Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

HHS Public Access

Author manuscript

Transl Stroke Res. Author manuscript; available in PMC 2024 August 01.

Published in final edited form as:

Transl Stroke Res. 2023 August ; 14(4): 513–529. doi:10.1007/s12975-022-01050-3.

Circulating plasma miRNA homologs in mice and humans reflect familial cerebral cavernous malformation disease

Sharbel G. Romanos[1], **Abhinav Srinath**[1,#], **Ying Li**[1,2,#], **Bingqing Xie**[3], **Chang Chen**[1,4], **Yan Li**[1,4], **Thomas Moore**[1], **Dehua Bi**[1,5], **Je Yeong Sone**[1], **Rhonda Lightle**[1], **Nick Hobson**[1], **Dongdong Zhang**[1], **Janne Koskimäki**[1], **Le Shen**[6], **Sara McCurdy**[7], **Catherine Chinhchu Lai**[7], **Agnieszka Stadnik**[1], **Kristina Piedad**[1], **Julián Carrión-Penagos**[1], **Abdallah Shkoukani**[1], **Daniel Snellings**[8], **Robert Shenkar**[1], **Dinanath Sulakhe**[4], **Yuan Ji**[1,5], **Miguel A. Lopez-Ramirez**[7,9], **Mark L. Kahn**[10], **Douglas A. Marchuk**[8], **Mark H. Ginsberg**[7], **Romuald Girard**[1,‡], **Issam A. Awad**[1,‡]

[1]Neurovascular Surgery Program, Department of Neurological Surgery, University of Chicago Medicine and Biological Sciences, Chicago, IL, USA

[2]Department of Neurosurgery, First Affiliated Hospital of Harbin Medical University, Harbin, Heilongjiang, 150001, China

[3]Section of Genetic Medicine, Department of Medicine, University of Chicago, Chicago, IL, USA

[4]Bioinformatics Core, Biological Sciences Division, University of Chicago, Chicago, IL, USA

[5]Department of Public Health Sciences, University of Chicago, Chicago, IL, USA

[6]Department of Surgery, The University of Chicago, Chicago, IL, USA.

[7]Department of Medicine, University of California San Diego, La Jolla, California, USA.

[8]Molecular Genetics and Microbiology Department, Duke University Medical Center, Durham, NC, USA.

[9]Department of Pharmacology, University of California San Diego, La Jolla, California, USA.

Correspondence to: Issam A. Awad, MD, MSc, FACS, MA (hon), Department of Neurological Surgery, University of Chicago Medicine, 5841 S. Maryland, MC3026/Neurosurgery J341, Chicago, IL 60637 USA; Telephone +1 773-702-2123; Fax +1 773-702-3518; iawad@uchicago.edu.

[#]Equal Contributions

[‡]Equal Contributions

Author contributions: IAA and RG designed and conceptualized the study, oversaw data analyses, and edited the final manuscript. SR, AbS, and RG helped optimize the study design. SR, AbS, YiL, TM, JYS, RL, NH, DZ, JK, LS, SM, AgS, KP, JCP, AbSh, DS, RS, MLR, CCL, MK, DM, MG acquired the data. SR, AbS, YiL, BX, CC, DB, YaL, DS, YJ, RG, and IAA analyzed and interpreted the data. The manuscript was written by SR, AbS, RG, IAA, and all authors edited and gave final approval of the published work.

Conflict of Interest: The authors have declared that no conflict of interest exists.

Ethics Approval: All animal experiments adhered to the NIH Guide for the Care and Use of Laboratory Animals and were approved by the respective Institutional Animal Care and Use Committees at the University of Chicago and University of California San Diego. All human studies were approved by the University of Chicago Institutional Review Board, which is guided by ethical principles consistent with the Belmont Report and comply with the rules and regulations of the US Department of Health and Human Services Federal Policy for the Protection of Human Subjects (56 FR 28003).

Consent to participate: All human subjects gave written informed consent in compliance with the Declaration of Helsinki.

Consent for publication: Human research participants provided informed consent for publication of any images within the manuscript.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

[10]Department of Medicine and Cardiovascular Institute, University of Pennsylvania, Philadelphia, PA, USA

Abstract

Patients with familial cerebral cavernous malformation (CCM) inherit germline loss of function mutations and are susceptible to progressive development of brain lesions and neurological sequelae during their lifetime. To date, no homologous circulating molecules have been identified that can reflect the presence of germ line pathogenetic CCM mutations, either in animal models or patients. We hypothesize that homologous differentially expressed (DE) plasma miRNAs can reflect the CCM germline mutation in preclinical murine models and patients. Herein, homologous DE plasma miRNAs with mechanistic putative gene targets within the transcriptome of preclinical and human CCM lesions were identified. Several of these gene targets were additionally found to be associated with CCM-enriched pathways identified using the Kyoto Encyclopedia of Genes and Genomes. DE miRNAs were also identified in familial CCM patients who developed new brain lesions within the year following blood sample collection. The miRNome results were then validated in an independent cohort of human subjects with real-time-qPCR quantification, a technique facilitating plasma assays. Finally, a Bayesian-informed machine learning approach showed that a combination of plasma levels of miRNAs and circulating proteins improves the association with familial-CCM disease in human subjects to 95% accuracy. These findings act as an important proof of concept for the future development of translatable circulating biomarkers to be tested in preclinical studies and human trials aimed at monitoring and restoring gene function in CCM and other diseases.

Keywords

Cerebral Cavernous Malformations; microRNA; machine learning; genotype

Introduction

Cerebral cavernous malformations (CCMs), also known as cavernous angiomas, are enlarged blood-filled vascular caverns prone to hemorrhage due to dysfunctional vessel wall angioarchitecture affecting up to 0.5% of the population [1, 2]. Approximately 30% of CCM patients manifest familial disease, with an autosomal dominant inheritance of a heterozygous germline loss of function mutation in one of three genes, *CCM1/KRIT1, CCM2/Malcavernin,* or *CCM3/PDCD10* [1]. These familial cases develop new hemorrhagic brain lesions throughout their life, predisposing them to a risk of symptomatic hemorrhage, seizures, and/or focal neurological deficits [1]. The *CCM3/PDCD10* disease manifests exceptional aggressiveness with greater lesion burden and earlier disease manifestations than other genotypes [3]. Surgical excision of symptomatic CCM lesions is the only current therapeutic option, with serious morbidity and costs, and obvious limitations with multiple lesions [1]. Thus, there is an intense effort at developing novel therapies [4, 5], including gene restoration therapies with viral vectors for familial disease [6].

There is an ongoing search for facile biomarkers that can accurately reflect disease status and lesional activity of CCMs, to guide the selection of aggressive cases for clinical trials

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

and to monitor the impact of novel therapies [7-10]. Peripheral blood plasma provides a unique window into greater systemic biological processes, offering a promising opportunity for biomarker development [7, 11, 12]. In fact, dysregulated lesional pathways identified in CCM lesions may be reflected within the plasma [13]. Recent studies in preclinical murine models and human CCM patients have reported a differential plasma proteome in CCM disease, including proteins which have previously been implicated in CCM mechanisms [14-20], and reflect the transcriptome of micro-dissected human and murine CCM lesions [13, 21]. However, circulating molecules other than plasma proteins have not been systematically examined in CCM disease.

Micro-ribonucleic acids (miRNAs) are small non-coding molecules ~21-24 nucleotides in length that regulate gene expression through mRNA silencing and have been shown to modulate several biological processes [22]. Plasma miRNAs are stable and can be isolated, and their levels have been shown to be affected by disease states [23, 24]. The intra-cellular miRNome of surgically resected CCMs identified 5 top-differentially expressed (DE) miRNAs, several of which were related to CCM processes including VEGF, PI3K/Akt, and MAPK signaling [25]. Of interest, Koskimäki et al. (2019), from our team, also identified that *miR-3472a* was DE in the plasma of *Ccm3* heterozygous mice [21]. Additionally, a pilot study identified 13 circulating plasma DE miRNAs in CCM patients who experienced a recent symptomatic hemorrhage [26]. One of these 13, *miR-185-5p*, targets the *IL10RA* gene, a receptor for IL10, which is DE in the transcriptome of human CCM lesions [13, 26]. Thus, plasma miRNAs associated with disease characteristics may contribute further pathobiological insights.

Overall, it would be beneficial to identify circulating homologous miRNAs with postulated mechanistic links, which can efficiently be cross validated in preclinical models of disease and translated to human subjects [27]. This is especially advantageous where experimental therapeutics could be tested, and their efficacy monitored using the same biomarker in preclinical studies, in human trials and ultimately in clinical practice. To date, homologous plasma miRNAs which reflect disease states in murine models and humans have not yet been identified. We hypothesize that homologous DE plasma miRNAs are associated with the inherited CCM germline mutation in preclinical murine models and patients, can distinguish cases who would develop new lesions, and can be assayed in peripheral blood and thus could be tested as potential translatable circulating biomarkers in specific clinical contexts of use.

Methods

Mouse sample collection

Animal studies for miRNA sequencing and differential expression were conducted in N=41 mice. Blood was collected via the submandibular vein route from 3 to 5 month-old mice including *Ccm1*$^{+/-}$ mice (n=7), Cc*m3*$^{+/-}$ mice (n=7) and sex matched wild type mice (n=7), which were bred at the University of Chicago (Table S1). The pre-clinical heterozygous mice have been used to reflect germ line heterozygocity, as in the human familial disease [3].

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Ccm1$^{flox/flox}$ mice (n=7) and *Ccm3*$^{flox/flox}$ mice (n=5) with a Pdgfb-Cre^{+} driver, along with sex matched wild type mice (n=8) were bred at the University of California San Diego (Table S1). These mice were injected with 50μg of 4-hydroxy-tamoxifen intraperitoneally at P1 to induce Cre activity and gene loss, bypassing embryonic lethality of the homozygous state [28]. The mice and their controls were sacrificed uniformly at 7-10 weeks of age through cardiac puncture and blood was collected simultaneously.

After collection, plasma was separated using Z-gel tubes (Sarstedt, Nümbrecht, Germany) through centrifugation (AllegraX-30R, Beckman Coulter). One hundred ul of plasma was then aliquoted into 1.7-ml microcentrifuge tubes and stored at −80°C until RNA extraction. The sample size for this study was estimated based on a previous study that identified DE miRNAs in the plasma of only 3 *Ccm3*$^{+/-}$ mice when compared to 3 wild type mice [21].

Human CCM cohort characteristics

For this study, the plasma miRNome of N=30 subjects, including 23 familial-CCM patients (n=13 *CCM1* and n=10 *CCM3*) and healthy non-CCM subjects (n=7), enrolled between July 2014 and July 2019, was sequenced (Table 1). The diagnosis of CCM was established via magnetic resonance imaging (MRI) at a single referral center (uchicagomedicine.org/ccm) by a senior neurosurgeon (IAA) with more than 30 years of experience in CCM disease management. Patients with the CCM genotype defined as either *CCM1* or *CCM3* were included, confirmed through genetic testing performed by PreventionGenetics (Marshfield, WI, USA), utilizing Sanger and NextGen sequencing followed by deletion/duplication analysis [2]. Patients with partial or complete resection of CCM or any prior brain irradiation were excluded. Due to low numbers of *CCM2* cases, which only make up 20 % of familial cases, this genotype was unable to be analyzed separately [29]. Among the 23 familial-CCM patients enrolled, 10 (n=5 *CCM1* and n=5 *CCM3*) harbored at least 1 new CCM on their routine clinical follow-up susceptibility weighted imaging (SWI) MRI in the year following their blood sample collection. SWI is recommended for the diagnosis of CCM and assessment of lesion burden as this sequence is more sensitive in detecting perilesional hemosiderin and capturing smaller lesions than traditional T1- and T2-weighted MRI sequences [2, 30].

An independent cohort (N=37) including an additional 24 familial-CCM patients (n=12 *CCM1* and n=12 *CCM3*) and healthy non-CCM subjects (n=13) was also enrolled between July 2014 and July 2019 for validation of the miRNA sequencing using real time quantitative PCR (RT-qPCR) (Table 2). Within this independent cohort, 25 subjects (n=19 familial-CCM; n=6 healthy non-CCM) had miRNA levels measured through RT-qPCR as well as proteins assayed within the plasma. Demographic analyses showed that a greater proportion of white/Caucasian was enrolled in both cohorts of CCM patients [$p<0.05$]. No other statistically significant demographic difference was observed between any of the human CCM cohorts, including age and sex. Previous statistical simulations and clinical studies have reported several DE plasma miRNAs in CCM patients using similar sample sizes [7, 26].

Healthy non-CCM subjects were recruited concurrently, specifically excluding (a) any medical or neurologic condition requiring ongoing follow-up or medical treatment in the

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

preceding year, (b) a history of concussion or brain trauma in the preceding year, (c) a history of prior brain irradiation at any time, (d) been pregnant or lactating in the preceding year, (e) used recreational, psychoactive, or neuroleptic drugs in the prior year. The healthy non-CCM subjects had an MRI performed to ensure that they did not have any unknown neurological conditions that could confound the study.

miRNA sequencing and differential expression analyses

Total RNAs from the serum of both heterozygous and homozygous mice (200μl), and from the plasma of familial-CCM patients (up to100μl) were extracted using the miRNeasy Serum/Plasma Kit (Qiagen, Hilden, Germany) following the manufacturer recommendation [21, 26]. cDNA libraries were then generated with commercially available Illumina small RNA-Seq kits (Clontech, Mountain View, CA, USA) and sequenced with the Illumina HiSeq 4000 platform (Illumina, San Diego, CA, USA) using single-end 50-bp reads, at the University of Chicago Genomics Core for sequencing. The sequencing depth median for the plasma miRNomes was 10 million and 12 million reads per sample for mouse and human analyses, respectively. The sequencing depth used is in the same range as that reported in orther studies examining known miRNAs that are differentially expressed between groups [31, 32].

The differential expression analyses were first performed in the plasma of the preclinical mouse models between (1) $Ccm1^{-/-}$, (2) $Ccm1^{+/-}$, (3) $Ccm3^{-/-}$, (4) $Ccm3^{+/-}$, and their respective wild type controls [$p<0.1$, false discovery rate (FDR) corrected]. Additional differential expression analyses were done between (1) $Ccm1^{-/-}$ and $Ccm3^{-/-}$, and (2) $Ccm1^{+/-}$ and $Ccm3^{+/-}$ (Fig. 1).

Differential expression analyses were performed in human between healthy non-CCM controls and (1) *CCM1* patients, (2) *CCM3* patients, and then (3) between *CCM1* and *CCM3* patients (Fig. 1). An additional analysis was performed between the CCM patients showing New Lesion Formation and No New Lesion Formation. Though these results were not FDR corrected, additional stringency was added as only homologous DE miRNA with (1) gene targets within CCM transcriptomes and (2) mechanistic links reported in CCM disease were considered most relevant. All differential expression analyses were conducted using R bioconductor package DESeq2. Additionally, a partial least-squares discriminant analysis (PLS-DA) was used to validate the capacity of DE miRNAs ($p<0.05$, non-FDR corrected) to distinguish between CCM1, CCM3, and healthy non-CCM [33-35]. Using the mixOmics package in R [36], PLS-DA components were sorted based on the variance explained, and the area under the curve (AUC) and *p*-value were reported for the top 2 components of each genotype comparison.

Putative targets of homologous mouse-human DE miRNAs

The human homologs of DE miRNAs identified in preclinical murine models were queried using the MirGeneDB database (https://mirgenedb.org/) (Fig. 1). Putative gene targets of homologous mouse-human DE miRNAs were identified using miRWalk 3.0 [37, 38] (Fig. 1). For each miRNA, gene targets were identified for the 3 different gene locations (3′ untranslated region (UTR), 5′ UTR, and coding sequence (CDS)) using a random forest tree

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

algorithm with a bonding prediction probability higher than 95%. All known gene sequences are scanned and any matches to known miRNA sequences are flagged. These matches are then compared and validated using 8 other established miRNA-target prediction programs [39]. MiRWalk consolidates data from several databases, including TargetScan, miRDB, and miRTarBase, and reports known miRNA-gene interactions from the literature to provide both predicted and validated gene target binding sites of miRNAs [39]. Only genes that appeared in at least 2 of the 3 databases on miRWalk were considered to be putative gene targets of the homologous DE miRNA, ensuring these interactions were validated by more than one source [13, 21, 26].

The putative targets of each DE miRNA identified in the plasma of mouse *Ccm1*$^{-/-}$ and *Ccm1*$^{+/-}$ were queried within the previously published transcriptome of *in vitro Ccm1*$^{-/-}$ mouse brain microvascular endothelial cells (BMECs) [13]. While the putative targets of the DE miRNAs identified in the plasma of *Ccm3*$^{-/-}$ and *Ccm3*$^{+/-}$ mice were queried in the previously published transcriptome of (1) *in vitro Ccm3*$^{-/-}$ mouse BMECs, and (2) laser micro-dissected lesional neurovascular units (NVUs) of *Ccm3*$^{+/-}$ mice [13, 21]. Gene targets of DE miRNAs identified in the plasma of *CCM1* and *CCM3* patients were queried within the previously published transcriptome of laser micro-dissected NVUs of human surgically resected CCM lesions [13].

CCM-enriched Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analyses

The enriched KEGG pathways of each transcriptome were derived independently using the differentially expressed genes (DEGs) [$p<0.05$, FDR corrected; $|FC|>1.5$] identified in (1) *Ccm1*$^{-/-}$ BMECs, (2) *Ccm3*$^{-/-}$ BMECs, as well as micro-dissected lesional NVUs of (3) *Ccm3*$^{+/-}$ mice, and (4) human surgically resected CCMs. Enriched KEGG pathways [$p<0.05$, FDR corrected] were generated using a database and knowledge extraction engine [40] with a Bayes factor cutoff of 3. The putative target DEGs of each homologous DE miRNA were then associated with the list of CCM-enriched KEGG pathways within their respective transcriptomes. Finally, KEGG pathways and putative target DEGs were categorized into 6 biological processes related to CCM disease identified after a comprehensive literature search [21]. These biological processes include: (1) Neuron, Glia, Pericyte Function, (2) Permeability/Adhesion, (3) Apoptosis and Oxidative Stress, (4) Inflammation/Immune Response, (5) Vascular Processes, (6) Cellular Proliferation.

Selection of miRNAs for validation using RT-qPCR

RT-qPCR was not only used to validate the miRNA sequencing results but also serve as proof of feasibility for direct measurement of miRNAs in plasma for future biomarker use [41]. Among the homologous DE plasma miRNA with gene targets within CCM transcriptomes, 4 were chosen for RT-qPCR validation as (1) they have at least 10 DEGs [$p<0.05$, FDR corrected] as putative targets within the human CCM lesional transcriptome, and (2) had putative gene targets previously implicated in CCM disease. In addition, *miR-375-5p* was assessed because it was dysregulated across both *CCM1* and *CCM3* genotypes in the plasma of both preclinical murine models and CCM patients. Finally, not DE in mice, *let-7e-5p* was selected since it was the only miRNA dysregulated in *CCM1*

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

patients when compared to healthy controls as well as compared to *CCM3* patients, making it a potential specific marker for the *CCM1* genotype.

RT-qPCR assay of putative DE miRNAs

Relative quantifications of a panel of 6 pre-selected DE miRNAs were assessed using RT-qPCR. The total RNA was first extracted from plasma using the MagMAX *mir*Vana Total RNA Isolation Kit (Thermo Fisher, Waltman, MA, USA). An exogenous spike-in control, *miR-cel-39-3p* ($1.5x10^{10}$ copies of *cel-miR-39-3p* in 5μl) (Integrated DNA Technologies, Inc., Coralville, IA, USA), was added to all plasma samples prior to extraction and then measured to correct for extraction efficiency, as well as to ensure the inter-plate reproducibility of the RT-qPCR [42]. RT-qPCR was performed using the TaqMan Advanced miRNA Assays Kit (Thermo Fisher). A miR-Amp reaction was performed to amplify miRNAs prior to real-time PCR using the QuantStudio3 (Thermo Fisher).

For the relative quantification, the Cq values of each selected plasma miRNA as well as of two human endogenous controls, *miR423-5p and miR16-5p,* and of exogenous control *cel-miR39-3p* were measured. *MiR-423-5p* was eventually used as the endogenous control as it is expected to be expressed at equal levels across tissue types and samples [43]. This was validated using the NormFinder software, as *miR-423-5p* was shown to have the most stable expression across the controls tested (https://moma.dk/normfinder-software) [44]. The Cq values of the endogenous control (i.e., *miR-423-5p*) were then used to calculate the relative plasma quantification [45]:

$$\Delta Cq = Cq_{\text{miRNA of interest}} - Cq_{\text{endogenous control}}$$

The comparison of the relative quantification of each pre-defined miRNA was performed between (1) familial-*CCM1* patients, (2) familial-*CCM3* patients, and (3) healthy non-CCM controls. The Cq and the plasma level of miRNA in a sample are inversely related (i.e.. lower Cq represents a higher concentration), thus ΔCq would be in the opposite direction of the FC reported through sequencing [46]. For each miRNA, a relative quantification value greater than ±3 standard deviations away from the mean was defined as an outlier [47, 48]. A non-parametric Mann-Whitney was performed between the ΔCq for each comparison using Prism (GraphPad, San Diego, California) [49]. The FC for each miRNA of interest was also calculated using $2^{-\Delta\Delta Cq}$. The $\log_2$(FC) were then compared using an unpaired Student's t-test using Prism (GraphPad) [49-51].

Combined plasma miRNA and protein levels in association with familial CCM

A three-step Bayesian-informed machine learning (ML) approach was implemented to assess if a weighted combination of plasma levels of miRNAs (i.e., assessed by RT-qPCR), and proteins (i.e., assessed by ELISA) shows better performance to distinguish familial-CCM. This Bayesian paradigm relies on prior information, such as the biological relevance of candidate molecules, to inform future development [52]. Additionally, ML approaches allow for the combination of the best performing plasma molecules, given that individual molecules may have imperfect sensitivity/specificity [7]. A thorough literature search was conducted to identify 21 plasma proteins related to angiogenesis, endothelial

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

permeability, inflammation, immune response, and extracellular matrix remodeling, with reported associations to CCM or neurovascular diseases (Table S2) [7, 26, 53, 54]. As a proof of concept, we assessed familial disease reflecting germline heterozygosity of any *CCM* gene, since the sample sizes of specific CCM genotypes were insufficient for separate statistical analyses (Fig. S5-6). Two different approaches were developed using 10-fold and leave-one out cross validation to first select the best individual proteins and miRNAs which reflected familial-CCM disease (Fig. S1). These cross validations reduce the overfitting effects on a limited dataset, while excluding any molecules which would not individually contribute to the model's diagnostic association [55]. Selected miRNAs or proteins were then used to create the best models including (1) only plasma levels of miRNA, (2) only plasma levels of proteins, and finally (3) combining miRNAs and proteins to diagnose familial-CCM (Fig. S1). In each case, the best model was identified as having the lowest Akaike information criterion (AIC) score. Feature selection minimized the dependency among features under the assumption of logistic regression, while selecting the model with AIC scores reduced the overall complexity of the model to prevent overfitting. Finally, all three models were tested on the same cohort of patients to compare performance through sensitivity and specificity (Fig. S1). *Refer to* supplemental material *for additional Methods.*

Results

DE miRNAs in peripheral blood plasma of Ccm1 and Ccm3 murine models

Differential expression analyses identified 17 DE miRNAs in the plasma of $Ccm1^{-/-}$ mice (Data File S1), and 44 in $Ccm1^{+/-}$ mice [all: $p<0.1$, FDR corrected] (Data File S2) when compared to their respective wild type mice. In the *Ccm3* models, 142 plasma miRNAs were DE in the homozygous mice (Data File S3), and 9 in the heterozygous mice when compared to their respective wild type mice [all: $p<0.1$, FDR corrected] (Data File S4). Finally, 99 DE plasma miRNAs were identified between $Ccm1^{-/-}$ and $Ccm3^{-/-}$ mice [all: $p<0.1$, FDR corrected] (Data File S5). No plasma miRNAs were DE between $Ccm1^{+/-}$ and $Ccm3^{+/-}$ mice. Non-sensitized *Ccm1* and *Ccm3* heterozygous mice have been shown to harbor a negligible lesion burden. Thus, these DE miRNAs were discovered to reflect disease genotype, rather than disease severity.

Homologous mouse-human DE miRNAs in peripheral blood plasma of patients with CCM1 and CCM3 disease

Differential analyses identified 16 DE plasma miRNAs between familial-*CCM1* patients and healthy non-CCM controls [$p<0.05$] (Data File S6). The PLS-DA results showed that the DE miRNAs ($p<0.05$) identified through RNA-seq are able to distinguish *CCM1* from healthy non-CCM subjects (top 2 components, both: AUC 100%, $p=7.5\text{x}10^{-4}$). Two of these 16 miRNAs, *miR-20b-5p* [fold change (FC)=25.9] and *miR-375-3p* [FC=3.5] had their murine homolog also DE in the plasma of $Ccm1^{+/-}$ mice [FC=2.6, 5.2 respectively; both: $p<0.1$, FDR corrected] (Fig. 2). No plasma miRNAs were commonly DE between familial-*CCM1* patients and $Ccm1^{-/-}$ mice.

In the familial-*CCM3* patients, 12 DE plasma miRNAs were identified compared to healthy non-CCM controls [$p<0.05$] (Data File S7). PLS-DA results confirmed that DE miRNAs

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

(p<0.05) can differentiate *CCM3* from healthy non-CCM subjects (top 2 components, both: AUC 100%, p=6.4x10^{-4}). Five of these 12 DE plasma miRNAs, *miR-93-5p* [FC=0.6], *miR-370-3p* [FC=4.3], *miR-487b-3p* [FC=13.3], *miR-369-5p* [FC=19.2], and *miR-375-3p* [FC=2.8] had their murine homolog also DE in the plasma of *Ccm3*$^{-/-}$ mice [FC= 0.5, 2.7, 6.7, 2.7, 3.5 respectively; all: p<0.1, FDR corrected] (Fig. 2). No plasma miRNAs were commonly DE between familial-*CCM3* patients and *Ccm3*$^{+/-}$ mice. It is unclear why homologous DE miRNAs associated with germline *CCM1* and *CCM3* heterozygosity in humans were shared with heterozygous *Ccm1*$^{+/-}$ mice and homozygous *Ccm3*$^{-/-}$ mice, respectively. The number of plasma DE miRNAs in heterozygous *Ccm3*$^{+/-}$ and postnatally induced *Ccm1*$^{-/-}$ mice was low, which may have been due to a low read depth during sequencing. In addition, postnatally induced gene loss in mouse models probably reflects an acute phase of the pathogenesis of CCMs, and not the chronic aspect seen in the human disease; these differences are exaggerated in the more aggressive *Ccm3* models.

We further identified 24 plasma DE miRNAs between *CCM1* and *CCM3* patients [p<0.05] (Data File S8). DE miRNAs (p<0.05) dentified through RNA-seq were shown to distinguish *CCM1* from *CCM3* through an additional PLS-DA (component 1, AUC 98.15%, p=2.2x10^{-4}; component 2, AUC 100%, p=1.2x10^{-4}). Of these, *miR-128-3p* [FC=0.6], *miR-410-3p* [FC=2.0], *miR-9-5p* [FC=0.07], *miR-323b-3p* [FC=13.7], and *miR-93-5p* [FC=0.7] had their murine homolog also DE between *Ccm1*$^{-/-}$ and *Ccm3*$^{-/-}$ murine models [FC=1.8, 1.8, 28.2, 6.6, and 0.6 respectively; all: p<0.1, FDR corrected] (Fig. 2).

Gene targets of homologous mouse-human DE plasma miRNAs

Of the 10 homologous DE miRNAs, 8 were DE in the same direction in familial-CCM patients and preclinical mouse models (all genotypes) (Fig. 2). Seven homologous DE miRNAs, targeting a range of 0.1%-2.4% of the total human genome, had DEG targets within the published transcriptome of laser micro-dissected NVUs of human surgically resected CCMs (Data File S9, Fig. 3). In addition, 5 of those 7 miRNAs also had DEG targets within either the transcriptome of (1) *in vitro Ccm1*$^{-/-}$ *or Ccm3*$^{-/-}$ BMECs, or (2) laser micro-dissected lesional NVUs of *Ccm3*$^{+/-}$ mice. To further classify these gene interactions, the KEGG was used to generate enriched KEGG pathways [p<0.05, FDR corrected; Bayes factor≥3] for each CCM transcriptome (Data File S10-13).

DE *miR-20b-5p,* which was increased in the plasma of familial-*CCM1* patients, and *miR-93-5p,* which was decreased in familial-*CCM3* patients compared to healthy non-CCM controls, respectively had 8.9% and 9.1% of their gene targets within the transcriptome of human micro-dissected CCM lesional NVUs. CCM-enriched KEGG pathway analyses showed that these DEGs were associated with PI3K-Akt signaling, focal adhesion, HIF-1, Rap1 signaling, and other pathways involved in CCM mechanisms [all: p<0.05, FDR corrected] (Data File S10). Specifically, both *miR-20b-5p* and *miR-93-5p* target *VEGFA* and *ADAMTS5*. Additionally, *miR-93-5p* targets *ROCK2* and *MAP3K14,* which are associated with CCM pathogenesis. *miR-93-5p* provides a unique translational opportunity for biomarker development as it also targets *Vegfa* in mice, which is a DEG in the transcriptome of laser micro-dissected lesional NVUs of *Ccm3*$^{+/-}$ mice.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

DE *miR-370-3p* and *miR-487b-3p*, both found to be increased in the plasma of familial-*CCM3* patients, were also found to target several DEGs in the transcriptome of human micro-dissected CCM lesional NVUs. Specifically, 9.4% of *miR-370-3p*'s gene targets, and 33.3% of *miR-487b*-3p's targets were within the human CCM lesional transcriptome. *miR-370-3p* targets *NFASC* and *FGF7*, which were both found to be associated with PI3K-Akt signaling, Rap1 signaling, and cell adhesion molecules [all: $p<0.05$, FDR corrected] (Data File S10). It also targets *Klhl21,* a DEG in the transcriptome of micro-dissected lesional NVUs of $Ccm3^{+/-}$ mice. In addition, *miR-487b-3p* was found to target *NRARP,* which is DE in the human CCM lesional transcriptome.

Finally, DE *miR-128-3p* and *miR-9-5p,* which were decreased, and *miR-410-3p,* which was increased (i.e., all between familial-*CCM1* compared to familial-*CCM3* patients) also target DEGs within the human CCM lesional transcriptome. *miR-128-3p*, *miR-9-5p*, and *miR-410-3p* had 10.8%, 13.5%, and 9.0% of their gene targets in the human CCM lesional transcriptome, respectively. Among these, *IGF1* and *NRXN1* are targeted by *mir-128-3p*. Of interest, *IGF1* is associated with enriched-KEGG pathways such as PI3K-Akt signaling, HIF-1 signaling, and Rap1 signaling, while *NRXN1* is related to cell adhesion molecules [all: $p<0.05$, FDR corrected] (Data File S10). Also, *miR-128-3p* targets *Foxq1* in mice, which is DE in the transcriptome of $Ccm1^{-/-}$ mouse BMECs. Furthermore, *miR-9-5p* targets *TNC*, *VAV3*, and *VCAN*, which are associated with focal adhesion and cell adhesion KEGG pathways [all: $p<0.05$, FDR corrected] (Data File S10). Additionally, *miR-9-5p* targets *Rap1b* which is DE in the transcriptome of $Ccm1^{-/-}$ mouse BMECs. Finally, *miR-410-3p* targets two DEGs within the human CCM lesional transcriptome, both of which are part of the solute carrier group of membrane transport proteins. Interestingly, one of these genes, *SLC8A1*, is also targeted by *miR-410-3p* in mice and is a DEG in the transcriptome of laser micro-dissected lesional NVUs of $Ccm3^{+/-}$ mice.

Plasma miRNAs are DE in familial CCM patients with new lesion formation

Ninety-seven plasma miRNAs [$p<0.05$, non-FDR corrected] were DE in familial-CCM patients that developed a new lesion within a year after their blood draw, as compared to those with stable lesion burden (Data File S14). Fifteen of these 97 were also DE in familial-*CCM3* patients, when compared to healthy non-CCM subjects or familial-*CCM1* patients. Of these 15 plasma miRNAs, 9 (*miR-141-3p, miR-409-3p, miR-431-5p, miR-485-3p, miR-9-5p, miR-128-3p, miR-379-5p, miR-205-5p,* and *miR-1271-5p*) were found to target DEGs within the transcriptome of laser micro-dissected NVUs of human surgically resected CCMs. These DEGs include *ROCK2*, *VEGFA*, *ADAMTS5*, *VCAN*, and *NRXN1*, which are associated with several CCM-enriched KEGG pathways such as PI3K-Akt signaling, HIF-1 signaling, Rap1 signaling, focal adhesion, and cell adhesion molecules [all: $p<0.05$, FDR corrected] (Data File S10).

Peripheral blood levels of DE miRNAs are associated with familial CCM disease

A validation of the miRNA sequencing differential analyses was performed by assessing the plasma levels of a selected panel of six DE miRNAs using RT-qPCR within an independent cohort of 12 familial-*CCM3*, 12 familial-*CCM1* patients, and 13 healthy non-CCM controls

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

(Fig. 4, Fig. S2). Relative quantification ΔCq levels assessed through RT-qPCR supported the dysregulation observed in the miRNome sequencing data (Fig. 4).

Lower RT-qPCR relative quantification values (i.e., ΔCq) of *miR-20b-5p* in the plasma of familial-*CCM1* were observed compared to healthy non-CCM controls [p=0.04] (Fig. 4a). A trend toward lower relative quantification values of *miR-20b-5p* was also observed in the plasma of familial-*CCM1* patients compared to familial-*CCM3* patients [p=0.07] (Fig. 4a). No difference was observed between familial-*CCM3* patients and healthy non-CCM subjects.

Familial-*CCM1* patients also showed greater RT-qPCR relative quantification values of plasma *miR-93-5p* compared to healthy non-CCM subjects [p=0.007] (Fig. 4b). In addition, a trend toward higher relative quantification values was observed between familial-*CCM3* compared to healthy non-CCM subjects [p=0.06] (Fig. 4b). No difference between familial-*CCM3* and familial-*CCM1* patients was observed.

The relative plasma quantification RT-qPCR values of *miR-9-5p* were higher in the plasma of familial-*CCM1* patients [p=0.01] as well as in healthy non-CCM subjects compared to familial-*CCM3* patients [p=0.01] (Fig. 4c). The relative quantification values of *miR-375-3p* were also higher in the plasma of familial-*CCM1* [p=0.05] and familial-*CCM3* patients [p=0.02] compared to healthy non-CCM subjects (Fig. 4d).

Finally, the relative quantification plasma values for *miR-128-3p* did not show differences in any of the comparisons (Fig. 4e), while relative quantification values of the miRNA *let-7e-5p* were higher in the plasma of familial-*CCM1* [p=0.01] and familial-*CCM3* [p=0.05] compared to healthy non-CCM subjects (Fig. 4f). The $\log_2$(FC) of DE miRNAs, as calculated using $2^{-\Delta\Delta Cq}$ also support the differences between groups identified through sequencing (Fig. S2).

Further linear Pearson correlation analyses showed no correlation between relative quantification of any of these miRNAs and the lesion burden. These results suggest that these DE plasma miRNAs reflect disease genotype, rather than severity.

Integration of plasma levels of DE miRNAs with proteins enhances their association with familial-CCM

The relative quantification plasma levels of the six miRNAs were used to create the most optimal weighted combination (i.e., achieving the lowest AIC score) able to diagnose familial-CCM patients in comparison to non-CCM controls. This model's accuracy was defined as fair [56] in distinguishing familial-CCM patients with 84% sensitivity and 67% specificity [AUC=75.4%, CI = 52.4% - 98.5%] (Fig. S3).

A similar approach was performed using plasma levels of proteins in familial-CCM patients. Using two independent ML approaches, the best weighted combination of plasma proteins identified familial-CCM patients from healthy non-CCM with sensitivity up to 89.5% and specificity up to 100% [optimal AUC =96.5%, CI = 90.1% - 100%] (Fig. S4).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Given the imperfect diagnostic performance of miRNAs or proteins alone, combining the two may offer a more holistic model reflecting various aspects of the disease. Thus, an integrated biomarker including proteins and miRNAs was developed using familial-CCM patients (n=19) with both plasma levels of protein and relative quantification of miRNAs. Using a 10-fold cross validation ML approach, the weighted combination of proteins and relative quantification levels of miRNAs diagnosed familial-CCM patients with an accuracy of 95%, with a sensitivity of 94.7% and sensitivity of 100%. (Fig. 5a and b):

$$\text{Canonical Value} = -114.88^{*}[let-7e-5p]_{r} + 359.47^{*}[miR-93-5p]_{r} - 118.16^{*}[miR-20b-5p]_{r} + 88.93^{*}[miR-128-3p]_{r} + 199.47^{*}[\text{IL}-10] - 19.7^{*}[\text{Tsp}-2] - 463.18$$

with $[X]_r$ denoting relative quantification of miRNAs. This model estimated higher canonical values for familial-CCM patients compared to non-CCM healthy controls [$p<0.0001$] (Fig. 5a).

In order to validate these results, another ML approach, a leave-one-out cross validation, was also applied. This approach yielded a combination with the same performance of 94.7% sensitivity and 100% specificity (Fig. 5c and d) in diagnosing familial-CCM, however it included less compounds making it potentially more feasible for future clinical applicability:

$$\text{Canonical Value} = 2.05^{*}[miR-9-5p]_{r} - 3.02^{*}[miR-93-5p]_{r} - 0.01^{*}[\text{IL}-2] + 0.1^{*}[\text{TNFRI}] + 0.46^{*}[\text{Tsp}-2]$$

This additional weighted combination calculated lower canonical scores for familial-CCM patients compared to non-CCM healthy controls [$p<0.0001$] (Fig. 5c). *Refer to* supplemental material *for additional results.*

Discussion

Herein, we identify several homologous DE plasma miRNAs in preclinical murine models and CCM patients, with links to genes dysregulated in CCM disease. Additionally, we show that the integration of readily measurable plasma proteins and plasma levels of miRNAs into a combined biomarker can increase the accuracy of association with familial-CCM disease up to 95%. To our knowledge, such a translational approach to circulating molecule discovery in mouse models and humans, with mechanistic links to the disease transcriptome, validation in an independent cohort of human subjects, and the combination of miRNA and protein levels to enhance diagnostic accuracy have not been previously described in this or any other disease.

The role of the PI3K/Akt pathway, and specifically a mutation in *PIK3CA*, have long been associated with genesis and maturation of malignant cancers as well as vascular malformations [57-61]. Recently, a gain of function of *PIK3CA* was directly linked to CCM development and maturation [62-64]. The differential plasma miRNome identified several miRNAs that target DEGs with a reported role in PI3K/Akt signaling. *miR-20b-5p* and *miR-93-5p* both target *VEGFA*, which acts as a substrate and effector of the PI3K/Akt pathway, and whose involvement in angiogenesis is well elucidated [60].

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Moreover, *miR-93-5p* targets *ROCK2,* which has been shown to be involved in CCM pathogenesis [65, 66]. A mutation in one of the three CCM genes can cause an increase in RhoA/ROCK signaling, which in turn disrupts intercellular junctions and vascular permeability [67]. Interestingly, *ROCK2* is the rho-associated protein kinase isoform most highly expressed in the central nervous system and is a key regulator of vascular remodeling and endothelial cell barrier function [66]. Atorvastatin, which has pleotropic ROCK inhibitor properties, is currently being tested in a clinical trial to evaluate its ability to stabilize CCM lesions in patients with symptomatic hemorrhage [68]. Another target of *miR-93-5p, MAP3K14*, is involved in MAPK signaling. CCMs arise from an upregulation of the MAPK/MEKK/ERK3 signaling pathway, with downstream overexpression of transcription factors Klf2 and Klf4 [69]. These findings on plasma circulating *miR-93-5p* may suggest a mechanistic role in CCM pathogenesis, and therefore deserve further investigation.

Furthermore, *miR-20b-5p*, *miR-128-3p,* and *miR-93-5p* target DEGs associated with HIF-1 signaling. It has recently been shown that CCMs exhibit an increase in HIF activity. In fact, an increase in COX-2, a HIF-1α target, leads to increased CCM lesion genesis [70]. Additionally, hypoxia increases *VEGF* expression, leading to an upregulation of HIF-1 and PI3K/Akt signaling, which promote angiogenesis [3, 71]. Further mechanistic studies on circulating *miR-20b-5p, miR-128-3p,* and *miR-93-5p* may identify their role in HIF-1 signaling.

Of the homologous DE miRNAs, *miR-9-5p* targets genes associated with CCM-enriched KEGG pathways related to cell adhesion molecules and focal adhesion such as *TNC*, *VAV3*, and *VCAN,* which are also DEGs in the transcriptome of human micro-dissected CCM NVUs. An endothelial secretion of ADAMTS5 and cleavage of versican, coded by *VCAN*, have recently been reported as downstream mechanisms of CCM pathogenesis [72]. Additionally, an endothelial gain of *ADAMTS5* was shown to synergize with *CCM1* loss of function to create larger vascular malformations [72]. Interestingly, *ADAMTS5* is also a putative target of not only *miR-20b-5p*, which was DE in the plasma of *CCM1* patients and *Ccm1* mice, but also *let-7e-5p,* which was specifically DE in the plasma of *CCM1*-familial patients.

Finally, Rap1 signaling was found to be associated with several DEGs targeted by *miR-20b-5p*, *miR-93-5p, miR-128-3p*, and *miR-370-3p*. Rap1 signaling is involved in endothelial cell migration, proliferation, and regulation of vascular permeability [54, 73, 74]. *RAP1*, which is downstream of *VEGF*, has been shown to directly interact with *CCM1/KRIT1*, relocating it from microtubules to the cell membrane and impacting integrin activation [75]. Of interest, *miR-9-5p* was identified as putatively targeting *Rap1b* within the transcriptome of *in vitro Ccm1*$^{-/-}$ mouse BMECs.

These suggested interactions between DE miRNAs and DEGs within CCM transcriptomes could also define new readily translatable therapeutic targets, potentially using miRNAs for mRNA silencing therapy [76, 77]. DE miRNAs may also be associated with other disease features, such as new lesion formation, and may help generate mechanistic hypotheses for processes mediating *de novo* lesion genesis. Herein, 15 plasma miRNAs were commonly DE in *CCM3* patients, as well as patients which formed new lesions within the following

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

year. Given that *CCM3* patients experience a more aggressive form of the disease [1], these common DE miRNAs may offer new hypotheses surrounding lesion genesis and disease aggressiveness.

This study is the first to show a weighted combination of levels of plasma miRNAs and proteins associated with genotype in a Mendelian disease, based on ML approaches and Bayesian concepts. Two cross-validation methods, which reduced overfitting effects, confirmed that integrating plasma miRNAs and proteins improves the diagnostic association with familial-CCM compared to each circulating compound (e.g., plasma miRNA and protein) alone. This is an important proof of concept for integration of several types of molecules in plasma biomarker development [7]. A homologous biomarker in mice and humans, which can differentiate familial-CCM with germline loss of function mutations could be immediately applied in preclinical studies and early trials, monitoring effectiveness of gene restoration therapeutics, which are currently of great interest. As patients undergo gene therapy, one could monitor their progress and the duration of therapeutic effect as biomarker canonical values shift in comparison to a healthy non-CCM profile.

In summary, we have shown for the first time in any disease, that homologous DE plasma miRNAs can reflect the germline loss of function mutation in preclinical murine models and patients. We have additionally shown that plasma miRNAs can also distinguish cases who would develop new lesions, and plasma levels of these miRNAs can be assessed for facile diagnostic assays. We describe a novel approach of circulating biomarker development which can be applied in specific clinical contexts of use in familial-CCM and other Mendelian diseases.

There are limitations to our study. We had not previously defined a specific transcriptome of *in vivo Ccm1* mouse lesions as with *Ccm3* mice [21]. Therefore, it remains unclear if the plasma miRNAs identified in preclinical *Ccm1* mice have specific putative targets within the lesional transcriptome in that genotype. However putative targets were identified within the transcriptome of murine BMECs with induced *Ccm1* loss [13]. Furthermore, gene targets for human DE miRNAs were only queried within the CCM lesional NVU transcriptome, and connections to individual cell lines within the NVU cannot be described. Current studies are being conducted to identify specific genetic dysregulation in individual cell lines of lesional NVUs.

Additionally, preclinical mouse models may reflect a specific stage of CCM pathogenesis, and therefore may not perfectly mimic the chronicity of human disease, and this may impact miRNA discovery [78, 79]. The phenotypes of the respective mouse models have been described previously, including detailed neuropathologic correlates [80, 81]. The heterozygous mice were included as they most closely mimic the human disease state, with germ line heterozygosity. We also queried circulating miRNAs in mice with induced homozygous gene loss in view of their greater lesion burden. The biomarkers were identified in age matched cohorts but did not assess their stability or differences over time at various ages. And while there were no overt sex differences among the cohorts, sample sizes were small, and future studies should address potential sex-related biomarker differences.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

The DE miRNAs in human plasma were not significant after FDR correction, and thus were not as statistically stringent as with the mouse miRNA discovery. However, DE miRNAs in humans were statistically validated through an additional PLS-DA analysis, confirming their ability to distinguish between CCM genotypes. FDR corrected data in mice was used to guide the selection of homologous DE miRNAs in humans for additional analyses. Both their homologous nature, and additional mechanistic links within the CCM lesional transcriptome, offered biological relevance endorsing the selection of a subset of miRNAs for further validation. Indeed, these were validated with significant differences and the same directionality of respective plasma levels by RT-qPCR in an independent human cohort.

Mere discovery of an association does not validate a biomarker. Validation requires substantial additional development, focusing on specific analytic hypotheses and clinical contexts of use. Further investigation with larger sample sizes is needed to confirm the validity of association of these DE miRNAs with familial CCM disease. However, the homologous miRNAs identified herein , their validation in an independent cohort, and their accurate performance when combined with plasma proteins is highly encouraging. In fact, a large NIH-funded initiative to enroll patients for plasma protein and miRNA biomarker development has recently been launched, and this study is powered to examine effects of sex, age, lesion location and genotype, and patient recruitment at multiple sites [7].

Finally, the identification of potential gene targets of DE miRNAs was limited to databases available within miRWalk. However, miRWalk employs an extensive comparative search of several reputable databases and literature sources to ensure both predicted and validated miRNA-gene interactions are reported [39]. Out of the 10 homologous DE miRNAs queried in miRWalk, only 1 did not have any putative gene targets identified (Data File S9).

This study identified potential biomarker candidates that reflect CCM genotype (i.e., germline *CCM*/Ccm gene loss), regardless of lesion burden in patients and mice [81, 82]. The preliminary results shown herein for the development of biomarkers reflecting genotype do not imply direct applicability in other contexts of use such as lesion burden or bleeding. The discovery of miRNAs related to new lesion formation provides a proof of concept for a potential prognostic context of use, to be tested and validated in future studies. Other studies are underway to identify miRNAs in correlation with lesional hemorrhage.

Finally, while the mechanistic links of the DE miRNAs are compelling, they are merely hypothesis generating. Causation was only investigated in relation to gene loss in mouse models. Isolation of miRNA from lesional endothelial cells or endothelial specific circulating exosomes may provide more direct mechanistic insights, however this falls outside the scope of our current study, which is primarily focused on associations with molecules in the circulating plasma. Since the origin and role of circulating miRNAs remains unclear, it is likely that lesional and circulating miRNAs may show different signatures of dysregulation [83, 84]. Thus, no concrete claims can be made about these plasma miRNAs and their intra-lesional levels. Other research is underway aimed at metabolomic discovery, which might identify other small circulating molecules that can further enhance biomarker performance.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements:

This work was supported by grants from the National Institutes of Health, the William and Judith Davis Fund in Neurovascular Research, the Be Brave for Life Foundation, and the Safadi Program at the University of Chicago Translational Fellowship.

Funding:

This work was supported by grants from the NIH (R21NS087328, 5U01NS104157-02, 1R01NS107887-01, 2 P01 NS092521-06), K01 HL133530 to MLR, William and Judith Davis Fund in Neurovascular Research to IAA, by the Be Brave for Life Foundation to RG, and the Safadi Program at the University of Chicago Translational Fellowship to RG.

Data Availability:

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References:

1. Awad IA, Polster SP. Cavernous angiomas: deconstructing a neurosurgical disease. J Neurosurg. 2019;131:1–13. 10.3171/2019.3.JNS181724. [PubMed: 31261134]
2. Akers A, Al-Shahi Salman R, I AA, Dahlem K, Flemming K, Hart B, et al. 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. 2017;80:665–80. 10.1093/neuros/nyx091. [PubMed: 28387823]
3. Shenkar R, Shi C, Rebeiz T, Stockton RA, McDonald DA, Mikati AG, et al. Exceptional aggressiveness of cerebral cavernous malformation disease associated with PDCD10 mutations. Genet Med. 2015;17:188–96. 10.1038/gim.2014.97. [PubMed: 25122144]
4. Polster SP, Cao Y, Carroll T, Flemming K, Girard R, Hanley D, et al. Trial Readiness in Cavernous Angiomas With Symptomatic Hemorrhage (CASH). Neurosurgery. 2019;84:954–64. 10.1093/neuros/nyy108. [PubMed: 29660039]
5. Snellings DA, Hong CC, Ren AA, Lopez-Ramirez MA, Girard R, Srinath A, et al. Cerebral Cavernous Malformation: From Mechanism to Therapy. Circ Res. 2021;129:195–215. 10.1161/CIRCRESAHA.121.318174. [PubMed: 34166073]
6. Zhu W, Shen F, Mao L, Zhan L, Kang S, Sun Z, et al. Soluble FLT1 Gene Therapy Alleviates Brain Arteriovenous Malformation Severity. Stroke. 2017;48:1420–3. 10.1161/STROKEAHA.116.015713. [PubMed: 28325846]
7. Girard R, Li Y, Stadnik A, Shenkar R, Hobson N, Romanos S, et al. A Roadmap for Developing Plasma Diagnostic and Prognostic Biomarkers of Cerebral Cavernous Angioma With Symptomatic Hemorrhage (CASH). Neurosurgery. 2021;88:686–97. 10.1093/neuros/nyaa478. [PubMed: 33469662]
8. Mikati AG, Khanna O, Zhang L, Girard R, Shenkar R, Guo X, et al. Vascular permeability in cerebral cavernous malformations. J Cereb Blood Flow Metab. 2015;35:1632–9. 10.1038/jcbfm.2015.98. [PubMed: 25966944]
9. Zeineddine HA, Girard R, Cao Y, Hobson N, Fam MD, Stadnik A, et al. Quantitative susceptibility mapping as a monitoring biomarker in cerebral cavernous malformations with recent hemorrhage. J Magn Reson Imaging. 2018;47:1133–8. 10.1002/jmri.25831. [PubMed: 28791783]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

10. Girard R, Fam MD, Zeineddine HA, Tan H, Mikati AG, Shi C, et al. Vascular permeability and iron deposition biomarkers in longitudinal follow-up of cerebral cavernous malformations. J Neurosurg. 2017;127:102–10. 10.3171/2016.5.JNS16687. [PubMed: 27494817]
11. Huang Z, Ma L, Huang C, Li Q, Nice EC. Proteomic profiling of human plasma for cancer biomarker discovery. Proteomics. 2017;17. 10.1002/pmic.201600240.
12. Loke SY, Lee ASG. The future of blood-based biomarkers for the early detection of breast cancer. Eur J Cancer. 2018;92:54–68. 10.1016/j.ejca.2017.12.025. [PubMed: 29413690]
13. Koskimaki J, Girard R, Li Y, Saadat L, Zeineddine HA, Lightle R, et al. Comprehensive transcriptome analysis of cerebral cavernous malformation across multiple species and genotypes. JCI Insight. 2019;4. 10.1172/jci.insight.126167.
14. Koskimaki J, Polster SP, Li Y, Romanos S, Srinath A, Zhang D, et al. Common transcriptome, plasma molecules, and imaging signatures in the aging brain and a Mendelian neurovascular disease, cerebral cavernous malformation. Geroscience. 2020;42:1351–63. 10.1007/s11357-020-00201-4. [PubMed: 32556941]
15. Lopez-Ramirez MA, Pham A, Girard R, Wyseure T, Hale P, Yamashita A, et al. Cerebral cavernous malformations form an anticoagulant vascular domain in humans and mice. Blood. 2019;133:193–204. 10.1182/blood-2018-06-856062. [PubMed: 30442679]
16. Lopez-Ramirez MA, Fonseca G, Zeineddine HA, Girard R, Moore T, Pham A, et al. Thrombospondin1 (TSP1) replacement prevents cerebral cavernous malformations. J Exp Med. 2017;214:3331–46. 10.1084/jem.20171178. [PubMed: 28970240]
17. Jenny Zhou H, Qin L, Zhang H, Tang W, Ji W, He Y, et al. Endothelial exocytosis of angiopoietin-2 resulting from CCM3 deficiency contributes to cerebral cavernous malformation. Nat Med. 2016;22:1033–42. 10.1038/nm.4169. [PubMed: 27548575]
18. Wustehube J, Bartol A, Liebler SS, Brutsch R, Zhu Y, Felbor U, et al. Cerebral cavernous malformation protein CCM1 inhibits sprouting angiogenesis by activating DELTA-NOTCH signaling. Proc Natl Acad Sci U S A. 2010;107:12640–5. 10.1073/pnas.1000132107. [PubMed: 20616044]
19. Stockton RA, Shenkar R, Awad IA, Ginsberg MH. Cerebral cavernous malformations proteins inhibit Rho kinase to stabilize vascular integrity. J Exp Med. 2010;207:881–96. 10.1084/jem.20091258. [PubMed: 20308363]
20. Wetzel-Strong SE, Weinsheimer S, Nelson J, Pawlikowska L, Clark D, Starr MD, et al. Pilot investigation of circulating angiogenic and inflammatory biomarkers associated with vascular malformations. Orphanet J Rare Dis. 2021;16:372. 10.1186/s13023-021-02009-7. [PubMed: 34479577]
21. Koskimaki J, Zhang D, Li Y, Saadat L, Moore T, Lightle R, et al. Transcriptome clarifies mechanisms of lesion genesis versus progression in models of Ccm3 cerebral cavernous malformations. Acta Neuropathol Commun. 2019;7:132. 10.1186/s40478-019-0789-0. [PubMed: 31426861]
22. Correia de Sousa M, Gjorgjieva M, Dolicka D, Sobolewski C, Foti M. Deciphering miRNAs' Action through miRNA Editing. Int J Mol Sci. 2019;20. 10.3390/ijms20246249.
23. Backes C, Meese E, Keller A. Specific miRNA Disease Biomarkers in Blood, Serum and Plasma: Challenges and Prospects. Mol Diagn Ther. 2016;20:509–18. 10.1007/s40291-016-0221-4. [PubMed: 27378479]
24. Lugli G, Cohen AM, Bennett DA, Shah RC, Fields CJ, Hernandez AG, et al. Plasma Exosomal miRNAs in Persons with and without Alzheimer Disease: Altered Expression and Prospects for Biomarkers. PLoS One. 2015;10:e0139233. 10.1371/journal.pone.0139233. [PubMed: 26426747]
25. Kar S, Bali KK, Baisantry A, Geffers R, Samii A, Bertalanffy H. Genome-Wide Sequencing Reveals MicroRNAs Downregulated in Cerebral Cavernous Malformations. J Mol Neurosci. 2017;61:178–88. 10.1007/s12031-017-0880-6. [PubMed: 28181149]
26. Lyne SB, Girard R, Koskimaki J, Zeineddine HA, Zhang D, Cao Y, et al. Biomarkers of cavernous angioma with symptomatic hemorrhage. JCI Insight. 2019;4. 10.1172/jci.insight.128577.
27. Wendler A, Wehling M. The translatability of animal models for clinical development: biomarkers and disease models. Curr Opin Pharmacol. 2010;10:601–6. 10.1016/j.coph.2010.05.009. [PubMed: 20542730]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

28. Tang AT, Choi JP, Kotzin JJ, Yang Y, Hong CC, Hobson N, et al. Endothelial TLR4 and the microbiome drive cerebral cavernous malformations. Nature. 2017;545:305–10. 10.1038/nature22075. [PubMed: 28489816]
29. Morrison L, Akers A. Cerebral Cavernous Malformation, Familial. In: Adam MP, Ardinger HH, Pagon RA, Wallace SE, Bean LJH, Gripp KW, et al., editors. GeneReviews((R)). Seattle (WA)1993.
30. de Champfleur NM, Langlois C, Ankenbrandt WJ, Le Bars E, Leroy MA, Duffau H, et al. Magnetic resonance imaging evaluation of cerebral cavernous malformations with susceptibility-weighted imaging. Neurosurgery. 2011;68:641–7; discussion 7-8. 10.1227/NEU.0b013e31820773cf. [PubMed: 21164377]
31. Blondal T, Brunetto MR, Cavallone D, Mikkelsen M, Thorsen M, Mang Y, et al. Genome-Wide Comparison of Next-Generation Sequencing and qPCR Platforms for microRNA Profiling in Serum. Methods Mol Biol. 2017;1580:21–44. 10.1007/978-1-4939-6866-4_3. [PubMed: 28439824]
32. Potla P, Ali SA, Kapoor M. A bioinformatics approach to microRNA-sequencing analysis. Osteoarthritis and Cartilage Open. 2021;3:100131. https:/doi.org/10.1016/j.ocarto.2020.100131. [PubMed: 36475076]
33. Decker JT, Hall MS, Blaisdell RB, Schwark K, Jeruss JS, Shea LD. Dynamic microRNA activity identifies therapeutic targets in trastuzumab-resistant HER2(+) breast cancer. Biotechnol Bioeng. 2018;115:2613–23. 10.1002/bit.26791. [PubMed: 29981261]
34. Kim SH, MacIntyre DA, Binkhamis R, Cook J, Sykes L, Bennett PR, et al. Maternal plasma miRNAs as potential biomarkers for detecting risk of small-for-gestational-age births. EBioMedicine. 2020;62:103145. 10.1016/j.ebiom.2020.103145. [PubMed: 33260001]
35. Johnson JJ, Loeffert AC, Stokes J, Olympia RP, Bramley H, Hicks SD. Association of Salivary MicroRNA Changes With Prolonged Concussion Symptoms. JAMA Pediatr. 2018;172:65–73. 10.1001/jamapediatrics.2017.3884. [PubMed: 29159407]
36. Rohart F, Gautier B, Singh A, Le Cao KA. mixOmics: An R package for 'omics feature selection and multiple data integration. PLoS Comput Biol. 2017;13:e1005752. 10.1371/journal.pcbi.1005752. [PubMed: 29099853]
37. Dweep H, Gretz N, Sticht C. miRWalk database for miRNA-target interactions. Methods Mol Biol. 2014;1182:289–305. 10.1007/978-1-4939-1062-5_25. [PubMed: 25055920]
38. Sticht C, De La Torre C, Parveen A, Gretz N. miRWalk: An online resource for prediction of microRNA binding sites. PLoS One. 2018;13:e0206239. 10.1371/journal.pone.0206239. [PubMed: 30335862]
39. Dweep H, Sticht C, Pandey P, Gretz N. miRWalk--database: prediction of possible miRNA binding sites by "walking" the genes of three genomes. J Biomed Inform. 2011;44:839–47. 10.1016/j.jbi.2011.05.002. [PubMed: 21605702]
40. Sulakhe D, Balasubramanian S, Xie B, Feng B, Taylor A, Wang S, et al. Lynx: a database and knowledge extraction engine for integrative medicine. Nucleic Acids Res. 2014;42:D1007–12. 10.1093/nar/gkt1166. [PubMed: 24270788]
41. Amur S, LaVange L, Zineh I, Buckman-Garner S, Woodcock J. Biomarker Qualification: Toward a Multiple Stakeholder Framework for Biomarker Development, Regulatory Acceptance, and Utilization. Clin Pharmacol Ther. 2015;98:34–46. 10.1002/cpt.136. [PubMed: 25868461]
42. Johnston S, Gallaher Z, Czaja K. Exogenous reference gene normalization for real-time reverse transcription-polymerase chain reaction analysis under dynamic endogenous transcription. Neural Regen Res. 2012;7:1064–72. 10.3969/j.issn.1673-5374.2012.14.004. [PubMed: 25722696]
43. Dakterzada F, Targa A, Benitez ID, Romero-ElKhayat L, de Gonzalo-Calvo D, Torres G, et al. Identification and validation of endogenous control miRNAs in plasma samples for normalization of qPCR data for Alzheimer's disease. Alzheimers Res Ther. 2020;12:163. 10.1186/s13195-020-00735-x. [PubMed: 33278902]
44. Andersen CL, Jensen JL, Orntoft TF. Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. Cancer Res. 2004;64:5245–50. 10.1158/0008-5472.CAN-04-0496. [PubMed: 15289330]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

45. Pfaffl MW. A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res. 2001;29:e45. 10.1093/nar/29.9.e45. [PubMed: 11328886]
46. Ruiz-Villalba A, Ruijter JM, van den Hoff MJB. Use and Misuse of Cq in qPCR Data Analysis and Reporting. Life (Basel). 2021;11. 10.3390/life11060496.
47. Selst MV, Jolicoeur P. A solution to the effect of sample size on outlier elimination. The quarterly journal of experimental psychology. 1994;47:631–50.
48. Davies L, Gather U. The identification of multiple outliers. Journal of the American Statistical Association. 1993;88:782–92.
49. Yuan JS, Reed A, Chen F, Stewart CN Jr. Statistical analysis of real-time PCR data. BMC Bioinformatics. 2006;7:85. 10.1186/1471-2105-7-85. [PubMed: 16504059]
50. Leti F, DiStefano JK. miRNA Quantification Method Using Quantitative Polymerase Chain Reaction in Conjunction with C q Method. Methods Mol Biol. 2018;1706:257–65. 10.1007/978-1-4939-7471-9_14. [PubMed: 29423803]
51. Peng F, He Ja, Loo JFC, Kong SK, Li B, Gu D. Identification of serum MicroRNAs as diagnostic biomarkers for influenza H7N9 infection. Virology Reports. 2017;7:1–8. https:/doi.org/10.1016/j.virep.2016.11.001.
52. Dridi N, Giremus A, Giovannelli JF, Truntzer C, Hadzagic M, Charrier JP, et al. Bayesian inference for biomarker discovery in proteomics: an analytic solution. EURASIP J Bioinform Syst Biol. 2017;2017:9. 10.1186/s13637-017-0062-4. [PubMed: 28710702]
53. Girard R, Zeineddine HA, Koskimaki J, Fam MD, Cao Y, Shi C, et al. Plasma Biomarkers of Inflammation and Angiogenesis Predict Cerebral Cavernous Malformation Symptomatic Hemorrhage or Lesional Growth. Circ Res. 2018;122:1716–21. 10.1161/CIRCRESAHA.118.312680. [PubMed: 29720384]
54. Girard R, Zeineddine HA, Fam MD, Mayampurath A, Cao Y, Shi C, et al. Plasma Biomarkers of Inflammation Reflect Seizures and Hemorrhagic Activity of Cerebral Cavernous Malformations. Transl Stroke Res. 2018;9:34–43. 10.1007/s12975-017-0561-3. [PubMed: 28819935]
55. Arlot S, Celisse A. A survey of cross-validation procedures for model selection. Statistics Surveys. 2010;4:40–79, 40.
56. Xia J, Broadhurst DI, Wilson M, Wishart DS. Translational biomarker discovery in clinical metabolomics: an introductory tutorial. Metabolomics. 2013;9:280–99. 10.1007/s11306-012-0482-9. [PubMed: 23543913]
57. Lien EC, Dibble CC, Toker A. PI3K signaling in cancer: beyond AKT. Curr Opin Cell Biol. 2017;45:62–71. 10.1016/j.ceb.2017.02.007. [PubMed: 28343126]
58. Alzahrani AS. PI3K/Akt/mTOR inhibitors in cancer: At the bench and bedside. Semin Cancer Biol. 2019;59:125–32. 10.1016/j.semcancer.2019.07.009. [PubMed: 31323288]
59. Chen H, Zhou L, Wu X, Li R, Wen J, Sha J, et al. The PI3K/AKT pathway in the pathogenesis of prostate cancer. Front Biosci (Landmark Ed). 2016;21:1084–91. 10.2741/4443. [PubMed: 27100493]
60. Kar S, Samii A, Bertalanffy H. PTEN/PI3K/Akt/VEGF signaling and the cross talk to KRIT1, CCM2, and PDCD10 proteins in cerebral cavernous malformations. Neurosurg Rev. 2015;38:229–36; discussion 36-7. 10.1007/s10143-014-0597-8. [PubMed: 25403688]
61. Martinez-Lopez A, Salvador-Rodriguez L, Montero-Vilchez T, Molina-Leyva A, Tercedor-Sanchez J, Arias-Santiago S. Vascular malformations syndromes: an update. Curr Opin Pediatr. 2019;31:747–53. 10.1097/MOP.0000000000000812. [PubMed: 31693582]
62. Ren AA, Snellings DA, Su YS, Hong CC, Castro M, Tang AT, et al. PIK3CA and CCM mutations fuel cavernomas through a cancer-like mechanism. Nature. 2021. 10.1038/s41586-021-03562-8.
63. Hong T, Xiao X, Ren J, Cui B, Zong Y, Zou J, et al. Somatic MAP3K3 and PIK3CA mutations in sporadic cerebral and spinal cord cavernous malformations. Brain. 2021. 10.1093/brain/awab117.
64. Weng J, Yang Y, Song D, Huo R, Li H, Chen Y, et al. Somatic MAP3K3 mutation defines a subclass of cerebral cavernous malformation. Am J Hum Genet. 2021;108:942–50. 10.1016/j.ajhg.2021.04.005. [PubMed: 33891857]
65. Lisowska J, Rodel CJ, Manet S, Miroshnikova YA, Boyault C, Planus E, et al. The CCM1-CCM2 complex controls complementary functions of ROCK1 and ROCK2 that are required for endothelial integrity. J Cell Sci. 2018;131. 10.1242/jcs.216093.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

66. McKerracher L, Shenkar R, Abbinanti M, Cao Y, Peiper A, Liao JK, et al. A Brain-Targeted Orally Available ROCK2 Inhibitor Benefits Mild and Aggressive Cavernous Angioma Disease. Transl Stroke Res. 2020;11:365–76. 10.1007/s12975-019-00725-8. [PubMed: 31446620]

67. Wei S, Li Y, Polster SP, Weber CR, Awad IA, Shen L. Cerebral Cavernous Malformation Proteins in Barrier Maintenance and Regulation. Int J Mol Sci. 2020;21. 10.3390/ijms21020675.

68. Polster SP, Stadnik A, Akers AL, Cao Y, Christoforidis GA, Fam MD, et al. Atorvastatin Treatment of Cavernous Angiomas with Symptomatic Hemorrhage Exploratory Proof of Concept (AT CASH EPOC) Trial. Neurosurgery. 2019;85:843–53. 10.1093/neuros/nyy539. [PubMed: 30476251]

69. Zhou Z, Tang AT, Wong WY, Bamezai S, Goddard LM, Shenkar R, et al. Cerebral cavernous malformations arise from endothelial gain of MEKK3-KLF2/4 signalling. Nature. 2016;532:122–6. 10.1038/nature17178. [PubMed: 27027284]

70. Lopez-Ramirez MA, Lai CC, Soliman SI, Hale P, Pham A, Estrada EJ, et al. Astrocytes propel neurovascular dysfunction during cerebral cavernous malformation lesion formation. J Clin Invest. 2021;131. 10.1172/JCI139570.

71. Park JH, Lee JY, Shin DH, Jang KS, Kim HJ, Kong G. Loss of Mel-18 induces tumor angiogenesis through enhancing the activity and expression of HIF-1alpha mediated by the PTEN/PI3K/Akt pathway. Oncogene. 2011;30:4578–89. 10.1038/onc.2011.174. [PubMed: 21602890]

72. Hong CC, Tang AT, Detter MR, Choi JP, Wang R, Yang X, et al. Cerebral cavernous malformations are driven by ADAMTS5 proteolysis of versican. J Exp Med. 2020;217. 10.1084/jem.20200140.

73. Liu M, Banerjee R, Rossa C Jr., D'Silva NJ. RAP1-RAC1 Signaling Has an Important Role in Adhesion and Migration in HNSCC. J Dent Res. 2020;99:959–68. 10.1177/0022034520917058. [PubMed: 32401565]

74. Rho SS, Ando K, Fukuhara S. Dynamic Regulation of Vascular Permeability by Vascular Endothelial Cadherin-Mediated Endothelial Cell-Cell Junctions. J Nippon Med Sch. 2017;84:148–59. 10.1272/jnms.84.148. [PubMed: 28978894]

75. Li X, Zhang R, Draheim KM, Liu W, Calderwood DA, Boggon TJ. Structural basis for small G protein effector interaction of Ras-related protein 1 (Rap1) and adaptor protein Krev interaction trapped 1 (KRIT1). J Biol Chem. 2012;287:22317–27. 10.1074/jbc.M112.361295. [PubMed: 22577140]

76. Wu M, Wang G, Tian W, Deng Y, Xu Y. MiRNA-based Therapeutics for Lung Cancer. Curr Pharm Des. 2018;23:5989–96. 10.2174/1381612823666170714151715. [PubMed: 28714413]

77. Rupaimoole R, Slack FJ. MicroRNA therapeutics: towards a new era for the management of cancer and other diseases. Nat Rev Drug Discov. 2017;16:203–22. 10.1038/nrd.2016.246. [PubMed: 28209991]

78. Burns TC, Li MD, Mehta S, Awad AJ, Morgan AA. Mouse models rarely mimic the transcriptome of human neurodegenerative diseases: A systematic bioinformatics-based critique of preclinical models. Eur J Pharmacol. 2015;759:101–17. 10.1016/j.ejphar.2015.03.021. [PubMed: 25814260]

79. Mullane K, Williams M. Preclinical Models of Alzheimer's Disease: Relevance and Translational Validity. Curr Protoc Pharmacol. 2019;84:e57. 10.1002/cpph.57. [PubMed: 30802363]

80. Detter MR, Shenkar R, Benavides CR, Neilson CA, Moore T, Lightle R, et al. Novel Murine Models of Cerebral Cavernous Malformations. Angiogenesis. 2020;23:651–66. 10.1007/s10456-020-09736-8. [PubMed: 32710309]

81. Zeineddine HA, Girard R, Saadat L, Shen L, Lightle R, Moore T, et al. Phenotypic characterization of murine models of cerebral cavernous malformations. Lab Invest. 2019;99:319–30. 10.1038/s41374-018-0030-y. [PubMed: 29946133]

82. Shenkar R, Venkatasubramanian PN, Wyrwicz AM, Zhao JC, Shi C, Akers A, et al. Advanced magnetic resonance imaging of cerebral cavernous malformations: part II. Imaging of lesions in murine models. Neurosurgery. 2008;63:790–7; discussion 7-8. 10.1227/01.NEU.0000315862.24920.49. [PubMed: 18981891]

83. Nagy ZB, Bartak BK, Kalmar A, Galamb O, Wichmann B, Dank M, et al. Comparison of Circulating miRNAs Expression Alterations in Matched Tissue and Plasma Samples During Colorectal Cancer Progression. Pathol Oncol Res. 2019;25:97–105. 10.1007/s12253-017-0308-1. [PubMed: 28980150]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

84. Malczewska A, Kidd M, Matar S, Kos-Kudla B, Modlin IM. A Comprehensive Assessment of the Role of miRNAs as Biomarkers in Gastroenteropancreatic Neuroendocrine Tumors. Neuroendocrinology. 2018;107:73–90. 10.1159/000487326. [PubMed: 29566385]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

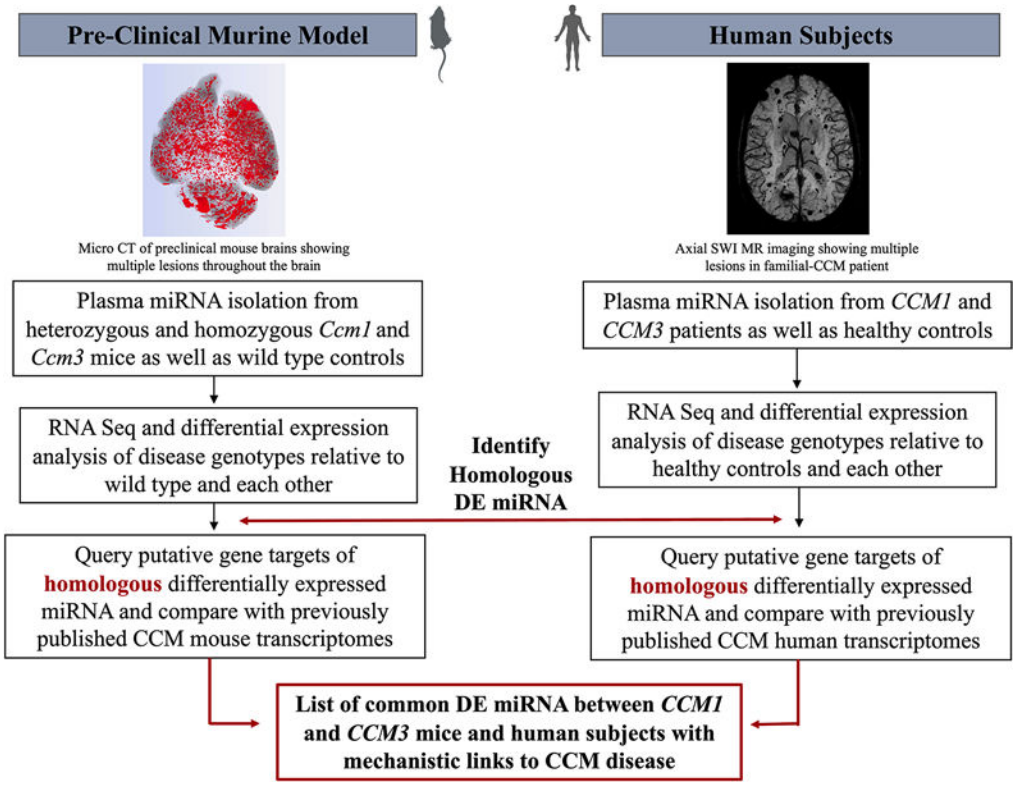

Fig. 1. Methodological Overview for Identification of Mechanistically Relevant Homologous Differentially Expressed (DE) miRNAs.

A parallel methodological pipeline was applied for murine models (N=41) and human subjects (N=30) to identify DE miRNA in the plasma of *CCM1* (n=14 mouse, n=13 human) and *CCM3* (n=12 mouse, n=10 human) genotypes compared to healthy controls (n=7), wild type (n=15), and each other. Once homologous DE miRNA were identified using the MirGeneDB database, MirWalk 3.0 was used to query putative gene targets of each homologous DE miRNA in mice and then in patients. Gene targets were compared to DE genes within previously published CCM transcriptomes, including the transcriptomes of mouse *in vitro* $Ccm1^{-/-}$ BMECs, and $Ccm3^{-/-}$ brain microvascular endothelial cells (BMECs), as well as *in vivo* $Ccm3^{+/-}$ lesional neurovascular units (NVUs) and laser micro-dissected NVUs of human surgically resected CCM lesions. Illustrative images represent microCT of $Ccm3^{-/-}$ mouse brain, and SWI (susceptibility weighted imaging) MRI of human brain with CCM disease and *CCM3* genotype.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

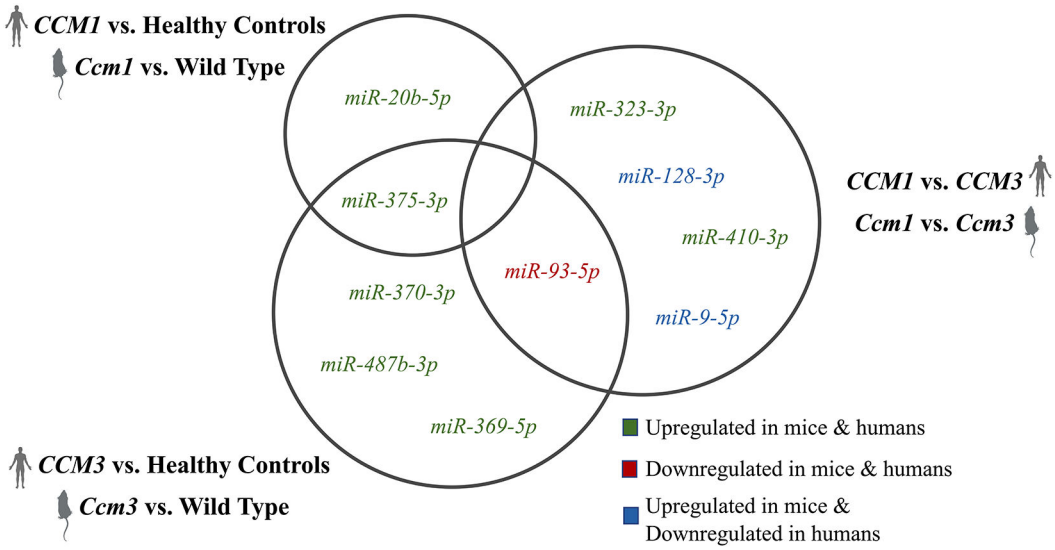

Fig. 2. Homologous Differentially Expressed (DE) Plasma miRNAs in Mouse and Human *CCM1* and *CCM3* Genotypes.

This was constructed to portray commonalities in DE miRNAs between human and mice in each differential expression comparison made. Ten DE plasma miRNAs were identified as homologous between mice (N=41) and human (N=30) *CCM1* (n=7 mice, n=13 human) and *CCM3* (n=7 mice, n=10 human) genotypes and non-CCM (n=7 mice, n=7 human) controls. *miR-375-3p* was upregulated in *CCM1* patients and *Ccm1* mice as well as in *CCM3* patients and *Ccm3* mice compared to healthy controls (HC) and wild type (WT) mice. *miR-93-5p* was downregulated in *CCM3* patients and *Ccm3* mice compared to HC and WT mice. It was also downregulated in *CCM1* patients and *Ccm1* mice compared to *CCM3* patients and *Ccm3* mice. *miR-20b-5p* was upregulated in *CCM1* patients and *Ccm1* mice compared to HC and WT mice. *miR-370-3p*, *miR-487b-3p*, and *miR-369-5p* were upregulated in *CCM3* patients and *Ccm3* mice compared to HC and WT mice. *miR-410-3p* and *miR-323-3p* were upregulated in *CCM1* patients and *Ccm1* mice compared to *CCM3* patients and *Ccm3* mice. *miR-128-3p* and *miR-9-5p* were upregulated in *Ccm1* compared to *Ccm3* in mice but downregulated in *CCM1* compared to *CCM3* in patients.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

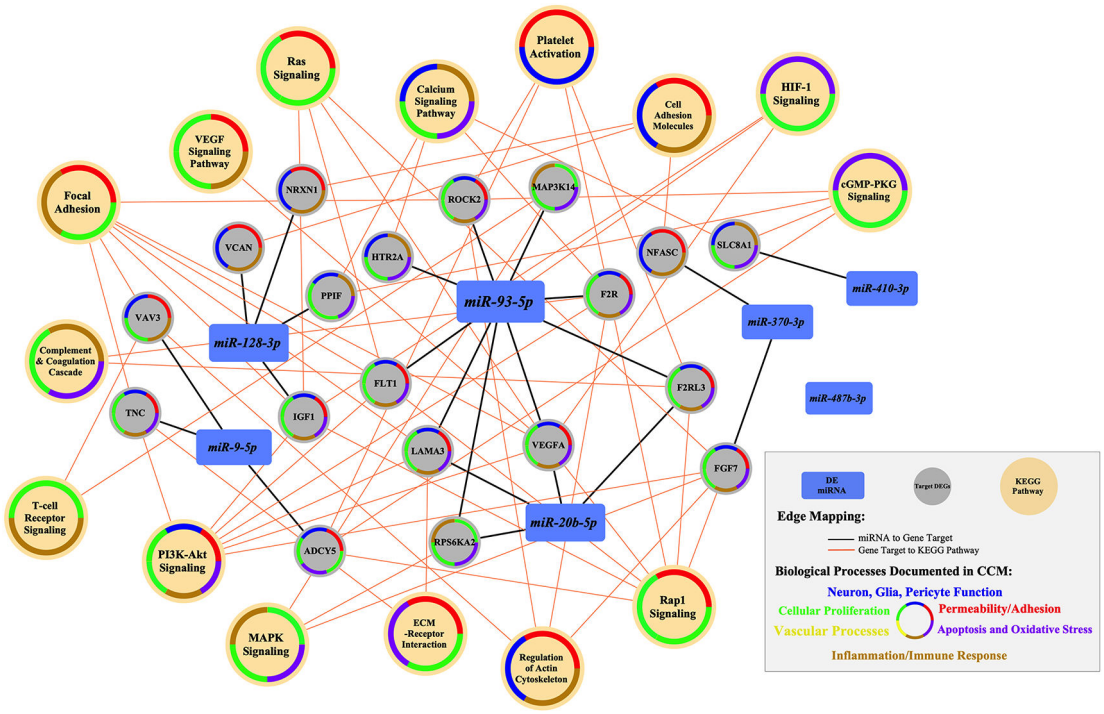

Fig. 3. Homologous Differentially Expressed (DE) Plasma miRNAs of *CCM1* and *CCM3* Target Genes Associated with CCM-Enriched KEGG Pathways.

Seven of the ten homologous DE plasma miRNAs identified between mice (N=41) and humans (N=30) target DE genes [|FC|>1.5; p<0.05, FDR corrected] within the previously published transcriptomes of (1) laser micro-dissected NVUs of human surgically resected CCM lesions, and/or (2) *in vitro Ccm1*$^{-/-}$ mouse BMECs, (2) *in vitro Ccm3*$^{-/-}$ mouse BMECs, as well as within (3) laser micro-dissected lesional NVUs of *Ccm3*$^{+/-}$ mice [13, 21]. The color-coded enriched KEGG pathways [p<0.05, FDR corrected; Bayes factor≥3] presented were limited to (1) those related to CCM processes identified through a comprehensive literature search and (2) linked to at least two homologous DE miRNAs. These biological processes include: (1) Neuron, Glia, Pericyte Function, (2) Permeability/Adhesion, (3) Apoptosis and Oxidative Stress, (4) Inflammation/Immune Response, (5) Vascular Processes, (6) Cellular Proliferation. Additionally, only DE genes linked to these KEGG pathways are shown. Homologous DE miRNAs are shown with a gradual increase in size for an increasing number of connections to DE genes.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

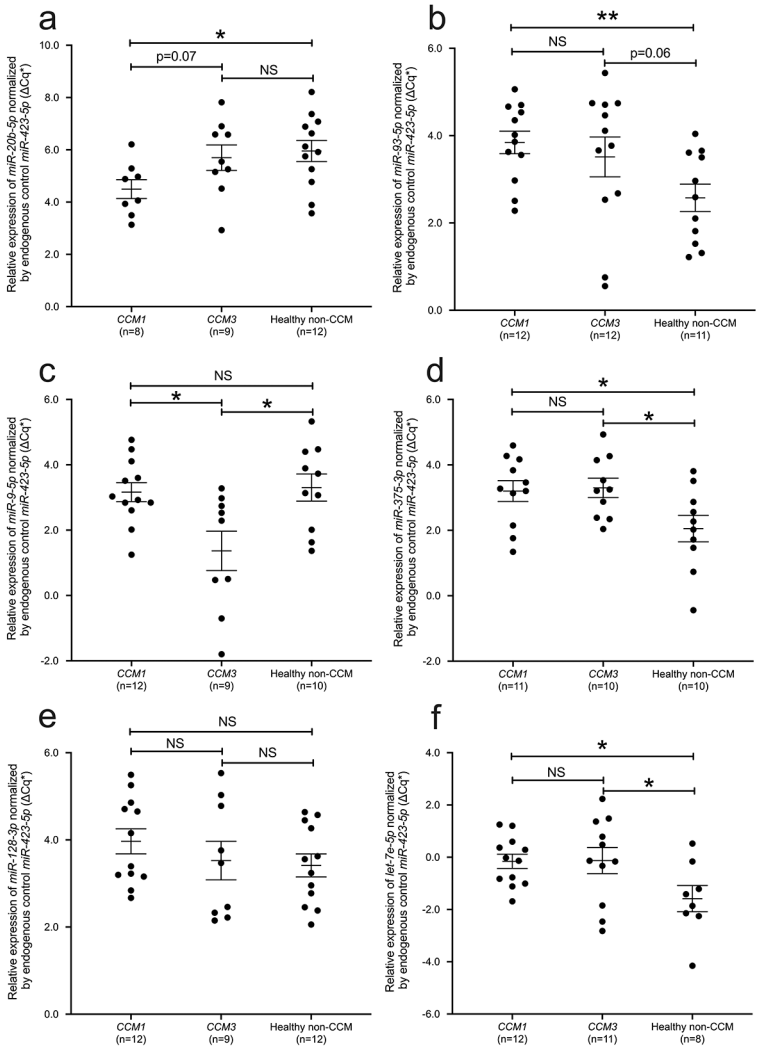

Fig. 4. Relative Quantification of Differentially Expressed (DE) miRNAs through RT-qPCR.

Validation of miRNome results were performed on a select panel of DE miRNAs in an independent cohort (N=37) of familial-CCM patients along with healthy non-CCM controls using RT-qPCR. The Cq value is inversely related with the plasma concentration of miRNA, therefore ΔCq values are opposite to the fold change. The averaged ΔCq of (**a**) *miR-20b-5p* was higher in non-CCM subjects (n=12) and *CCM3* patients (n=9) compared to *CCM1* patients (n=8), (**b**) *miR-93-5p* was higher in *CCM1* (n=12) and *CCM3* (n=12) patients compared to non-CCM subjects (n=11), (**c**) *miR-9-5p* was higher in non-CCM subjects (n=10) and *CCM1* patients (n=12) compared to *CCM3* patients (n=9), (**d**) *miR-375-3p* was higher in *CCM1* (n=11) and *CCM3* (n=10) patients compared to non-CCM subjects (n=10), (**e**) *miR-128-3p* showed no significant differences between groups (n=12 *CCM1*, n=9 *CCM3*, n=12 non-CCM) , (**f**) *let-7e-5p* levels was lower in *CCM1* (n=12) and *CCM3* (n=11) patients compared to non-CCM subjects (n=8). Statistical analyses were performed using a Mann-Whitney test. Values greater than 3 standard deviations from the mean were excluded as outliers. *p<0.05; **p<0.01; NS=not significant

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

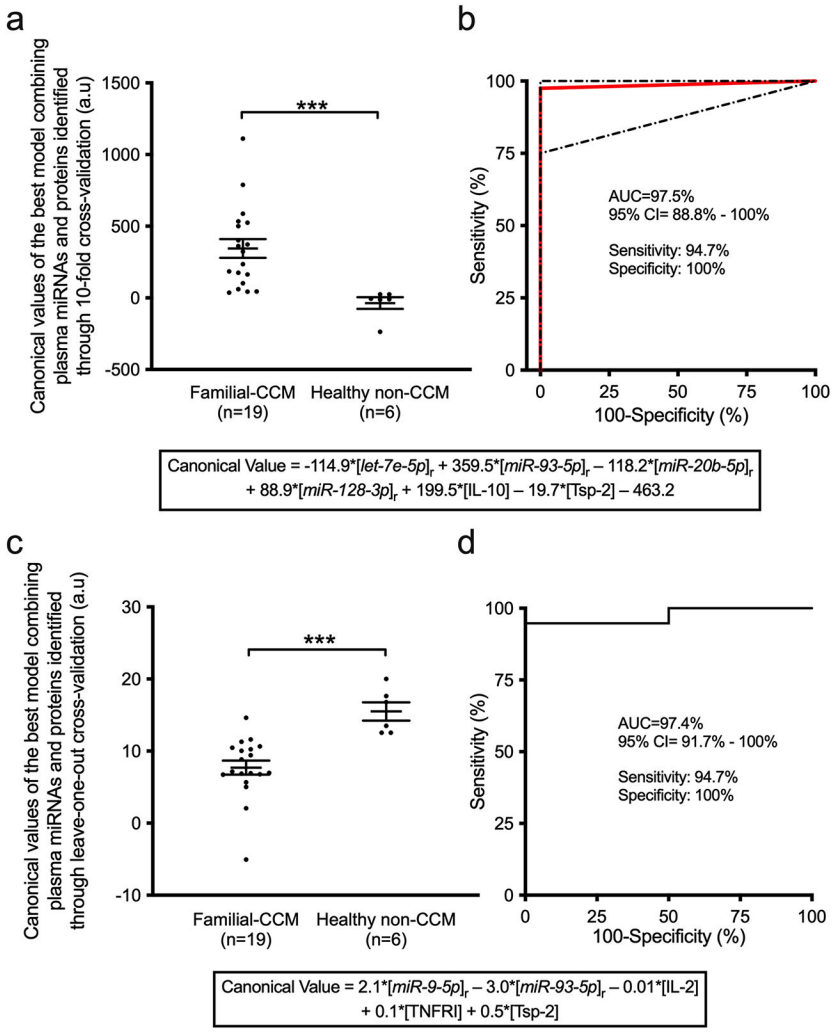

Fig. 5. Integrated Plasma miRNA and Protein Biomarker and Diagnosis of Familial CCM.

Two machine learning approaches were applied to integrate plasma miRNAs and proteins into a diagnostic biomarker for CCM and identify the best model with the lowest Akaike information criterion in 25 subjects. (**a**) Using a 10-fold cross validation method, the best weighted combination of molecules to diagnose familial CCM (n=19) from non-CCM (n=6) subjects included plasma relative quantification values of *let-7e-5p, miR-93-5p, miR-20b-5p, miR-128-3p*, and plasma concentrations of interleukin-10 (IL-10) and thrombospondin-2 (Tsp-2). (**b**) Receiver operating characteristic (ROC) analysis for this model yielded a 97.5% area under the curve (AUC), with a sensitivity of 94.7% and specificity of 100%. The red curve represents the average ROC of the cross validation, while the dash-dotted lines represent the unique ROC curves generated for each of the 10 iterations within the 10-fold cross validation. (**c**) Using a leave-one-out cross validation method, the best weighted combination of molecules to diagnose familial CCM (n=19) from non-CCM (n=6) subjects included plasma relative quantification values of *miR-93-5p, miR-9-5p*, and plasma concentrations of interleukin-2 (IL-2), tumor necrosis factor receptor 1 (TNFRI), and Tsp-2. (**d**) ROC analysis for this model yielded a 97.4% AUC, with a sensitivity of 94.7% and a specificity of 100%. $[X]_r$ denotes the relative quantification value for each miRNA.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Statistical analyses to compare the canonical values were implemented using an unpaired two samples Student's t-test, assessed with pooled standard deviation, or Mann-Whitney test according to the equality of the variance. ***$p<0.001$

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1.

Cohort Demographics of 23 CCM Patients (n=13 *CCM1*, n=10 *CCM3*) and Healthy Controls (n=7) for miRNome Sequencing in N=30 subjects.

Patient Characteristics	**CCM1**	**CCM3**	**Healthy non-CCM**
Sample Size	13	10	7
Age (mean ± SD)	30.9 ± 21.9	24.1 ± 13.5	22.1 ± 4.1
Range (yr.)	[4.3 – 59.5]	[10.2 - 62.3]	[19.1 – 30.6]
Female (%)	53.8	60.0	57.1
Lesion Characteristics			
Number of SWI-weighted lesions (mean ± SD)	32.2 ± 39.4	44.0 ± 26.5	NA
Range	[4 -105]	[10 -101]	
Number of T_2-weighted lesions (mean ± SD)	7.1 ± 7.4	18.9 ± 11.0	NA
Range	[2 - 23]	[5 - 39]	
Patients with T_2 brainstem lesions (%)	23.0	10.0	NA
Ethnicity			
White/European (%)	53.8	100	28.6
African American (%)	23.1	0	14.2
Hispanic (%)	23.1	0	28.6
Asian (%)	0	0	28.6

SWI=susceptibility weighted imaging; SD=standard deviation, NA=not applicable

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2.

Cohort Demographics for 37 subjects of 24 CCM Patients (n=12 *CCM1*, n=12 *CCM3*) and Healthy Controls (n=13) for RT-qPCR Validation.

Patient Characteristics	CCM1	CCM3	Healthy non-CCM
Sample Size	12	12	13
Age (mean ± SD)	43.7 ± 20.8	28.5 ± 19.0	39.7 ± 14.7
Range (yr.)	[12.2 - 67.5]	[7.0 - 64.7]	[19.3 - 63.6]
Female (%)	83.3	41.7	53.8
Lesion Characteristics			
Number of SWI-weighted lesions (mean ± SD)	49.3 ± 46.3	36.2 ± 33.0	NA
Range	[1 -106]	[9 -101]	
Number of T_2-weighted lesions (mean ± SD)	6.9 ± 7.6	11.0 ± 6.6	NA
Range	[0 - 23]	[3 - 22]	
Patients with T_2 brainstem lesions (%)	33.3	8.3	NA
Ethnicity			
White/European (%)	91.7	91.7	53.5
African American (%)	8.3	0	23.0
Hispanic (%)	0	0	8.2
Asian (%)	0	8.3	15.3

SWI=susceptibility weighted imaging; SD=standard deviation, NA=not applicable