMT in lesions of human bAVM (Figure 2F).

**Detection of DNMs**

Since germline variants in Hereditary Hemorrhagic Telangiectasia (HHT) genes could induce hereditary syndromes that were hallmarked by bAVM,\textsuperscript{21,22} we hypothesized that simplex bAVM patients without family history might develop as the consequence of DNMs that occurred in germ cells or during early embryonic development. To test this hypothesis, deep WES was performed for 60 case-unaffected-parental trios to detect DNMs in bAVM patients. The sequencing depth of these samples was summarized in Figure IVA in the Data Supplement. Due to intrinsic sampling bias (ie, parental DNAs of younger cases were easier to be acquired), the average AVM diagnosis age of patients enrolled for DNM study was younger than the rest of the enrolled cases in this study (Figure IVB in the Data Supplement). Furthermore, the DenovoSAVI pipeline was developed and used for DNM identification. As a result, a total of 63 DNMs were detected (with 49 confirmed by mass spectrometry or Sanger sequencing), averaging to 1.05 DNMs (63/60) per exome per generation (Figure 3A; Figure IVC and IVD in the Data Supplement; Table II in the Data Supplement). This number is compatible with the mutation rate...
estimated by other methods (ranging from $1.1 \times 10^{-8}$ to $2.5 \times 10^{-8}$ per site per generation), suggesting that AVM might not be driven by higher germline mutation rate. However, the missense-to-synonymous ratio was 3.75 (45/12), indicating a positive selection compared with neutral expectation (2.23 missense mutations).
Figure 3. Genomic sequencing study of case-unaffected parental trios uncovered the landscape of de novo germline mutations (DNMs) in brain arteriovenous malformations. 

A. Distribution of DNMs per patient in the 60 case-unaffected parent trios. 
B. Distribution of alterations in NH—single-nucleotide polymorphism (SNP) and DNMs. 
C. Gene ontology pathway analysis on 46 DNM genes. 
D. Enrichment of 46 DNM genes in blood vessel from GTEx database compared with random (all genes). 
E. Expression enrichment of 46 DNM genes in each cell type from the Panglao database. Length of bar represents enrichment P and those <0.05 are colored in red. 
F. Fisher exact test contingency table for association of 16 DNMs (Table) and endothelial-to-mesenchymal transition (EndMT). Total number of coding genes used for DNM detection is shown.

G. Representative confocal images for the brain of epas1 morpholinos (MOs) injected Tg(kdrl:GFP;globin:DsRedx) embryos and uninjected control (UIC). Arrows indicate kdrl:GFP+ vessel ECs with altered morphology, while the cranial hemorrhage indicated by accumulated globin:DsRedx+ red blood cells is pointed out by the asterisk; dorsal views. Scale bars, 50 μm.

H. Quantification of percentage of embryos with indicated phenotype in each condition.

I. Representative confocal images for the trunk of epas1 MOs injected embryos and UIC; lateral views. Scale bars, 100 μm. All images taken at 60 to 70 hpf. Data are representative of 3 independent experiments. n/n=number of embryos with representative phenotype/total number of embryos examined. AVM indicates arteriovenous malformation; GFP, green fluorescent protein; GTEx, the genotype-tissue expression; and VM, vascular malformation.
expected to occur per synonymous mutation). In addition, we calculated the ratio of different variants in new homozygous single-nucleotide polymorphism (SNP; ie, heterozygous in parents but homozygous in proband) and found that both missense and nonsense variants were remarkably enriched in DNMs compared with that in new homozygous SNP for both common and non-common SNPs. These results collectively implied the potential involvement of DNM in AVM's disease genesis (Figure 3B).

Characterization and Prioritization of DNMs

In the following analyses, synonymous mutations and common SNPs (allele frequency ≥1% in 1000Genomes population) were excluded, leaving 48 DNMs affecting 46 genes (Table III in the Data Supplement). Notably, EXPH5 gene was altered in 2 independent patients, and other genes were unique to each proband. Four mutations were nonsense or splicing site mutations with higher functional impact, including EXPH5 (R1111*), JUP (Q9*), CHEK2 (c.1008+2T>G), and ENG (c.360+1G>A). JUP encodes plakoglobin—a protein closely related to β-catenin that links desmosomal cadherins to intermediate filaments and can also substitute for β-catenin in adherens junctions;27 ENG—a coreceptor for TGF-β (transforming growth factor beta) in EC—is reported contributing to the pathology of HHT1 (hereditary hemorrhagic telangiectasia type 1), which is featured by bAVMs.28 Hence, the patient with ENG (c.360+1G>A) variant is likely a de novo HHT case without family history. This ENG splice site mutation was reported causing a reduction in the level of functional endoglin.29 EXPH5 encodes a ras-related protein Rab27B effector protein Slac2-b. Mutations in this gene have been associated with epidermolysis bullosa, an inherited skin fragility disorder, and Slac2-b knockdown reduced keratinocytes adhesion.30 Among the 48 DNMs, 23 variants were found with dbSNP reference SNP (rs) numbers, but based on population allele frequencies from public database,31,32 all of them were rare SNPs (allele frequencies below 0.1% in population). According to the genotyping validation (mass spectrometry or Sanger sequencing) performed on these 23 SNP-numbered DNMs, 18 of 20 successfully genotyped variants were confirmed to be de novo (Table III in the Data Supplement).

We then performed gene ontology enrichment analysis and found these 46 genes were enriched in cell adhesion, cell migration, and cell junction, all of which relate to EndMT AVM phenotype (Figure 3C). To investigate tissue expression of these DNMs genes, we analyzed gene expression profiles of 11 688 samples from a wide range of tissue types and defined tissue-specific gene clusters for 33 tissue types (Figure V in the Data Supplement). Interestingly, 6 genes including EPAS1 (related to vascular remodeling),26 DACT1, EMILIN1 (functions in vascular cell maintenance and inhibits TGF-β signaling),24 ENG, GAS5, and COL1A1 were in the vessel-specific gene expression cluster, relatively higher than random (Figure 3D; binomial test \( P=0.060 \)). Next, we performed cell type enrichment analysis (Figure VIA in the Data Supplement) and found most significant involvement of cell types constituting blood vessel and surrounding tissues, including ECs, pericytes, smooth muscle cells, and fibroblasts using both adult and nonadult single-cell datasets (Figure 3E; Figure VIB and VIC in the Data Supplement). These evidences collectively demonstrated that the gene list we identified from the 60 trios contained an elevated number of bAVM-associated genes.

In addition, we used in silico mutagenesis and protein homology modeling to construct a series of 3-dimensional structural models of the proteins in WT and mutant type. Limited by the availability of the protein structure database and accuracy of homology remodeling, this analysis was performed in 7 variants. We observed that most of the analyzed DNMs occurred on functional sites of proteins, and these mutations often changed local charge distributions that might affect protein-protein interactions, protein-ligand interactions, protein-RNA interactions, and protein-DNA interactions (Figure VII in the Data Supplement; Table III in the Data Supplement).

To identify bAVM-relevant DNM genes, we considered both mutational impact and the above enrichment analyses and eventually prioritized 16 genes in the Table as candidate bAVM-associated genes. Interestingly, we found that patients carrying DNMs in these genes had a higher chance to have epilepsy as the first symptom (Fisher exact test, \( P=0.023 \); Figure VIII in the Data Supplement). Remarkably, 3 (ie, JUP, ENG, and COL1A1) of the 16 genes were included in the known EndMT gene list,36 implying that DNMs might regulate the process of EndMT in bAVMs (Fisher exact test, \( P=6.5×10^{-5} \); Figure 3F).

Knockdown of EPAS1 in Zebrafish Alters Brain Vascular Morphology and Causes Cranial Hemorrhage

We then performed in vivo experiments on zebrafish (Danio Rerio) by injecting morpholino against the selected candidate gene into Tg(kdrl:GFP;globin:DsRedx) zebrafish embryos and then evaluated the corresponding phenotype in morphants at 60 to 70 hours post-fertilization by visualizing the green fluorescent protein (GFP)-positive ECs and DsRedx-positive red blood cells.

In particular, we set out to study a DNM gene that was vessel specific in our tissue-level expression analysis and reported to be related to vascular remodeling in mouse, EPAS1. EPAS1 has 2 orthologs in zebrafish, epas1a and epas1b. Previous studies have shown unaffected trunk vessel formation in zebrafish epas1a and epas1b mutants.38 We reinvestigated the vasculature by coinjecting epas1a and epas1b morpholinos into the embryos of Tg(kdrl:GFP;globin:DsRedx). Strikingly, while epas1a morphants indeed displayed normal trunk vasculature, their
brains showed robust and significant vascular malformations (Figure 3G through 3I), in which reticulated vasculature was formed (46/48) and some were accompanied with severe bleeding (5/48). These observations indicate that knockdown of the \( \text{EPAS1} \) gene may directly act to drive bAVM-like brain vascular malformations in vivo.

### Association Between DNMs and ENDMT

To explore the association between DNMs and EndMT, we manipulated human umbilical vein endothelial cells (HUVECs) to study the function of \( \text{JUP} \), \( \text{ENG} \), and \( \text{EXPH5} \), selected endothelial genes with high-impact DNMs (Figure IX in the Data Supplement; Table IV in the Data Supplement). Gene set enrichment analysis of RNA-seq in \( \text{JUP} \) and \( \text{ENG} \) knockdown HUVECs confirmed the occurrence of EndMT using not only the known upexpressed-in-EndMT gene set (\( \text{JUP} \): normalized enrichment score [NES], 3.004; false discovery rate [FDR], 0.010; \( \text{ENG} \): NES, 3.133; FDR, <0.001) but also the known downexpressed-in-EndMT gene set (\( \text{JUP} \): NES, −4.303; FDR, <0.001; \( \text{ENG} \): NES, −4.552; FDR, <0.001; Figure 4A; Figures X and XI in the Data Supplement). For \( \text{EXPH5} \) knockdown HUVECs, we found an increase in upexpressed-in-EndMT genes (NES, 2.798; FDR, 0.068) but not the decrease in downexpressed-in-EndMT genes, probably indicating a partial EndMT in \( \text{EXPH5} \) knockdown HUVECs (Figure XIE through XIG in the Data Supplement).

We further performed immunofluorescent staining and observed increased expression of mesenchymal markers (\( \alpha \)-SMA and \( \text{TAGLN} \) [transgelin]) and decreased expression of CD31 in HUVECs with \( \text{JUP} \) and \( \text{ENG} \) knockdown (Figure 4B; Figure XII in the Data Supplement). In consistency with RNA-seq, we found \( \text{EXPH5} \) knockdown HUVECs showed increased level of \( \alpha \)-SMA and \( \text{TAGLN} \) but no significant change in CD31. We also observed enhanced cell migration in HUVECs with \( \text{JUP} \), \( \text{ENG} \), and \( \text{EXPH5} \) knockdown by wound healing analysis (Figure 4C). Moreover, fibrin gel bead assay showed that knockdown of DNMs genes inhibited vascular-like structure formation in HUVECs (Figure 4D). Altogether, these evidences confirmed the involvement of DNMs genes in regulating EndMT in vitro.

### KRAS Mutation and ENDMT

Recent studies identified somatic \( \text{KRAS} \) and \( \text{BRAF} \) mutations in bAVM\(^2\)–7 and observed activation of MAPK-ERK (extracellular signal regulated kinase) signaling pathway, as well as expression changes of EndMT-related genes in \( \text{KRASG12V} \)-expressing HUVECs.\(^2\) To explore the underlying association among somatic alterations, DNMs and EndMT, we examined bAVM nidus from 219 patients for detecting somatic mutations. We first used WES from 75 nidus with matched blood to identify somatic variant candidates. We found \( \text{KRASG12D} \), \( \text{G12V} \), \( \text{G12A} \), and \( \text{Q61H} \) mutations with variant allele frequency (VAF) ranging from 1.87% to 10.26% (Figure XIII in the Data Supplement). Although median coverage of AVM WES was >250×, coverage for \( \text{KRAS} \) and \( \text{BRAF} \) hot spots was below 120×, which may have limited the detection of low-frequency mutations (Figure XIV in the Data Supplement). We, therefore, performed amplicon sequencing (median coverage, 13 000×) targeting \( \text{KRASG12V} \) and \( \text{c34G} \) in 179 bAVM nidus and 106 blood. A model using \( \beta \)-binomial distribution was developed to identify variants with low allelic frequency (Figure XV in the Data Supplement). Using this method,
KRAS mutations were identified in 69.8% (125/179) of the nidus samples with the lowest VAF $10^{-4}$. Yet, 2.8% (3/106) of the blood samples were KRAS mutant, implying potential false discovery rate (Figure 5A; Table V in the Data Supplement).

To confirm the results via another approach, we performed droplet digital polymerase chain reaction on 61 bAVMs and 30 blood samples on 10 selected variant candidates (Figure XIIIB in the Data Supplement). While no variants were detected in blood, 41 of 42 KRAS variants (97.6%) detected by amplicon sequencing were confirmed by droplet digital polymerase chain reaction with high VAF correlation (Figure XVI in the Data Supplement; Pearson r, 0.952). Moreover, 2 additional KRAS variants (G12V and G12S) were separately detected in patients A060 and A226. In addition, 1 BRAF V600E (patient A228) and 1 KRAS Q61H variant (patient A123) were detected. Collectively, 72% (129/179) of the patients were KRAS mutant via WES, amplicon sequencing, or droplet digital polymerase chain reaction. This result was comparable to the previously reported incidence in the Canadian cohort.2

To explore the prevalence of KRAS mutations in non-nidus regions, SA (n=4), scalp SAs (n=3), feeding arteries (n=7), and draining veins (n=5) from 7 patients were sequenced. One patient (A084) was KRAS G12D mutant in nidus (VAF, 6.6%), feeding arteries (VAF, 0.24%), and draining veins (VAF, 0.69%) but KRAS-WT in SA (Figure 5B). All other non-nidus samples were KRAS-WT (Figure XVIIA and XVIIIB in the Data Supplement). Our data suggested that KRAS mutations are enriched in the nidus and less present in nearby vessels but are unlikely occurring at distant vessels. Moreover,
the large sample set in our study allowed us to explore clinical and phenotypic relevance of the somatic KRAS mutations. Interestingly, we found that KRAS-mutant cases had higher incidence of having bleeding as the first symptom (Fisher exact test, \( P = 0.0072 \); Figure 5C).

No association was observed between somatic KRAS mutation and preoperative radiotherapy (Fisher exact test, \( P = 0.3 \); Figure XVIIC in the Data Supplement).

Furthermore, recent study reported that 2 patients with somatic BRAF mutation demonstrated relatively older ages.\(^7\) Similarly, our A228 patient with somatic BRAF V600E mutation (VAF, 1.03%) was diagnosed at age 43, higher than the median age 23 of all the bAVM participants in our study.

To explore the association between somatic KRAS mutation and DNMs, we investigated 43 cases that were enrolled in both DNM and somatic mutation study. Figure 5D showed the landscape of the 60 case-unaffected parent trios. Top, Clinical features and somatic KRAS mutation. KRAS mutation is determined by WES, amplicon-seq, or ddPCR of bulk bAVM. Middle, Selected de novo germline mutation (DNM) genes. Color indicates the protein impact of the mutation. Bottom, Number of noncommon and nonsynonymous DNMs in each patient. Right, DNM genes enriched in gene ontology (GO) pathways and enriched in blood vessel from tissue-specific analysis (GTEx). High (impact), nonsense, or splicing site mutations; moderate-SD, missense mutation and predicted to be deleterious by SIFT; moderate, missense mutation. DV indicates draining vein; Dx, diagnosis; FA, feeding artery; GTEx, the genotype-tissue expression; N/A, not available; and SSA, scalp superficial temporal artery.
significant trend was observed between somatic \( \text{KRAS} \) mutation and the DNMs.

We generated HUVECs overexpressing \( \text{KRAS}^{4B^{G12D}} \) to explore potential role of \( \text{KRAS} \) in regulating EndMT. As expected, elevated phosphorylated ERK was detected in manipulated HUVECs (Figure XVIII in the Data Supplement). We also observed increased level of mesenchymal markers and suppression of endothelial markers in \( \text{KRAS}^{4B^{G12D}} \)-overexpressing HUVECs by Western blot (Figure 6A), immunofluorescence staining (Figure 6B; Figure XVIIIIB in the Data Supplement), and qPCR (quantitative reverse transcription) (Figure XVIIIC in the Data Supplement; Table VI in the Data Supplement).

Morphologically, the cobblestone appearance of ECs was transformed into an elongated and spindle-like shape in immunofluorescence staining by \( \text{KRAS} \) activation, consistent with reported EndMT observations\(^{40} \) (Figure 6B). Moreover, \( \text{KRAS}^{4B^{G12D}} \)-HUVECs acquired an enhanced migratory ability (Figure 6C). We then blocked the MAPK-ERK pathway utilizing the MAPK-specific inhibitor U0126 and observed that \( \text{KRAS}^{4B^{G12D}} \)-driven changes were rescued, confirming the role of MAPK-ERK signaling in mediating EndMT in HUVECs (Figure 6A through 6C). Furthermore, we analyzed \( \text{KRAS}^{G12D} \)-induced morphogenesis in a bead-based morphogenesis assay.\(^{41} \) Control ECs sprouted numerous tubular vascular-like structures mimicking angiogenesis, while \( \text{KRAS}^{4B^{G12D}} \)-HUVECs exhibited sheet-like structures, rather than normal vascular elongation and lumogenesis (Figure 6D). These results collectively confirmed the role of activating \( \text{KRAS} \) mutation in driving EndMT in vitro.

**Lovastatin Attenuates EndMT In Vitro and Ex Vivo**

Lovastatin—an agent widely used in treating lipid disorders—may inhibit EndMT in glomerular ECs.\(^{42} \) To explore...
whether lovastatin reverses EndMT induced by KRAS mutation in ECs, we treated HUVECs overexpressing mutant KRAS with lovastatin and examined EndMT markers by immunofluorescent staining. Induced mesenchymal marker α-SMA was abrogated and the suppressed endothelial marker CD31 was upregulated upon treatment (Figure 7A; Figure XIXA in the Data Supplement). This result was further confirmed by RT-qPCR and Western blot (Figure 7B; Figure XIXB in the Data Supplement). Moreover, we demonstrated that lovastatin reversed the enhanced migration induced by KRAS activation (Figure 7C; Figure XIXC in the Data Supplement). Similarly, we examined whether Lovastatin could reverse EndMT in the JUP, EXPH5, and ENG knockdown HUVECs. As exhibited in Figure 7D, when treated with lovastatin, the mesenchymal markers α-SMA, TAGLN, and SLUG were substantially decreased in all 3 gene knockdown HUVECs, while CD31 was moderately increased in JUP and ENG knockdown HUVECs (Figure XIXD and XIXE in the Data Supplement). Meanwhile, enhanced migration induced by 3 gene knockdowns was reversed by lovastatin treatment (Figure 7E; Figure XIXC in the Data Supplement). These results collectively demonstrated the role of lovastatin in reversing EndMT induced by either KRAS mutations or DNMs in vitro.

Next, we explored whether lovastatin could inhibit EndMT in patient-derived ECs. Isolated bAVM ECs cultured in endothelial medium expressed α-SMA, SLUG (SNAI2, snail family transcriptional repressor 2), and TAGLN. When treated with lovastatin, α-SMA, TAGLN, and SLUG expression was decreased, implying a suppression of EndMT in those cells (Figure 7F; Figure XX in the Data Supplement).

**DISCUSSION**

Here, we performed DNA sequencing on 219 bAVMs and blood samples of 60 case-unaffected-parental trios. The largest bAVM cohort in the Chinese population enabled the discovery of association between somatic KRAS mutation (in 72% cases) and bleeding as the first symptom. The discrepancy in KRAS mutation frequency of different studies may be due to different population group, clinical practice, sample status, and the detection techniques.

In addition, we identified 63 DNMs from 60 case-unaffected-parental trios with an average 1.05 (range from 0 to 5) mutations per patient. Previous studies have reported the increase in the number of DNMs with parent’s age at conception, but probably due to the limited number of samples, no significant association has been observed in our cohort. Furthermore, no obvious patterns were detected in patients who were carrying over 1 DN姆. We note that most of the DNMs are not functional in the disease process. Therefore, only 48 nonsynonymous noncommon DNMs that affect 46 genes were selected for further investigation. We propose that whereas not all variants were disease relevant, some of these mutations would increase the risk or predispose bAVM in individuals. To interrogate the involvement of these genes in bAVM pathogenesis, we pooled these genes together to perform a gene set enrichment type of analysis. We found cell adhesion, cell migration, and cell junction terms enriched in GO analysis, 6 DNMs significantly enriched in vessel-specific gene expression cluster in tissue-level analysis, and significantly enriched expression of the DNMs genes in blood vessel–related cell types through public single-cell database. These imply that a significant number of the DNMs genes are potentially associated with bAVM pathogenesis. Based on the variant impact on the protein, we identified crucial DNMs such as ENG (a splicing variant), JUP (stop codon), and EXPH5 (recurrence in 2 independent cases with 1 stop codon and 1 missense variant) and showed that alterations in these genes in vitro were associated with EndMT-like change. Moreover, knockdown of epsa1, 1 of the 6 DNMs genes in vessel-specific gene cluster, driven brain vascular malformation and induced cranial hemorrhage in zebrafish. However, more comprehensive in vivo models such as variant heterozygous mouse should be utilized to completely illuminate the effects of the DNMs and corresponding mechanisms in future research.

Patients with DNMs in the prioritized 16 genes were found to be associated with epilepsy as the first symptom. The potential mechanism regarding this observation may need further investigation. However, it is believed that the DNMs occur early in the embryonic development and the cerebrovascular malformations caused by early embryonic development mutations can affect the blood supply to the cortex chronically leading to cortical dysgenesis, which often results in epilepsy. Moreover, patient A095 had 2 in phase mutations or DNMs in the DLGAP3 gene, where substitution at 2 adjacent nucleotides occurred resulting in the change of 2 amino acids (Figure XXI in the Data Supplement). Previous study has generated SAPAP3(DLGAP3)+mutant mice that showed deficiency in corticostriatal synapses and led to obsessive compulsive disorder–like behaviors. Further follow-up will be needed to study the effects of these 2 variants.

We observed EndMT in simplex bAVM and propose that the identified DNMs and somatic KRAS mutations may convergently provoke EndMT, but additional studies would be needed to understand potential cross talk between these alterations. EndMT is a collective course in which gradients (ie, partial versus more complete EndMT, reversible, transient) and extent of reduced endothelial markers versus increased mesenchymal markers should be considered. Different patterns of EndMT were reported in various physiological and pathological processes.

In bAVMs, by utilizing single-cell analysis, we found 3 subgroups of ECs, and each group expressed different pattern of mesenchymal markers, indicating a complicated EndMT in bAVM endothelium. Further investigation would be needed to decipher the heterogeneity in vivo. Interestingly, we found KRAS-mutant bAVMs were prone to rupture. We speculated this correlation could be explained by the severe EndMT induced by the KRAS mutation. It is also in consistency with our previous finding that a more prominent EndMT was correlated with microhemorrhage of bAVMs.
Figure 7. In vitro and ex vivo studies demonstrated repurposing lovastatin that targets endothelial-to-mesenchymal transition as a potential treatment of brain arteriovenous malformations (bAVMs).

A. Immunofluorescence staining of CD31 and α-SMA in human umbilical vein endothelial cells (HUVECs) overexpressing KRAS4B<sup>G12D</sup> and HUVECs overexpressing KRAS4B<sup>G12D</sup> treated with lovastatin. Scale bar, 100 μm. B. Expression of endothelial and mesenchymal markers in HUVECs overexpressing KRAS4B<sup>G12D</sup> HUVECs overexpressing KRAS4B<sup>G12D</sup> treated with lovastatin, and control as shown by real-time PCR (n=3). The P values of * mark are the following: KRAS vs CTRL (PECAM1: P=0.0013; CDH5: P=6.8×10<sup>−4</sup>; VWF: P=9.6×10<sup>−6</sup>; ACTA2: P=0.0011; SNAI2: P=0.0050; TAGLN: P=2.3×10<sup>−4</sup>); KRAS vs Kras+lovastatin (PECAM1: P=4.9×10<sup>−5</sup>; CDH5: P=0.0018; VWF: P=0.0023; ACTA2: P=6.2×10<sup>−4</sup>; SNAI2: P=0.040; TAGLN: P=0.011). Analyzed by unpaired t-test. The T bars represent SD.

C. Statistical analysis of reduced area in HUVECs overexpressing KRAS4B<sup>G12D</sup> treated with lovastatin (KRAS) and HUVECs overexpressing KRAS4B<sup>G12D</sup> treated with lovastatin (KRAS+lovastatin) and HUVECs transfected with control adenovirus (CTRL) in wound healing assays. The T bars represent SD. P by Mann-Whitney nonparametric test (n=5).

D. Expression of endothelial marker CD31 and mesenchymal marker α-SMA in HUVECs with JUP, ENG, and EXPH5 knockdown and with lovastatin treatment by immunofluorescence staining. Scale bar, 100 μm. E. Statistical analysis of reduced area in HUVECs with JUP, ENG, and EXPH5 knockdown (siJUP/ENG/EXPH5), knockdown HUVECs treated with lovastatin (siJUP/ENG/EXPH5+lovastatin), and HUVECs transfected with control siRNA (CTRL) in wound healing assays. The T bars represent SD. P by Mann-Whitney nonparametric test (n=5).

F. Representative immunofluorescence staining images of primary endothelial cells (ECs) derived from bAVM. The overexpressed mesenchymal markers (α-SMA, SLUG, and TAGLN) are observed in these cells. Lovastatin treatment inhibits the expression of these mesenchymal markers. Green, CD31; red, mesenchymal markers; blue, DAPI. Scale bar, 100 μm. CTRL in A, B, and C indicate HUVECs infected with control adenovirus. CTRL in D and E indicate HUVECs transfected with control siRNA. α-SMA indicates alpha-smooth muscle actin; CD31, cluster of differentiation 31; DAPI, 4',6-diamidino-2-phenylindole; GFP, green fluorescent protein; HUVEC, human umbilical vein endothelial cells; SLUG, SNAI2 snail family transcriptional repressor 2; TAGLN, Transgelin; and VWF, von Willebrand factor.
Since EndMT is ubiquitously observed in bAVMs with heterogeneous genetic background, we argue that the EndMT, rather than the KRAS or ERK pathways, is playing the most crucial role in the development of bAVMs, and future studies could concentrate on targeting EndMT but not limited to KRAS or ERK pathways to R&D bAVM novel therapeutic drugs. The 3-hydroxy-3-methylglutaryl coenzyme A reductase inhibitors (statins) have widely been in clinical use for the reduction of total cholesterol and LDL-C (low-density lipoprotein cholesterol) levels and have been shown to be safe. Here, we presented the efficacy of lovastatin in reversing EndMT using in vitro and ex vivo experiments, providing new preclinical evidence for future clinical trials that might benefit numerous patients. Regarding the mechanism, it has previously been revealed that TGF-β signaling and oxidative stress are able to promote the EndMT process and lovastatin could significantly suppress TGF-β signaling and oxidative stress. Subsequent in vitro study found that lovastatin could suppress high glucose–induced EndMT in glomerular ECs through downregulating TGF-β1 expression and Smad2/3 phosphorylation. In addition, lovastatin could reduce malondialdehyde and superoxide dismutase decreasing oxidative stress in the kidney of diabetic mouse. Therefore, we suspect that lovastatin may alleviate EndMT in ECs through the inhibition of TGF-β signaling. However, further experimental studies will be needed to consolidate the molecular mechanisms before clinical applications.

ARTICLE INFORMATION

Received February 2, 2021; revision received September 3, 2021; accepted September 16, 2021.

Affiliations

Neurosurgey, Beijing Tiantan Hospital (H.L., RH, WF, Y.J., Jiancong Weng, ZY, Jie Wang, HX, SW, JZ, YC) and Beijing Neurosurgical Institute (YC), Capital Medical University, China. China National Clinical Research Center for Neurological Diseases, Beijing, China (H.L., RH, WF, Y.J., Jiancong Weng, ZY, Jie Wang, HX, SW, JZ, YC). Division of Life Science, State Key Laboratory of Molecular Neuroscience (YN, QZ, D.S., Z.Y., Jie Wang, H.X., S.W., J.Z., Y.C.) and Beijing Neurosurgical Institute (Y.C.), Capital University, Beijing, China (J.L.).

Sources of Funding

This article is funded by the project “Genomics Platform Construction for Chinese Major Brain Disease-AVM” (No. PX2019_026280_000002-AVM), Beijing Advanced Innovation Center for Big Data-based Precision Medicine (PM2020_014226_000066), NSFC Excellent Young Scientists Fund (No. 31922088), Hong Kong RGC Fund (C6002-17GF), Hong Kong Epigenomics Project (LJKCFL18SC01-E), and the Hong Kong ITC Fund (ITCPD/17-9).

Disclosures

None.

Supplemental Materials

Expanded Methods

Data Supplement Figures I–XXI

Data Supplement Tables I–VIII

References 49–59

REFERENCES


16. Jie Wang, H.X., S.W., J.Z., Y.C.) and Beijing Neurosurgical Institute (Y.C.), Capital University, Beijing, China (J.L.).