# Mutations in chromatin modifiers and Ephrin signaling components in Vein of Galen malformation

Daniel Duran[1*], Jungmin Choi[2,22*], Xue Zeng[2*], Sheng Chih Jin[2,22*], Carol Nelson-Williams[2], Bogdan Yatsula[4], Jonathan Gaillard[1], Charuta Gavankar Furey[1], Andrew T. Timberlake[2], Weilai Dong[2], Michelle A. Sorscher[5], Erin Loring[2], Jennifer Klein[6], Qiongshi Lu[7], August Allocco[1], Ava Hunt[1], Jason K. Karimy[1], Mark W. Youngblood[1,2], Jinwei Zhang[8], Michael L. DiLuna[1], Charles C. Matouk[1], Shrikant Mane[9], Irina Tikhonova[9], Christopher Castaldi[9], Francesc López-Giráldez[9], James Knight[9], Shozeb Haider[10], Mariya Soban[10,11], Seth L. Alper[12], Masaki Komiyama[13], Andrew F. Ducruet[14], Joseph M. Zabramski[14], Alan Dardik[4], Brian P. Walcott[15], Christopher J. Stapleton[16], Beverly Aagaard-Kienitz[17], Georges Rodesch[18], Eric Jackson[19], Edward R. Smith[20], Darren B. Orbach[6,20], Alejandro Berenstein[5], Kaya Bilguvar[2,9], Miikka Vikkula[21], Murat Gunel[1,2], Richard P. Lifton[2,22#,] and Kristopher T. Kahle[1,23,24#]

[1]Department of Neurosurgery, Yale School of Medicine, New Haven CT, USA.
[2]Department of Genetics, Yale School of Medicine, New Haven CT, USA.
[3]Howard Hughes Medical Institute, Chevy Chase MD, USA.
[4]Department of Surgery, Yale School of Medicine, New Haven CT, USA.
[5]Department of Neurosurgery, Icahn School of Medicine at Mount Sinai, New York NY, USA.
[6]Department of Neurosurgery, Boston Children's Hospital, Boston MA, USA.
[7]Department of Biostatistics & Medical Informatics, University of Wisconsin-Madison, Madison WI, USA.
[8]Institute of Biomedical and Clinical Sciences, University of Exeter Medical School, Hatherly Laboratory, Exeter, UK.
[9]Yale Center for Genome Analysis, West Haven CT, USA.
[10]University College London, School of Pharmacy, London, England
[11]Department of Biochemistry, Aligarh Muslim University, Aligarh, India
[12]Division of Nephrology and Center for Vascular Biology Research, Beth Israel Deaconess Medical Center; and Department of Medicine, Harvard Medical School, Boston, MA USA.
[13]Department of Neurointervention, Osaka City General Hospital, Osaka, Japan.
[14]Department of Neurosurgery, Barrow Neurological Institute, Phoenix AZ, USA.
[15]Department of Neurological Surgery, University of Southern California, Los Angeles, CA, USA.
[16]Department of Neurological Surgery, Massachusetts General Hospital and Harvard Medical School, Boston, MA USA.
[17]Department of Neurological Surgery, University of Wisconsin, Madison, Wisconsin
[18]Service de Neuroradiologie Diagnostique et Thérapeutique, Hôpital Foch, Suresnes, France.
[19]Department of Neurosurgery, Johns Hopkins University School of Medicine, Baltimore, MD USA.
[20]Department of Neurointerventional Radiology, Boston Children's Hospital, Boston MA, USA.
[21]Human Molecular Genetics, de Duve Institute, Université catholique de Louvain, Brussels, Belgium.
[22]Laboratory of Human Genetics and Genomics, The Rockefeller University, New York, NY, USA.
[23]Department of Pediatrics, Yale School of Medicine, New Haven CT, USA.
[24]Department of Cellular & Molecular Physiology, Yale School of Medicine, New Haven CT, USA.

*Equal contributors
[#]Corresponding authors

Lead contact: kristopher.kahle@yale.edu

## ABSTRACT

**Normal vascular development includes the formation and specification of arteries, veins, and intervening capillaries. Vein of Galen malformations (VOGM) are often catastrophic congenital cerebrovascular anomalies, shunting arterial blood into the brain's deep venous system through aberrant direct connections. Exome sequencing of 55 VOGM probands, including 52 parent-offspring trios, revealed VOGM patients harbor an excess of rare damaging *de novo* mutations, with enrichment in chromatin modifiers that regulate brain and vascular development. VOGM probands also exhibited a significant excess of rare inherited damaging mutations in Ephrin signalling genes and other genes implicated in Mendelian AVM syndromes. *EPHB4* and *CLDN14* each showed a significant mutation burden. Inherited mutations showed incomplete penetrance and variable expressivity, with mutation carriers often exhibiting cutaneous vascular abnormalities, suggesting a two-hit mechanism. Collectively, identified mutations can account for ~30% of VOGM cases studied. These findings have implications for risk assessment and provide insight into disease biology.**

## INTRODUCTION

During early embryogenesis, the vasculature becomes highly specialized to meet the hemodynamic and nutritive demands of tissues. Studies in vertebrate model organisms have shown that arterio-venous (A-V) specification is determined by a genetic program that results in differential gene expression and differentiation of arteries and veins prior to the establishment of circulation (Fish and Wythe, 2015). For example, Ephrin-B2 and its cognate receptor tyrosine kinase Eph-B4 are exclusively expressed in either arteries or veins, respectively, prior to blood flow in the developing embryo (Wang et al., 1998; Adams et al., 1999; Gerety et al., 1999), and deletion of *Efnb2* or *EphB4* in mouse results in impaired A-V specification, the development of arterio-venous malformations (AVMs), and embryonic lethality in mice. Nonetheless, A-V specification is complex, and involves carefully orchestrated activity of multiple signaling cascades (e.g., Eph-Ephrin, Hedgehog, VEGF, TGF-beta, Wnt, and Notch) and transcriptional networks (e.g., HEY and HES, SOX factors, Forkhead factors, COUP-TFII, and others)(Fish and Wythe, 2015). The genetic determinants of neurovascular development and A-V specification in humans are incompletely understood.

Arteriovenous malformations (AVMs) resulting from failure of normal specification and differentiation of arteries and veins occur in humans, and result in high-flow vascular malformations characterized by direct connection of arteries to veins without intervening capillaries (Spetzler and Martin, 1986). During normal brain development, primitive choroidal and subependymal arteries that perfuse deep brain structures are connected via an intervening capillary network to the embryonic precursor of the Vein of Galen, known as the median prosencephalic vein of Markowski (MPV). The MPV is responsible for returning the majority of deep cerebral venous blood (~40% of total blood flow in the brain) to the dural sinuses that ultimately drain into the internal jugular veins (Raybaud et al., 1989). Vein of Galen malformations (VOGMs) affect ~1/25,000 live births, account for up to 30% of all pediatric cerebral AVMs, and are among the most severe congenital cerebrovascular anomalies (Long et al., 1974; Deloison et al., 2012). VOGMs directly connect primitive choroidal or subependymal cerebral arteries to the MPV, exposing it to dangerously high blood flow and pressures, often with catastrophic consequences. VOGMs are often first detected prenatally and become symptomatic in the neonatal period. If untreated, VOGMs result in high morbidity and mortality due to high-output cardiac failure, hydrocephalus, and/or brain hemorrhage from blood vessel rupture (Recinos et al., 2012). Additionally, a significant percentage of VOGM patients have congenital heart disease (CHD), including sinus venosus atrial septal defects and aortic coarctation (McElhinney et al., 1998). Although VOGM treatment has greatly benefited from advances in endovascular therapies that partially obliterate the anomalous A-V connection (Mitchell et al., 2001), (Altschul et al., 2014), the lesions of many patients remain refractory to intervention. Moreover, the highly specialized care centers and physicians that perform these interventions remain inaccessible to many patients.

Limited knowledge of the genetics and molecular mechanisms of VOGM has hindered development of early diagnostic and targeted therapeutic strategies (Duran et al., 2018). Although classically considered isolated, sporadic congenital lesions (Duran et al., 2018), VOGMs have rarely been reported to occur in the context of known Mendelian disorders, including 8 cases of autosomal dominant capillary malformation-arteriovenous malformation

syndrome type 1 (CM-AVM1) caused by *RASA1* mutation (OMIM# 608354) (Revencu et al., 2013; Duran et al., 2018), 2 cases of autosomal dominant CM-AVM2 caused by *EPHB4* mutation (Amyere et al., 2017), and single cases of autosomal dominant Hereditary Hemorrhagic Telangiectasia (HHT) type 1 (HHT1) due to *ENG* mutation (OMIM# 187300) (Tsutsumi et al., 2011), and autosomal dominant HHT type 2 (HHT2) caused by *ACVRL1* mutation (OMIM# 600376; 1 case) (Chida et al., 2013).

Traditional genetic approaches have been limited in their ability to identify additional causative genes for VOGM and other types of brain AVMs because cases are rare and most often sporadic in families (Xu et al., 2010). This limitation motivates use of whole-exome sequencing (WES) in large numbers of affected subjects and their families, searching for genes mutated in affected subjects more often than expected by chance. This powerful unbiased approach has proven successful in the study of genetically heterogeneous neurodevelopmental disorders (Neale et al., 2012; Sanders et al., 2012; Krumm et al., 2015; Timberlake et al., 2016; Timberlake et al., 2017; Willsey et al., 2017), including those associated with brain malformations (Bilguvar et al., 2010; Barak et al., 2011; Mishra-Gorur et al., 2014; Nikolaev et al., 2018), congenital heart defects (Zaidi et al., 2013; Homsy et al., 2015; Jin et al., 2017), and most recently, congenital hydrocephalus (Furey et al., 2018). We hypothesized that the apparently sporadic occurrence of VOGM might reflect damaging *de novo* mutation events or incomplete penetrance of rare transmitted variants.

## RESULTS

### VOGM cohort characteristics and whole exome sequencing

We recruited 55 probands with radiographically confirmed VOGMs (see **Methods**). All probands underwent endovascular therapy. The cohort included 52 parent-offspring trios with a single affected offspring and 3 singleton cases. 62% of probands were diagnosed prenatally or in the first month after birth, and all but one was diagnosed by age 2 years. Common features at the time of diagnosis included high output cardiac failure (62%), macrocephaly (64%), hydrocephalus (60%) and prominent vasculature of the face or scalp (49%). Other congenital heart defects were found in 9%; these included partial anomalous pulmonary venous return, patent ductus arteriosus, and pulmonary valve stenosis. **Supplementary Table 1** summarizes cohort demographics, clinical features, and radiographic classification, including angio-architectural subgroups as classically defined by Lasjaunias, et. al (Lasjaunias et al., 2006). Of the 55 patients in our cohort, 38 presented with "choroidal" type lesions, defined as VOGMs with numerous feeder vessels and "pseudoniduses" that communicate with the MPV (Lasjaunias et al., 2006). The remaining 17 probands presented with "mural" VOGMs, characterized by a smaller number of larger caliber vessels that fistulize into the MPV (Lasjaunias et al., 2006). Representative images of VOGMs are shown in **Supplementary Figure 1**.

DNA was isolated and whole exome sequencing (WES) was performed as described (Timberlake et al., 2016). In parallel, WES from 1,789 control trios that comprise parents and unaffected siblings of autism probands was analyzed (Fischbach and Lord, 2010; Krumm et al., 2015). Cases and controls showed similar sequencing metrics, with ≥94.6% of targeted bases having at least 8 independent reads and ≥89.8% having at least 15 independent reads. Detailed exome

metrics are shown in **Supplementary Table 2**. Variants were called using the Genome Analysis Toolkit (GATK) HaplotypeCaller (McKenna et al., 2010; Van der Auwera et al., 2013), and allele frequencies were annotated in the Exome Aggregation Consortium (ExAC) and the Genome Aggregation Database (gnomAD) (Lek et al., 2016) (see **Methods**). TrioDeNovo was used to identify *de novo* mutations (Wei et al., 2015). MetaSVM was used to infer the impact of missense mutations (Dong et al., 2015). Missense variants were considered damaging if predicted deleterious by MetaSVM (D-mis). Inferred loss-of-function (LoF) mutations, including stop-gains, frameshift insertions/deletions, and canonical splice site mutations were also considered damaging. Mutations in genes of interest were validated by PCR amplification and Sanger sequencing (**Supplementary Figure 2**).

**Enrichment in damaging *de novo* mutations in VOGM probands**

The *de novo* mutation rate in probands was $1.73 \times 10^{-8}$ per base pair, with 1.35 *de novo* coding region mutations per proband (**Supplementary Table 3**). These results are consistent with expectation and prior experimental results (Homsy et al., 2015; Ware et al., 2015; Timberlake et al., 2017). The total burden of *de novo* mutations in controls was similar to a previous report (Timberlake et al., 2017). The distribution of types of *de novo* coding mutations observed in probands was compared to that expected from the probability of mutation of each base in the coding region and flanking splice sites (Samocha et al., 2014). While there was no significant enrichment of synonymous mutations or missense mutations inferred to be tolerated in probands, there was a significant enrichment of inferred damaging *de novo* missense (D-mis) mutations in probands ($P = 4.7 \times 10^{-3}$; enrichment = 2.19-fold; **Supplementary Table 3**). In contrast, no enrichment was observed in any class of coding region mutation in control subjects **(Supplementary Table 3**). From the observed excess of damaging *de novo* mutations in probands compared with expectation, we infer that these mutations contribute to ~17% of VOGM cases.

Only one gene, *KEL*, had more than one protein-altering *de novo* mutation (**Figure 1**). With analysis of 52 trios, we observed a near-significant enrichment of protein-altering *de novo* mutations in *KEL* (one-tailed *Poisson* test $P = 4.2 \times 10^{-6}$; Benjamini-Hochberg false discovery rate [BH-FDR] = *0.08*) using the denovolyzeR program (Ware et al., 2015). One of these was a mutation causing premature termination (p.Gln321*) and the other was a missense mutation p.Gly202Ser (CADD = 22.6). *KEL* encodes the Kell blood group transmembrane glycoprotein. Both the premature termination and p.Gly202Ser mutations in Kell alter its peptidase domain that generates vasoactive endothelin peptides (Lee et al., 1999), which provide vascular-derived axonal guidance cues (Makita et al., 2008).

We analyzed the burden of *de novo* mutations in genes that are highly intolerant to heterozygous LoF (pLI $\geq$ 0.9) (LoF-intolerant genes) (Lek et al., 2016), consistent with loss of a single dose of these genes causing markedly impaired reproductive fitness. We observed a significant enrichment of *de novo* D-mis mutations (6 mutations; *P = 0.02*; enrichment = 2.82-fold; **Table 1)**. Most interestingly, four of these *de novo* mutations (3 D-mis and the sole LoF mutation in this set) occurred in genes encoding chromatin modification, which have been implicated in the pathogenic mechanism of other congenital diseases including CHD, autism, and intellectual impairment (Zaidi et al., 2013; De Rubeis et al., 2014; Iossifov et al., 2014; McCarthy et al.,

2014; Homsy et al., 2015; Jin et al., 2017). There are 269 LoF-intolerant chromatin modifier genes (among the 547 genes in the chromatin modifier gene ontology term GO:0016569); mutations in this set are highly enriched ($P = 8.9 \times 10^{-4}$; enrichment = 9.63-fold; **Table 1**). No significant enrichment of *de novo* mutations in LoF-intolerant chromatin modifier gene sets was observed in controls (**Table 1)**. A case-control analysis using the two-tailed binomial exat test as described previously (Sanders et al., 2012; Willsey et al., 2017) further supported this result (**Supplementary Table 4**). In an orthogonal analysis of all genes using the permutation-based test (1M iterations), the probability of finding 4 or more damaging *de novo* mutations in intolerant chromatin modifiers among a total of only 21 damaging *de novo* mutations in all genes was also low (empirical $P = 5.1 \times 10^{-3}$; expected number = 0.72; see **Methods**). Together, damaging *de novo* mutations in chromatin modifiers were identified in ~8% of VOGM probands. The high enrichment of these mutations is consistent with each of these mutations contributing to disease pathogenesis.

The four mutated genes (**Supplementary Table 5**) include *KMT2D*, which encodes a histone H3K4 methyltransferase; *SMARCA2*, which encodes the ATP-dependent helicase of the BAF chromatin remodeling complex; *SIRT1*, encoding an NAD-dependent histone deacetylase; *KAT6A*, encoding a histone lysine acetyltransferase. All four genes are highly expressed in the developing human and murine brain (Machida et al., 2001; Ogawa et al., 2011; Pollen et al., 2015; Tham et al., 2015) and are known play essential roles in neuronal and vascular development (Potente et al., 2007; Griffin et al., 2008; Van Laarhoven et al., 2015). Missense mutations occur at positions that are completely conserved through orthologs in worm (*KAT6A* and *KMT2D*) and yeast (*SMARCA2*) and reside in key domains of protein function (**Figure 2, panel B; Supplementary Table 5**). Clinical characteristics of probands harboring these mutations are shown in **Figure 2** and **Supplementary Table 5**.

**Enrichment of rare damaging transmitted mutations in *EPHB4***

We next assessed the total burden of rare (MAF $\leq 2 \times 10^{-5}$) *de novo* and transmitted D-mis and LoF mutations in LoF-intolerant genes in all probands as described in **Methods**. The probability of the observed number of rare variants in each gene occurring by chance was compared to the expected burden, adjusting for the length of each gene as previously described (Besse et al., 2017). Analysis of damaging variants in LoF-intolerant genes revealed genome-wide significant enrichment (Bonferroni multiple testing threshold = $2.6 \times 10^{-6}$) of mutations in one gene: *EPHB4* (pLI = 0.99) had one LoF and three different D-mis mutations (one-tailed binomial $P = 7.47 \times 10^{-10}$; BH-FDR = $1.43 \times 10^{-5}$; enrichment = 341.13-fold; **Figure 3, panel A; Table 2**). Independent case-control analyses of gene burden for the damaging variants in all probands versus a cohort of 3,578 autism parental controls and ExAC controls showed a significant mutation burden in *EPHB4* vs. autism parental controls (one-tailed Fisher's $P = 1.68 \times 10^{-6}$ odds ratio = 89.97, 95% CI [19.29, Inf]) and vs. ExAC controls (one-tailed Fisher's $P = 1.98 \times 10^{-6}$ odds ratio = 49.76, 95% CI [16.39, Inf]) **Supplementary Table 6**). There are 3,230 genes in the intolerant gene set; the Bonferroni-corrected threshold for intolerant genes is therefore $1.54 \times 10^{-5}$. At this threshold, in addition to *EPHB4*, one additional gene, *CAD*, surpassed this threshold (one-tailed binomial $P = 4.17 \times 10^{-6}$, enrichment = 101.08-fold).

All *EPHB4* mutations were transmitted. Three of these have allele frequencies of zero in ExAC and gnomAD, while one has MAF of $1.13 \times 10^{-5}$ in gnomAD (Lek et al., 2016). All D-mis mutations in *EPHB4* alter highly conserved amino acid residues (**Figure 4, panel C; Table 2**). *EPHB4* encodes the Ephrin type-B receptor 4 (Eph-B4), a receptor tyrosine kinase preferentially expressed in vein endothelium and a critical regulator of arterio-venous identity (Gerety et al., 1999). The p.Lys650Asn and p.Phe867Leu mutations lie in the Eph-B4 receptor tyrosine kinase domain. p.Ala509Gly lies in one of two extracellular fibronectin III domains, which is believed to bind to extracellular matrix (**Figure 4, panel C**).

In addition to the 4 VOGM probands with rare damaging variants in *EPHB4*, 7 additional family members in these kindreds carried the same mutations; none had a diagnosis of VOGM. However, three *EPHB4* mutation carriers displayed uncommon cutaneous vascular lesions. For example, in kindred VGAM-115, the mutation-carrying father had an abdominal port wine stain, and the proband's mutation carrier sibling had atypical capillary malformations on the left cheek and posterior thigh (**Supplementary figure 3, panels A & B**). Similarly, the mutation-carrying mother in kindred KVGAM-33 had an atypical capillary malformation on her left arm. Additionally, the proband's mutation carrier sibling had an atrial septal defect (**Figure 4, panel B**). Vascular and cardiac abnormalities were absent among non-mutation carriers in these families. Notably, all VOGMs associated with *EPHB4* mutations were of the choroidal subtype (**Figure 4, panels A & B**). These findings provide evidence of incomplete penetrance and variable expressivity of *EPHB4* mutations. Interestingly, *EPHB4* mutations have been previously implicated in autosomal dominant capillary malformation-arteriovenous malformation (CM-AVM) syndrome type 2 (CM-AVM2); 2 of 54 patients were noted to have VOGM (Amyere et al., 2017).

Wild type and VOGM-mutant Eph-B4 (p.Ala509Gly; p.Lys650Asn; p.Phe867Leu) were expressed in mammalian cells (see **Methods**), and levels of Eph-B4 tyrosine (Tyr)-phosphorylation, which is correlated with receptor tyrosine kinase activity (Lisabeth et al., 2013; Ferguson et al., 2015), was compared between mutant and wild-type. Eph-B4 p.Lys650Asn showed sharply reduced phosphorylation, and no Tyr phosphorylation was detected on the Eph-B4 p.Phe867Leu mutant (**Figure 4, panel D**). These mutations lie in the Eph-B4 kinase domain. In contrast, the p.Ala509Gly, which lies in the extracellular fibronectin III domain, showed no significant difference in phosphorylation.

Eph-B4 Tyr phosphorylation creates docking sites for signaling molecules via a phosphotyrosine-SH2 domain interaction, thereby mediating Eph-B4 downstream signaling (Pawson and Scott, 1997; Wang et al., 2002). RASA1-encoded Ras-GAP binds to these docking sites and regulates downstream signaling pathways (Kawasaki et al., 2014; Roth Flach et al., 2016). We immunoprecipitated wild type or VOGM-mutant Eph-B4 from mammalian cells co-expressing Ras-GAP. Immunoblotting showed that the amount of Ras-GAP immunoprecipated with *EPHB4* p.Lys650Asn mutation was significantly reduced, and Ras-GAP failed to bind Eph-B4 p.Phe867Leu mutation (**Figure 4, panel E**). In contrast, wild type Eph-B4 and Eph-B4 p.Ala509Gly bound Ras-GAP with similar affinity. These results are consistent with the cytoplasmic mutations impairing receptor kinase activity; the extracellular mutation alters the fibronectin domains, which are likely to affect signaling by altering binding to ligands in extracellular matrix; this has not been directly tested.

**Recurrent rare damaging transmitted mutations in *CLDN14***

Expansion of the analysis to include the burden of rare (MAF $\leq$ $2 \times 10^{-5}$) D-mis and LoF mutations in *all* genes, identified one additional significant gene that showed genome-wide significant burden of damaging mutation, *CLDN14* (one-tailed binomial vs. expected $P = 6.44 \times 10^{-7}$; BH-FDR = 0.01; enrichment = 190.22-fold; **Figure 3, panel B; Table 2**). Comparing the burden of mutations in *CLDN14* by case-control analysis against both autism parental controls (odds ratio = Inf, 95% CI [38.42, Inf]; one-tailed Fisher's $P = 3.38 \times 10^{-6}$; **Supplementary Table 6a**) and ExAC controls, odds ratio = 67.82, 95% CI [17.66, Inf]; one-tailed Fisher's $P = 1.65 \times 10^{-6}$; **Supplementary Table 6b**) supported the significance of this finding.

The three rare damaging mutations in *CLDN14* were all heterozygous transmitted missense mutations (5.5% of probands; **Figure 5**). Strikingly, two of the three mutations are identical (p.Val143Met) and present in unrelated VOGM probands. p.Val143Met has a MAF of *0* in Non-Finnish Europeans in ExAC. Kinship analysis using Beagle v3.3.2 (Browning and Browning, 2011; Stuart et al., 2015) and trio analysis using 139 phased genotypes flanking the mutation revealed that p.Val143Met variant lies on a segment shared identically by-descent from a common ancestor by the two probands with a minimum shared segment of 0.34 Mb. Nonetheless, these probands did not share any other rare variants, indicating that they do not share a recent common ancestor. The other rare damaging variant, p.Ala113Pro, is absent in ExAC, and has a frequency of $8.28 \times 10^{-6}$ in gnomAD.

*CLDN14* encodes Claudin-14, a tight junction protein expressed in epithelia and endothelial cells in the brain and kidney (Kniesel and Wolburg, 2000; Wattenhofer et al., 2005). The extracellular loops of claudins are believed to make homo- or heterotypic interactions with adjacent cells to form the tight junction barrier (Van Itallie and Anderson, 2013). p.Val143Met alters a highly conserved residue in the second extracellular loop of Claudin 14, and is predicted to be deleterious by MetaSVM, CADD, Polyphen2, and SIFT. Claudin-14 p.Ala113Pro lies in the second intracellular loop of the protein (**Figure 5, panel C; Supplementary Figure 4**).

Interestingly, in one of the families harboring the recurrent Claudin-14 mutation p.Val143Met, cutaneous vascular lesions segregated with the mutation in family members. The mother who transmitted the mutation had a port wine stain on her left thigh; and both her children (proband and carrier sibling) had similar atypical capillary malformations in the nuchal area. Additionally, this VOGM proband had a midline atypical capillary malformation on her lower back (**Supplementary Figure 3, panels G-J**).

**Enrichment of mutations in genes in Ephrin signaling pathway**

To search for pathways enriched for rare damaging mutations in VOGM, LoF-intolerant genes harboring damaging *de novo* and/or rare (MAF $\leq$ $2 \times 10^{-5}$) damaging transmitted mutations, along with *KEL and CLDN14*, were input into IPA analysis (May 2018 version; total 128 input genes) (Kramer et al., 2014) (**Supplementary Table 7**). Axonal guidance signaling, which is essential for vascular patterning and critically regulated by Ephrin-Eph receptor forward and reverse signaling (Adams and Eichmann, 2010), was the most significantly enriched canonical

pathway ($P = 6.61 \times 10^{-8}$; BH-FDR = $1.33 \times 10^{-5}$); this pathway includes genes of the Ephrin receptor signaling pathway, which was also significant. Nine other pathways were significant after B-H correction in IPA pathway analysis. Because IPA does not adjust for gene length, we tested each of these pathways in case-control analysis (see **Methods**) using autism parents and ExAC controls that were matched for ethnicity; the only pathways that showed significant enrichment in cases were the axonal guidance and Ephrin receptor signaling pathways (**Supplementary Tables 8**). We further analyzed the burden of mutation in genes in these 2 pathways by binomial analysis comparing observed mutations to the expected after correcting for gene size (see **Methods**); there was observed significant enrichment in both axonal guidance and Ephrin receptor signaling pathways in cases, whereas analyses of synonymous variants in cases as well as rare damaging variants in autism parents and ExAC controls showed no significant enrichment (**Supplementary Tables 9**). Lastly, we found that most of the signal in the axonal guidance pathway was attributable to the Ephrin receptor signaling pathway, because after the latter genes were removed, the axonal guidance pathway was no longer significantly enriched (**Supplementary Table 10**). Detailed variant information of damaging mutations in genes that contributed to the significant result is described in **Supplementary Table 11**.

Mutations in the Ephrin receptor signaling pathway with rare damaging mutations included *EPHB4*, *RASA1*, *EPHA4*, *EPHA6*, *ITSN1*, *NGEF*, and *ITGB1* (**Supplementary Table 11**). The products of these genes are highly expressed in the embryonic human brain and vasculature (Guo et al., 2012), function in signaling networks that regulate neural and vascular development (see **Discussion**), and can be mapped into a single experimentally-supported STRING interactome (**Supplementary Figure 5**, see **Methods)**. Additionally, *EFNB2* (pLI = 0.94, mis-Z = 2.01), encoding the Eph-B4 ligand Ephrin-B2, harbored the rare D-mis mutation p.Arg277His in a neonate with a choroidal VOGM (**Supplementary Figure 6**). *EFNB2* is represented in the GO and IPA axonal guidance and Ephrin receptor signaling gene sets, but this specific *EFNB2* variant was not included as input in our pathway analysis because its MAF slightly exceeded our threshold of $2 \times 10^{-5}$ (MAF of $1.88 \times 10^{-5}$ and $2.89 \times 10^{-5}$ in ExAC and gnomAD, respectively). Interestingly, the proband's *EFNB2* mutation-carrying mother, while not having VOGM, exhibited multiple atypical nuchal capillary malformations. *Ephrin-b2* knockout mice phenocopy *eph-b4* knockout mice, exhibiting CNS and systemic AVMs, defects in cerebral angiogenesis, and embryonic lethality (Wang et al., 2010b).

**Rare damaging mutations in genes implicated in Mendelian forms of cutaneous vascular disease**

VOGM is a rare feature in patients with CM-AVM types 1 and 2 due to mutation in *RASA1* (Revencu et al., 2013; Duran et al., 2018) and *EPHB4* (Amyere et al., 2017), respectively. VOGM has also previously been associated with single cases of Hereditary Hemorrhagic Telangiectasia (HHT) type 1 due to *ENG* mutation (Tsutsumi et al., 2011) and HHT type 2 due to *ACVRL1* mutation (OMIM# 600376) (Chida et al., 2013).

In our VOGM cohort we found one damaging mutation in *RASA1* (p.Arg709*, **Supplementary Table 11**), encoding the Eph-B4 binding partner and effector Ras GTPase-activating protein 1 (Ras-GAP) (Kawasaki et al., 2014), in a patient with a mural VOGM. This patient and mutation

has been previously reported (Revencu et al., 2013) and was independently ascertained in the present study.

There was a single damaging p.Arg484Gln mutation in *ACVRL1* (**Supplementary Table 12**) in a patient with a choroidal VOGM. *ACVRL1* encodes ALK1, a receptor kinase in the TGF-β signaling pathway highly expressed in developing human vasculature (Zhang et al., 2017). This mutation alters a conserved residue in the ALK1 kinase domain. This mutation has been reported in HHT2 with isolated pulmonary hypertension (ClinVar accession RCV000198604.1; rs863223408) (Harrison et al., 2005), however, neither the *ACVRL1*-mutant VOGM proband nor mutation carrier family members exhibited any HHT-associated findings (e.g., epistaxis or telangiectasia) or any other vascular abnormalities. *ACVRL1* (ALK1) is a known regulator of Ephrin-B2-Eph-B4 signaling (Zhang, 2009; Kim et al., 2012; Roman and Hinck, 2017), and *acvrl1* deficient mice exhibit markedly enlarged cerebral vessels with arteriovenous shunting and altered expression of Eph-B4 (Walker et al., 2011).

Additionally, two rare damaging mutations (p.Gly39Ser and p.Asn373Ser) in unrelated probands with mural VOGMs were identified in the *ACVRL1* paralog *ACVR1* (pLI = 0.96, mis-Z = 2.59; **Supplementary Table 12; Supplementary Figure 7, panels A & B**), a gene not previously implicated in VOGM. *ACVR1* encodes the receptor serine-threonine receptor kinase ALK2 that functions in conjunction with TGF-beta type 2 receptors and co-receptors such as Endoglin (*ENG*) (Chen et al., 1998; Barbara et al., 1999; Wolfe and Myers, 2010). Of note, 1 previously reported VOGM case has been associated with mutations in *ENG* in HHT type 1 (OMIM# 187300). (Luo et al., 2010; Pacifici and Shore, 2016). Interestingly, our *ACVR1* p.Asn373Ser VOGM proband also had an atrial septal defect and partial anomalous pulmonary venous return (**Supplementary Figure 7, panel B**), and a mutation carrier in the family had cutaneous capillary malformations (**Supplementary Figure 7, panel B**). Neither the proband nor other family members had common HHT features (Abdalla et al., 2003). p.Gly39Ser and p.Asn373Ser alter conserved residues in the extracellular ligand binding domain and the cytoplasmic serine-threonine kinase domain, respectively (**Supplementary Figure 7, panel C**). The p.Asn373Ser substitution is predicted to result in critical structural alterations in the kinase domain (**Supplementary Figure 7, panels D & E**).

## DISCUSSION

The rarity and the sporadic nature of VOGM have hindered its genetic understanding. This study, the largest genomic analysis of VOGM to date, has provided new insights into disease pathogenesis, in part from its trio-based design. The results implicate damaging *de novo* mutations in ~17% of VOGM cases, show that mutations in chromatin modifiers account for about half this signal, and demonstrate mutations in *KEL* are likely to account for another fraction. Rare transmitted mutations in *EPHB4* and related members of the Ephrin signaling pathway contribute to another ~18% of probands. Thus, while rare mutations with large effect contribute to a significant fraction of VOGM cases, it is clear that there are likely many genes whose mutation contributes to disease pathogenesis, illustrating that larger studies will be needed to further validate and extend these findings. Information for all probands, including subphenotypes of VOGM and identified likely pathogenic mutations, are shown in

**Supplementary Table 13**. We also analyzed rare homozygous and compound heterozygous genotypes (MAF < 0.001), and found no genes with more than one such genotype.

The role of Eph-B4 signaling in arterio-venous specification and vascular development is well established in model systems (Zhang and Hughes, 2006; Mosch et al., 2010). We found damaging mutations in *EPHB4* in 8% of VOGM in this cohort, implicating a failure in normal vascular differentiation in disease pathogenesis. Two of 52 prior patients with CM-AVM and *EPHB4* mutation have been reported to have VOGMs (Amyere et al., 2017), and during preparation of this manuscript, another paper reported mutations in *EPHB4* in VOGM (Vivanti et al., 2018). Additionally, we found significant enrichment of very rare heterozygous damaging mutations in 6 other genes involved in Ephrin signaling, implicating a broader role for mutations in these pathways. *KEL*, encoding the Kell blood group transmembrane glycoprotein, had two damaging *de novo* mutations. The Kell peptidase domain generates endothelins, vasoactive peptides via cleavage of the endothelin-3 pro-protein (Lee et al., 1999). Endothelins serve as signals in axonal guidance (Makita et al., 2008), which also plays a key role in vascular patterning (Adams and Eichmann, 2010). We also found rare mutations in genes previously implicated in HHT, another disease featuring potentially dangerous AVMs in VOGM probands. These findings implicate mutations in common pathways regulating vascular development in disease pathogenesis.

In addition to these genes in which known biology has pointed to a role in vascular development, we observed recurrent mutations in *CLDN14*, with a genome-wide significant burden of rare damaging mutation, and in *ACVR1*, which is in the TGF-beta signaling pathway previously implicated in HHT. Claudin-14 is a four-pass integral membrane protein and component of tight junctions in brain epithelia and endothelial cells (Kniesel and Wolburg, 2000; Wattenhofer et al., 2005). The recurrent mutation in *CLDN14* lies in the large 2$^{nd}$ extracellular loop that likely plays a role in intercellular contact and tight junction formation. The regulation of endothelial tight junction formation by Claudins can impact cell permeability, integrity, and proliferation (Morita et al., 1999; Gonzalez-Mariscal et al., 2007). Functional interactions have been reported between Claudins and Ephrin-B2-EphB4 and Ephrin-B1-Eph-A2 bi-directional signaling (Tanaka et al., 2005). Further work will be required to establish the role of *CLDN14* in vascular development.

It is striking that the transmitted disease-associated mutations identified herein show dramatic incomplete penetrance for VOGM, with none of the identified mutation carriers having VOGM. Moreover, many of the mutated VOGM genes have been implicated in other Mendelian diseases that play critical roles in brain and cardiovascular development, sometimes producing quite different phenotypes. These observations highlight the pleiotropy with variable expressivity resulting from these mutations. This feature has been described previously for a number of diseases. For example, haploinsufficiency for the same chromatin modifying genes results in congenital heart disease (Zaidi et al., 2013) or autism (Iossifov et al., 2014), or both.

Variable expressivity of VOGM and associated features could arise from specific genetic (Timberlake et al., 2016) or environmental modifiers (Garcia et al., 2015) working in concert with the rare mutations identified. Alternatively, the results could be explained by a "two-hit" mechanism in which phenotypic expression relies upon an inherited mutation and a second, post-zygotic mutation in the wild-type allele (Brouillard et al., 2002; Pagenstecher et al., 2009). The

observation that many VOGM patients, as well as their mutation-bearing family members, have sporadic cutaneous capillary malformations (**Supplementary Table 13**) is consistent with a two-hit mechanism. This mechanism has been shown for other hereditary multifocal vascular malformations such as *RASA1*-mutated CM-AVM1 (Revencu et al., 2013), glomuvenous malformations (OMIM# 138000) cutaneomucosal venous malformation (OMIM# 600195), and cerebral cavernous malformations (OMIM# 116860) (Pagenstecher et al., 2009). In this model, the phenotypic expression would depend on the cell types in which somatic mutations occur. Such a mechanism could obviously explain the low penetrance of VOGM arising from transmitted mutations. Further work will be required to explore this possibility.

Another possibility is that different mutations in the same gene produce distinctive phenotypes. In several instances, this seems unlikely. For example, damaging heterozygous *de novo* missense mutations in *SMARCA2* that strongly cluster in the ATPase domain cause Nicolaides-Baraitser syndrome [OMIM# 601358], featuring severe mental retardation, early-onset seizures, short stature, dysmorphic facial features, and sparse hair. VOGM has not, to our knowledge, been reported in this disease. Our *SMARCA2* mutated VOGM proband harboured a *de novo* substitution (p. Arg855Leu) at the identical position in which another *de novo* mutation (p. Arg855Gly) has been reported in a Nicolaides-Baraitser patient. Other *de novo* missense mutations in Nicolaides-Baraitser cluster in adjacent codons 851-854 (Van Houdt et al., 2012). The *SMARCA2*-mutated VOGM proband had a seizure history but otherwise is without overlapping phenotypes with Nicolaides-Baraitser syndrome. The observations that loss of *SMARCA2's* ortholog in mouse (*BRM*) impairs neurovascular development (Son and Crabtree, 2014; Wiley et al., 2015), and that its paralog *SMARCA4* (aka *BRG1*) is also essential for venous specification and neurovascular development, further supports a role for mutation in *SMARCA2* in neurovascular development (Davis et al., 2013).

Similarly, *de novo* heterozygous mutations in *KMT2D* cause autosomal dominant Kabuki syndrome (OMIM# 147920) featuring distinctive facial dysmorphology with frequent skeletal abnormalities, mental retardation, dermatoglyphic abnormalities, growth deficiency, and, in a third of patients, severe cardiovascular and vessel abnormalities (Hannibal et al., 2011; Dentici et al., 2015; Van Laarhoven et al., 2015). While LoF *KMT2D* mutations predominate in Kabuki syndrome, damaging missense mutations also occur, and do so most frequently in its short F/Y-rich domain (FYRN) (Yan et al., 2000; Holbert et al., 2007). The FYRN (F/Y-rich N terminus) and FYRC (F/Y-rich C terminus) domains are found in many histone-modifying proteins proteins (Garcia-Alai et al., 2010; Zhang et al., 2015) and the physical interaction of these domains is required for enzymatic function (Hsieh et al., 2003). Like many Kabuki-related missense mutations, our VOGM-associated *de novo* p.Cys5230Tyr mutation maps to the FYRN (F/Y-rich N-terminus) domain. Moreover, similar to a previously reported case of Kabuki syndrome (Moon et al., 2018), this VOGM proband has low set ears, long and everted palpebral fissures, and congenital hypothyroidism. This observation provides strong supporting evidence that our patient's *KMT2D* mutation is indeed Kabuki-related. Interestingly, another case of Kabuki syndrome was reported to have severe dilation (ectasis) of the Vein of Galen without clear fistulas to the MPV (Sanchez-Carpintero et al., 2012).

Alternatively, phenotypic heterogeneity, in which different mutations in the same gene can give rise to give rise to different phenotypes, may contribute to other VOGM cases. For example, LoF

c-terminal frameshift or premature termination mutations in *KAT6A* cause autosomal dominant mental retardation (MR) type 32 (OMIM# 601408), also featuring facial dysmorphism, seizures, and multiple vascular abnormalities including ventricular and atrial septal defects and patent ductus arteriosus (Arboleda et al., 2015). In contrast, our VOGM *KAT6A* mutant proband harbored a proximal missense mutation (p.Thr478Ile) between its PHD finger and histone acetyl transferase (HAT) domains (Yan et al., 2000; Holbert et al., 2007). This proband had a history of seizures but no other overlapping phenotypes with MR type 32; however, the patient did exhibit multifocal atypical cutaneous vascular lesions, including capillary malformations. Consistent with a role in VOGM pathogenesis, KAT6A regulates the expression of the T-box transcription factors TBX1 and TBX5, essential regulators of the pharyngeal arch vessels that give rise to the great arteries of the neck, head, and brain (Merscher et al., 2001; Vitelli et al., 2002; Zhang et al., 2005; Vanyai et al., 2015).

Two other potential examples of phenotypic heterogeneity in our cohort include probands with inherited *ACVR1* and *CLDN14* mutations. Heterozygous germline mutations in *ACVR1* cause autosomal dominant fibrodysplasia ossificans progressiva (OMIM# 135100) (Shore et al., 2006). However, most fibrodysplasia ossificans progressiva cases are due to a recurrent p.Arg206His gain-of-function mutation that leads to constitutive activation of TGF-beta signaling (Chaikuad et al., 2012; Mura et al., 2012). In contrast, *ACVR1* mutations in VOGM are remote damaging missense mutations (p.Gly39Ser and p.Asn373Ser) that localize to the extracellular Activin receptor domain and the cytoplasmic serine-threonine kinase domain, respectively. These VOGM probands do not exhibit features of fibrodysplasia ossificans progressiva. Similarly, recessive LoF genotypes in *CLDN14* cause sensorineural deafness type 29 (OMIM # 614035); in contrast, VOGM-associated *CLDN14* mutations are heterozygous and D-mis, and include a recurrent missense mutation, raising the possibility of gain of function or neomorphic effects.

Cutaneous vascular lesions are a common hallmark of several developmental vascular disorders including HHT, Parkes Weber syndrome, and CM-AVM (Brouillard and Vikkula, 2007; Revencu et al., 2013; Amyere et al., 2017). The observation that 68% of VOGM families with full clinical data (31/45) had capillary malformations or other uncommon cutaneous vascular lesions, and that identified mutations in probands were found in all family members with these cutaneous lesions, provides evidence linking both VOGM and these cutaneous lesions to the same mutations. In this study we found capillary malformations and related cutaneous lesions in VOGM probands and/or relatives harboring *EFNB2*, *EPHB4*, *EPHA4*, *ACVR1*, and *CLDN14* mutations. This is consistent with the notion that mutations in these VOGM-related genes may predispose to other organ-based AVMs as well (Brouillard and Vikkula, 2007).

These findings have clinical implications, indicating that the offspring of mutation carriers are at increased risk of VOGMs, as well as capillary malformations and potentially other AVMs. However, not all mutation carriers develop capillary malformations, making the absence of CMs an unreliable clinical marker for risk of transmission in families with affected members. These observations make family history and mutation-based screening important for risk assessment. Moreover, the narrow developmental window of gestational weeks 6-11 during which fistulas to the vein of Markowski form (Raybaud et al., 1989) poses a challenge to improvement of early therapeutic strategies for VOGM, further compounded by our incomplete understanding of VOGM pathogenesis. Thus, attempted genetic or biomarker diagnosis with intention to treat

must be performed earlier than the safe gestational age threshold for amniocentesis (Shulman et al., 1994). These difficulties highlight the need for continued genetic research on VOGM and CM-AVM, with focus on the mechanistic implications of recently discovered VOGM-associated mutations. Interestingly, CM-AVM2-associated Eph-B4 kinase domain mutations result in loss of inhibition, and consequent activation, of downstream RAS/MAPK/ERK1/2 and PI3K/AKT/mTORC1 signaling cascades (Kim et al., 2002; Salaita and Groves, 2010; Xiao et al., 2012). PI3K/AKT/mTORC1 up-regulation has been noted in capillary malformations of *RASA1*-mutant CM-AVM1 patients (Kawasaki et al., 2014). These data suggest that therapy targeting Eph-B4-Ras-GAP-mTOR signaling pathway may represent a therapeutic target for VOGM and possibly other vascular lesions in the CM-AVM spectrum.

## ACKNOWLEDGEMENTS

We are grateful to the patients and families who participated in this research. Support for this work was provided by the Yale-NIH Center for Mendelian Genomics (5U54HG006504) and an NIH NRCDP award to K.T.K. S.C.J. was supported by the James Hudson Brown-Alexander Brown Coxe Postdoctoral Fellowship at the Yale University School of Medicine and the American Heart Association Postdoctoral Fellowship. J.G. was supported by the Howard Hughes Institute Medical Research Fellowship. Last, we acknowledge Collin Whitmore and his family members for their inspiration, courage, and generous support of this research.

## MAIN FIGURES

**Figure 1. Damaging *de novo KEL* mutations in choroidal Vein of Galen malformation**

**(A)** Representative magnetic resonance angiography and digital subtraction angiography images demonstrating mural and choroidal Vein of Galen malformations in KVGAM45-1 and KVGAM10-1, respectively.

**(B)** Pedigrees depicting kindred structures. Note that probands carrying *de novo* mutations in *KEL* are the only family member in families KVGAM45 and KVGAM10 that demonstrates Vein of Galen malformations; none of the family members in these two families have any disease phenotypes or cutaneous manifestations.

(C) Mutations identified by exome sequencing were confirmed by direct PCR amplification with custom primers followed by Sanger sequencing.

**(D)** Linear representation of the KEL molecule, with functional domains represented as dark rectangles. Amino acid modifications are mapped (in red) on the protein structure. Conservation of the wild-type amino acid substituted by the missense mutations is depicted below. TM = Transmembrane domain.

**Figure 2. Damaging *de novo* mutations in chromatin modifiers in Vein of Galen malformation**

**(A)** Vascular imaging of probands. Magnetic resonance angiographies and a digital subtraction angiography reconstruction demonstrating choroidal Vein of Galen malformations are shown.

**(B)** Pedigree structures of VOGM kindreds. For each kindred, the gene and mutation, the angiographic image, and pedigree structure are shown. Subjects with atypical capillary malformations are denoted by blue symbols. Red 'D' denotes damaging mutation, '+' denotes wild type sequence.

**(C)** Linear representation of functional domains of SIRT1, KMT2D, SMARCA2, and KAT6A with location of VOGM mutations. Functional domains are represented by dark rectangles. Amino acid changes are located on the protein structure in red lettering. For missense mutations, conservation of the wild-type amino acid across phylogeny is shown. The mutated amino acid is highlighted in yellow. SIR2 = Sirtuin catalytic domain, SIR2 Domain; PHD = Zinc Finger PHD type; MOZ_SAS = Histone acetyltransferase domain, MYST-type; zf-H = PHD-like zinc binding domain; FYRN = F/Y-rich domain - F/Y-rich N-terminus motif; FYRC = F/Y-rich domain - F/Y-rich C-terminus motif; SET = Su(var)3-9, Enhancer-of-zeste and Trithorax; QLQ = Glutamine-Leucine-Glutamine domain; HAS = Helicase-SANT associated domain; BRK = BRK domain; SNF2_N = SNF2-related, N-terminal domain; Hel_C = Helicase C-terminal domain; SnAC = Snf2-ATP coupling, chromatin remodeling complex; BD = Bromodomain.

**(D)** Schematic representation of modifications on histone marks exerted by SIRT1, KMT2D, SMARCA2, and KAT6A.

**Figure 3. Exome-wide significant enrichment of rare damaging transmitted mutations in *EPHB4* and *CLDN14***

**(A)** Quantile-quantile plots of observed vs. expected binomial test p-values for rare damaging (D-mis+LoF) variants with MAF < $2 \times 10^{-5}$ in the Genome Aggregation database (gnomAD) in each gene that is intolerant to LoF (pLI > 0.9).

**(B)** Quantile-quantile plots of observed vs. expected binomial test p-values for rare damaging (D-mis+LoF) variants with MAF < $2 \times 10^{-5}$ in gnomAD for all genes. MAF = Minor allele frequency; D-mis = Missense mutations predicted to be deleterious per MetaSVM; LoF = Canonical loss-of-function mutations (stop-gains, frameshift insertions or deletions, canonical splice site mutations).

**Figure 4. Damaging *EPHB4* mutations in choroidal Vein of Galen malformation**

**(A)** Vascular imaging of probands from digital subtraction angiographies and magnetic resonance angiography demonstrating Vein of Galen malformations.

**(B)** Pedigree structures of the kindreds. For each kindred, the gene and mutation, the angiographic image, and pedigree structure are shown. A carrier harboring an atrial septal defect is represented by a yellow symbol in family KVGAM 33.

**(C)** Linear representation of Eph-B4 functional domains (dark rectangles) with location of VOGM mutations (in red). Conservation across phylogeny of the wild-type amino acid (in yellow) at each mutated position is shown. LB = ligand binding domain; CRD = Cysteine-rich domain; FNIII = Fibronectin III domain; TM = Transmembrane domain; TK = Tyrosine kinase domain; SAM = Sterile alpha motif.

**(D)** Representative immunoblots showing the effect of Ala509Gly, Lys650Asn, and Phe867Leu mutations on resting state Eph-B4 tyrosine phosphorylation in HEK 293T cells. Receptor phosphorylation was analyzed by immunoprecipitation (IP) with HA-Tag antibody and Immunoblotting (IB) using anti–p-tyrosine (p-Tyr) and HA-Tag antibodies. Blot demonstrates drastically reduced phosphotyrosine content in Lys650Asn mutant, and no detectable phosphotyrosine in the Phe867Leu mutant.

**(E)** Representative immunoblot demonstrating binding of Eph-B4 constructs by Ras-GAP. Ras-GAP protein was co-immunoprecipitated with Eph-B4 mutants. Ras-GAP binding was markedly reduced in Lys650Asn mutants and abrogated in Phe867Leu mutants.

**Figure 5. Damaging *CLDN14* mutations in Vein of Galen malformation**

**(A)** Representative images from digital subtraction angiographies and magnetic resonance angiography demonstrating Vein of Galen malformations of the probands.

**(B)** Pedigree structures indicating genotypes and phenotypes as described in Figure 1.

**(C)** Linear representation of Claudin-14 functional domains (dark rectangles) with location of VOGM mutations (in red). Conservation of the wild-type amino acid is shown. TM = Transmembrane domain; C = cytoplasmic loop; E = extracellular loop.

## TABLES

**Table 1. Enrichment of damaging *de novo* mutations in chromatin modifiers intolerant to LoF mutation in VOGM**

**Table 2. Transmitted mutations in *EPHB4* and *CLDN14* in VOGM**

## SUPPLEMENTARY FIGURES

**Supplementary figure 1. VOGM type and representative imaging for probands**
Representative images of 3-Tesla time-of-flight magnetic resonance angiography or digital subtraction angiography for all patients with available imaging, with patient codes and VOGM subtype. In cases for which imaging was not available, VOGM subtype was ascertained from detailed transcripts of angiographies during or before endovascular treatment of VOGM.

**Supplementary figure 2. Chromatogram alignments and segregation of candidate variants**
Mutations identified by exome sequencing were confirmed by direct PCR amplification with custom primers followed by Sanger sequencing.

**Supplementary figure 3. Cutaneous manifestations in VOGM probands and family members**
Atypical capillary malformation on the posterolateral aspect of the distal third of the thigh (A) and left cheek (B) of VGAM115-2. Panel C demonstrates an atypical capillary malformation on left flank of VGAM115-4 – both individuals depicted in panels A-C carry *EPHB4* c.1295_1296del; p.Glu432fs. Atypical capillary malformations on the angle of the jaw and superior right thorax of KVGAM42-1 (D) and KVGAM42-2 (E). Both individuals depicted in panels D-E carry *RASA1* c.2125C>T; p.Arg708*. (F). Atypical capillary malformation on the posterolateral aspect of the forearm in VGAM114-5, carrier of *EPHA4* c.2347G>A; p.Gly783Ser. (G). Nuccal atypical capillary malformation in a proband harboring *CLDN14* p.(Val143Met) (H). Lumbar atypical capillary malformation in the same proband depicted in panel G. (I). Nuccal atypical capillary malformation in the sister of the proband depicted in panels G and H, also harboring *CLDN14* p.Val143Met. (J). Atypical capillary malformation in the thigh of the mother of individuals depicted in panels G-I, carrying and transmitting *CLDN14* p.Val143Met.

**Supplementary figure 4. *In silico* modeling of *CLDN14* mutations**

- **(A)** *In silico* modeling of p.Ala113Pro. Ala113 is the first residue of Helix-3 and lies between P112 and K114 at the interface of Claudin-14 and the cytoplasmic side of the lipid bilayer. Just preceding helix-3 is the loop that connects helix-2 and -3, on which posttranslational modification sites are present including palmitoylation. The presence of Proline in an alpha helix can result in helix distortion since it lacks the backbone NH to establish hydrogen bond necessary to complete the H-bond chain. Additionally the steric or rotamer effects prevent Proline from adopting the helical geometry. Presence of two adjacent closed-cyclic side chains of Proline at 112 and 113 are presumed to structurally influence the conformation of the helix-2 and -3 connector loop and the destabilize the start of helix-3. The $\Delta\Delta G$ of A113P is calculated to be 1.08 Kcal/mol.
- **(B)** *In silico* modeling of p.Val143Met. Val143 is located towards the end of helix-3 at the outer interface of the lipid bilayer. The side chain of valine is tightly packed against beta sheets being surrounded by the hydrophobic side chains of L164, I160 and V144. The V143M mutation ($\Delta\Delta G$ = 0.6 Kcal/mol) results in the longer side chain of methionine having steric clashes with the beta sheets.

**Supplementary figure 5. Interactome of mutated Ephrin receptor signaling genes in VOGM**

All 7 genes (highlighted) contributing to the significant result of the Ephrin receptor signaling pathway in pathway analysis using IPA was used as input in String. These 7 genes are mapped into a single experimentally supported STRING interactome.

**Supplementary figure 6. Mutations affecting other members of the Ephrin family in Vein of Galen malformation patients displaying CM-AVM-like cutaneous manifestations**

- **(A)** Representative magnetic resonance angiography and digital subtraction angiography images demonstrating choroidal Vein of Galen malformations in probands.
- **(B)** Pedigrees depicting kindred structure and phenotypic variance. Note the presence of atypical capillary malformations represented by blue symbols in both parents transmitting *EFNB2* and *EPHB4* mutations in families VGAM101 and VGAM114.
- **(C)** Linear representation of affected Ephrin molecules. Functional domains are represented by dark rectangles. Amino acid modifications are mapped on the protein structure in red lettering. Conservation of the wild-type amino acid substituted by the four deleterious missense mutations is depicted on the right of each molecule diagram. LB = ligand binding domain; CRD = Cysteine-rich domain; FNIII = Fibronectin III domain; TM = Transmembrane domain; TK = Tyrosine kinase domain; SAM = Sterile alpha motif.
- **(D)** *In silico* modeling of the repercussion of Eph-A4 p.Gly783Ser. Gly783 lies at the base of a flexible loop in Eph-A4.
- **(E)** The Gly-to-Ser substitution at this position (ΔΔG 0.6Kcal/mol) is predicted to affect loop stability.

**Supplementary figure 7. Damaging *ACVR1* mutations in mural Vein of Galen malformation**

- **(A)** Representative images from digital subtraction angiographies demonstrating mural Vein of Galen Malformations in probands.
- **(B)** Pedigrees depicting kindred structure and phenotypic variance. Blue symbols represent atypical capillary malformations in individuals harboring the c.1118A>G; p.Asn373Ser mutation. Yellow coloring of half the symbol denotes concomitant presence of an atrial septal defect and (asterisk) partial anomalous pulmonary venous return in the proband of family KVGAM44.
- **(C)** Linear representation of the ACVR1 molecule, with functional domains represented as dark rectangles. Amino acid modifications are mapped (in red) on the protein structure. Conservation of the wild-type amino acid substituted by the two deleterious missense mutations is depicted below each. AR = Activin receptor; TM = Transmembrane domain; KD = Kinase domain.
- **(D)** *In silico* modeling of p.Asn373Ser. Asn373 in a 3D model of ACVR1.
- **(E)** Predicted energetically costly (ΔΔG 2.9Kcal/mol) loss of hydrogen bond formation between the Asn373 side chain and backbone atoms of Gly356, Lys243, Arg375 and the side chain of Ser244 caused by Asn373 substitution to serine.

**Supplementary figure 8. The genomic landscape of VOGM**

Schematic representation of signaling between two primitive endothelial cells showing contact-mediated interaction between the membrane-bound ligand Ephrin-B2 in the upper cell and its

receptor, Eph-B4, in the lower cell. The figure depicts crosstalk between Eph-B4, its effector and binding partner RASA1, and Activin type 1 receptors mediated by non-canonical TGF-β signaling. Mutations discovered in our cohort affecting *EPHB4*, *EFNB2*, and *ACVR1* are shown in red, and previously reported mutations in *RASA1*, *EPHB4*, and *ENG* in VOGM patients are indicated in black. *RASA1* p.Arg709* (highlighted in yellow) is the sole *RASA1* mutation detected in our cohort.

## SUPPLEMENTARY TABLES

**Supplementary table 1. VOGM patient clinical and demographic characteristics**

**Supplementary table 2. Summary sequencing statistics for the VOGM cases and control cohorts**

**Supplementary table 3. Enrichment of *de novo* variant types in 52 VOGM and 1,789 control trios**

**Supplementary table 4. *De novo* burden analysis comparing cases to controls**

**Supplementary table 5. Damaging *de novo* mutation in loss-of-function intolerant chromatin modifier genes in 55 cases**

**Supplementary table 6. Case-control analysis of gene burden of the damaging variants in all probands versus 3,578 Autism parental controls (panel a) and ExAC controls (panel b)**

**Supplementary table 7. Significant pathways in Ingenuity ® Pathway Analysis including intolerant genes harboring damaging *de novo* mutations, damaging rare transmitted mutations, and significantly mutated genes (*KEL*, *EPHB4*, and *CLDN14*) (n = 128)**

**Supplementary table 8. Mutation burden in cases compared to ethnicity matched (European) controls in pathways significant from IPA**

**Supplementary table 9. Gene burden in cases compared to expectation for axonal guidance and Ephrin receptor signaling pathways**

**Supplementary table 10. Gene burden in cases compared to expectation (panel a) and ethnicity matched (European) controls (panel b) for axonal guidance pathway after removing genes in the Ephrin signaling pathway**

**Supplementary table 11. Mutations in genes in the Ephrin receptor signaling pathway**

**Supplementary table 12. Transmitted VOGM-associated mutations in genes causing CM-AVM1, HHT, and paralog**

**Supplementary table 13. Distribution of highly pathogenic mutations in VOGM patients with different subphenotypes**

## METHODS

All procedures in this study comply with Yale University's Human Investigation Committee (HIC) and are Human Research Protection Program. Written informed consent was obtained from all adult participants. Parent or legal guardian authorization was obtained in writing for sample collection of all minors in this study.

### Whole exome sequencing and variant calling

Exon capture was performed on genomic DNA samples derived from saliva or blood using Roche SeqCap EZ MedExome Target Enrichment kit or IDT xGen target capture kit followed by 99 base paired-end sequencing on the Illumina HiSeq 2500 platform. Sequence reads were aligned to the human reference genome GRCh37/hg19 using BWA-MEM (Li, 2014) and further processed to call variants following the GATK Best Practices workflow (McKenna et al., 2010). All variants covered by independent aligned sequencing reads with a depth of 8x or greater were visualized *in silico* to eliminate false positives. *De novo* mutations were identified using TrioDeNovo (Venugopal, 2014). *De novo* and transmitted variants were annotated with ANNOVAR (Wang et al., 2010a). MetaSVM (Dong et al., 2015) was used to predict the deleteriousness of non-synonymous variants (herein referred to as D-mis). Only loss-of-function mutations (nonsense, canonical splice-site, frameshift indels) and D-mis mutations (missense mutations predicted deleterious by MetaSVM) with MAF $\leq 2 \times 10^{-5}$ in gnomAD (calculated based on combined dataset of WES and WGS data from gnomAD database) were considered potentially damaging for unbiased binomial analysis. Missense variants were further annotated for deleteriousness utilizing the Missense badness, PolyPhen-2, and Constraint (MPC) score (Samocha et al.). Candidate mutations were confirmed by PCR amplification followed by Sanger sequencing (primer sequences available on request).

### Kinship analysis

Pairwise proband relatedness and pedigree information of trios were confirmed using KING (Manichaikul et al., 2010) by estimating kinship coefficient and calculating identity-by-descent (IBD). The shared pairwise IBD segments in 45 European probands were detected using Beagle v3.3.2 (Browning and Browning, 2011).

### *De novo* enrichment analysis and variant stratification

We derived the probability of observing a *de novo* mutation in each gene based on a context specific model as described previously (Samocha et al., 2014). Due to the differences in exome capture reagent, sequencing platforms, and sequencing coverage between cases and controls, the expected number of *de novo* mutations was estimated by adjusting for sequencing depth in 52 case trios and 1,789 autism control trios separately. Rather than using the variant calls in controls published in the original Simons Simplex Collection (SSC) study (Krumm et al., 2015), we downloaded the bam files from the SSC, reanalyzed the data, and filtered the control vcf file using the same filter criteria as what was used in our case cohort. A one-tail *Poisson* test was used to compare observed number of *de novo* mutations across each variant class to expected number under the null hypothesis. The R package 'denovolyzeR' (Ware et al., 2015) was used to

perform all *de novo* analyses. All genes represented in this dataset were annotated with artery-specific and brain-specific expression values in a form of reads per kilobase transcript per million reads (RPKM) from the GTEx database (https://gtexportal.org/home/). Genes harboring *de novo* mutations are also annotated with human brain specific expression data obtained during the first four weeks of development (Gerrard et al., 2016) in the form of quantile rank based on transcript per kilobase million (TPM), indicating the relative rank of expression level within human genome. The final dataset was analyzed for recurrently affected genes, and all variants in genes affected by a single *de novo* mutation were stratified. LoF variants were ranked based on pLI (from highest to lowest).

**Estimation of the number of damaging *de novo* mutations in LoF-intolerant chromatin modifiers.**

One million permutations were performed to derive the empirical distribution of the number of damaging *de novo* mutation in the LoF-intolerant chromatin modifier genes. For each permutation, the number of observed damaging *de novo* mutations in all genes (n=21) was randomly distributed across the genome, weighted according to the gene mutability. The empirical P value is calculated as the proportion of times that the number of damaging *de novo* mutations from the permutation is greater than or equal to the observed number (n=4) of damaging *de novo* mutations in the LoF-intolerant chromatin modifier genes. The average number of expected number of damaging *de novo* mutations in the LoF-intolerant chromatin modifier genes is also calculated.

**Binomial analysis**

Independent binomial tests were used to compare expected and observed number of rare variants in each gene. The expected number of rare damaging variants is determined by taking the fractional length of a gene (in base pairs) relative to the entire exome and multiplying this by the total number of rare damaging variants. This represents the expected occurrence of sporadic mutations in each gene without considering the influences of selection pressure or precedents of ethnic background. Genome-wide significance threshold is taken as p-value less than $2.6 \times 10^{-6}$ after applying Bonferroni correction considering the examination of ~19,000 genes in the human genome. Inputs for this test were those with inferred pathogenicity, as defined by 1. Missense mutations called deleterious per MetaSVM, or 2. Inferred loss-of-function mutations, including stop gains, frameshift insertions/deletions, or canonical splice site mutations. Binomial analysis for mutational enrichment did not include non-frameshift insertions or deletions.

**Pathway analysis**

Input for this analysis were LoF-intolerant genes (pLI $\geq$ 0.9) harboring damaging *de novo* mutations and/or rare (MAF $\leq 2 \times 10^{-5}$) damaging transmitted mutations (n = 128 genes), as well as genes with significant burden of *de novo* (*KEL*) or transmitted mutation (*EPHB4* and *CLDN14*) into Ingenuity Pathway Analysis (IPA, Apr 2018). Core analysis using Ingenuity Knowledge Base (Gene Only) as the reference set was performed. P-value was calculated using a one-tailed Fisher's exact test reflecting the likelihood that the overlap between the input and a given gene set is due to random chance. In individual based case-control analysis, ethnicity-

matched case and control samples were filtered using the same criteria. Individuals carrying variants of interest in case and control groups were tallied separately, and the p-value was obtained from a one-tailed Fisher's exact test. In binomial pathway analysis, the observed number of rare damaging variants in LoF-intolerant genes that belong to statistically significant canonical pathways of interest were compared to the expected number of mutations in each set using a one-tailed binomial test. Gene sets of canonical pathways, axonal guidance signaling (n = 452) and Ephrin receptor signaling (n = 174), were obtained from IPA. The expected number of mutations in a given gene set is calculated as the formula below:

$$\text{Expected number of mutations} = \text{N} \times \frac{\sum_{Gene\ Set} \text{Gene Length}}{\sum_{Intolerant\ Genes} \text{Gene Length}}$$

Where N denotes total number of rare damaging *de novo* and transmitted mutations in intolerant genes as well as genes with significant burden of *de novo* and transmitted mutations.

**Interactome construction**

We input all genes contributing to the significantly enriched pathway to String (version 10.5) (Szklarczyk et al., 2015). For organism, *Homo sapiens* was selected. For each displayed interaction, active interaction sources were restricted to experiments, and the maximum number of interactors was limited to 50.

***In silico* modeling of mutational effects on protein structure**

The sequence for all available modeled human proteins was downloaded from Uniprot (Apweiler et al., 2004). The stereochemical parameters of VOGM-associated mutations were analyzed using PROCHECK (Laskowski et al., 1993) and PROSA (Wiederstein and Sippl, 2007), and the final models were chosen based on the lowest energy function score (Dope) within the modeling program. The mutations were constructed and the free energy of change calculated (ΔΔG) in silico using the ICM mutagenesis (Abagyan et al., 1994).

**Cell Culture**

HEK 293T cells were passaged at 80-90% confluence on high glucose DMEM (Dulbecco's modified Eagle's medium, Gibco Life Technologies, Waltham MA, USA) supplemented with 10% fetal bovine serum (FBS, Gibco Life Technologies Waltham MA, USA), L-glutamine, and penicillin/streptomycin.

**Mutagenesis of Eph-B4 and Plasmid Transfection**

A wild-type mouse *EPHB4* cDNA was sub-cloned into the pShuttle-IRES-hrGFP-2 plasmid vector with HA-Tag sequence (Protack et al., 2017). The QuikChange II Site-Directed Mutagenesis Kit (Aligent Technologies, Santa Clara CA, USA) was used to generate isolated single amino-acid changes within the *EPHB4* ORF (A509G, K650N, F867L). All mutant constructs were sequenced to confirm successful mutagenesis. The wild-type *EPHB4*, the mutant constructs, and empty vector were individually transiently transfected into HEK 293T cells using

Lipofectamine 2000 transfection reagent (Invitrogen, Carlsbad, CA, USA) according to its standard protocol. DNA complexes were removed after 5 h and replaced with fresh complete medium. After 48 hours, the medium was aspirated and the cells starved for 18 h in serum-free conditions.

**Co-immunoprecipitation and Western Blotting**

Cell lysates were prepared using NP40 lysis buffer (50 mM Tris, pH7.5; 1% Nonidet P-40; 150 mM NaCl; 10% Glycerol) containing protease and phosphatase inhibitor cocktail (Roche, Basel, Switzerland). Protein concentrations were measured with DC Protein Assay Reagents (Bio-Rad, Hercules CA, USA). For immunoprecipitation, equal amounts of cell lysates were incubated with Sepharose beads linked to anti-HA-Tag antibody (Cell Signaling Technology, Danvers, MA, USA) overnight at 4 °C. Immunoprecipitated protein complexes were separated on SDS-PAGE gel and analyzed by Western blotting using following antibodies: anti-phosphotyrosine P-Tyr-100, anti-HA-Tag, (Cell Signaling Technology, Danvers MA, USA), anti-Ras GAP (Santa Cruz Biotechnology, Dallas TX, USA). Band intensities were quantified using ImageJ software (Schneider et al., 2012). Statistical analyses were performed using Prism 7 software (GraphPad Software, La Jolla CA, USA).

## REFERENCES


Samocha, K., Kosmicki, J., Karczewski, K., O'Donnell-Luria, A., Hoffman, E., MacArthur, D.G., Neale, B., and Daly, M. Regional missense constraint improves variant deleteriousness prediction. bioRxiv.

Long, D.M., Seljeskog, E.L., Chou, S.N., and French, L.A. (1974). Giant arteriovenous malformations of infancy and childhood. J Neurosurg 40, 304-312.

Spetzler, R.F., and Martin, N.A. (1986). A proposed grading system for arteriovenous malformations. J Neurosurg 65, 476-483.

Raybaud, C.A., Strother, C.M., and Hald, J.K. (1989). Aneurysms of the vein of Galen: embryonic considerations and anatomical features relating to the pathogenesis of the malformation. Neuroradiology 31, 109-128.

Laskowski, R.A., Macarthur, M.W., Moss, D.S., and Thornton, J.M. (1993). Procheck - a Program to Check the Stereochemical Quality of Protein Structures. J Appl Crystallogr 26, 283-291.

Abagyan, R., Totrov, M., and Kuznetsov, D. (1994). Icm - a New Method for Protein Modeling and Design - Applications to Docking and Structure Prediction from the Distorted Native Conformation. J Comput Chem 15, 488-506.

Shulman, L.P., Elias, S., Phillips, O.P., Grevengood, C., Dungan, J.S., and Simpson, J.L. (1994). Amniocentesis performed at 14 weeks' gestation or earlier: comparison with first-trimester transabdominal chorionic villus sampling. Obstet Gynecol 83, 543-548.

Pawson, T., and Scott, J.D. (1997). Signaling through scaffold, anchoring, and adaptor proteins. Science 278, 2075-2080.

Chen, Y.G., Hata, A., Lo, R.S., Wotton, D., Shi, Y., Pavletich, N., and Massague, J. (1998). Determinants of specificity in TGF-beta signal transduction. Genes Dev 12, 2144-2152.

McElhinney, D.B., Halbach, V.V., Silverman, N.H., Dowd, C.F., and Hanley, F.L. (1998). Congenital cardiac anomalies with vein of Galen malformations in infants. Arch Dis Child 78, 548-551.

Barbara, N.P., Wrana, J.L., and Letarte, M. (1999). Endoglin is an accessory protein that interacts with the signaling receptor complex of multiple members of the transforming growth factor-beta superfamily. The Journal of biological chemistry 274, 584-594.

Gerety, S.S., Wang, H.U., Chen, Z.F., and Anderson, D.J. (1999). Symmetrical mutant phenotypes of the receptor EphB4 and its specific transmembrane ligand ephrin-B2 in cardiovascular development. Mol Cell 4, 403-414.

Lee, S., Lin, M., Mele, A., Cao, Y., Farmar, J., Russo, D., and Redman, C. (1999). Proteolytic processing of big endothelin-3 by the kell blood group protein. Blood 94, 1440-1450.

Morita, K., Sasaki, H., Furuse, M., and Tsukita, S. (1999). Endothelial claudin: claudin-5/TMVCF constitutes tight junction strands in endothelial cells. J Cell Biol 147, 185-194.

Kniesel, U., and Wolburg, H. (2000). Tight junctions of the blood-brain barrier. Cell Mol Neurobiol 20, 57-76.

Yan, Y., Barlev, N.A., Haley, R.H., Berger, S.L., and Marmorstein, R. (2000). Crystal structure of yeast Esa1 suggests a unified mechanism for catalysis and substrate binding by histone acetyltransferases. Mol Cell 6, 1195-1205.

Machida, Y., Murai, K., Miyake, K., and Iijima, S. (2001). Expression of chromatin remodeling factors during neural differentiation. J Biochem 129, 43-49.

Merscher, S., Funke, B., Epstein, J.A., Heyer, J., Puech, A., Lu, M.M., Xavier, R.J., Demay, M.B., Russell, R.G., Factor, S*., et al.* (2001). TBX1 is responsible for cardiovascular defects in velo-cardio-facial/DiGeorge syndrome. Cell 104, 619-629.

Mitchell, P.J., Rosenfeld, J.V., Dargaville, P., Loughnan, P., Ditchfield, M.R., Frawley, G., and Tress, B.M. (2001). Endovascular management of vein of Galen aneurysmal malformations presenting in the neonatal period. AJNR Am J Neuroradiol 22, 1403-1409.

Brouillard, P., Boon, L.M., Mulliken, J.B., Enjolras, O., Ghassibe, M., Warman, M.L., Tan, O.T., Olsen, B.R., and Vikkula, M. (2002). Mutations in a novel factor, glomulin, are responsible for glomuvenous malformations ("glomangiomas"). American journal of human genetics 70, 866-874.

Kim, I., Ryu, Y.S., Kwak, H.J., Ahn, S.Y., Oh, J.L., Yancopoulos, G.D., Gale, N.W., and Koh, G.Y. (2002). EphB ligand, ephrinB2, suppresses the VEGF- and angiopoietin 1-induced Ras/mitogen-activated protein kinase pathway in venous endothelial cells. FASEB J 16, 1126-1128.

Vitelli, F., Morishima, M., Taddei, I., Lindsay, E.A., and Baldini, A. (2002). Tbx1 mutation causes multiple cardiovascular defects and disrupts neural crest and cranial nerve migratory pathways. Human molecular genetics 11, 915-922.

Wang, Z., Miura, N., Bonelli, A., Mole, P., Carlesso, N., Olson, D.P., and Scadden, D.T. (2002). Receptor tyrosine kinase, EphB4 (HTK), accelerates differentiation of select human hematopoietic cells. Blood 99, 2740-2747.

Abdalla, S.A., Geisthoff, U.W., Bonneau, D., Plauchu, H., McDonald, J., Kennedy, S., Faughnan, M.E., and Letarte, M. (2003). Visceral manifestations in hereditary haemorrhagic telangiectasia type 2. J Med Genet 40, 494-502.

Hsieh, J.J., Ernst, P., Erdjument-Bromage, H., Tempst, P., and Korsmeyer, S.J. (2003). Proteolytic cleavage of MLL generates a complex of N- and C-terminal fragments that confers protein stability and subnuclear localization. Mol Cell Biol 23, 186-194.

Apweiler, R., Bairoch, A., Wu, C.H., Barker, W.C., Boeckmann, B., Ferro, S., Gasteiger, E., Huang, H., Lopez, R., Magrane, M.*, et al.* (2004). UniProt: the Universal Protein knowledgebase. Nucleic Acids Res 32, D115-119.

Harrison, R.E., Berger, R., Haworth, S.G., Tulloh, R., Mache, C.J., Morrell, N.W., Aldred, M.A., and Trembath, R.C. (2005). Transforming growth factor-beta receptor mutations and pulmonary arterial hypertension in childhood. Circulation 111, 435-441.

Tanaka, M., Kamata, R., and Sakai, R. (2005). EphA2 phosphorylates the cytoplasmic tail of Claudin-4 and mediates paracellular permeability. The Journal of biological chemistry 280, 42375-42382.

Wattenhofer, M., Reymond, A., Falciola, V., Charollais, A., Caille, D., Borel, C., Lyle, R., Estivill, X., Petersen, M.B., Meda, P., and Antonarakis, S.E. (2005). Different mechanisms preclude mutant CLDN14 proteins from forming tight junctions in vitro. Hum Mutat 25, 543-549.

Zhang, Z., Cerrato, F., Xu, H., Vitelli, F., Morishima, M., Vincentz, J., Furuta, Y., Ma, L., Martin, J.F., Baldini, A., and Lindsay, E. (2005). Tbx1 expression in pharyngeal epithelia is necessary for pharyngeal arch artery development. Development 132, 5307-5315.

Lasjaunias, P.L., Chng, S.M., Sachet, M., Alvarez, H., Rodesch, G., and Garcia-Monaco, R. (2006). The management of vein of Galen aneurysmal malformations. Neurosurgery 59, S184-194; discussion S183-113.

Shore, E.M., Xu, M., Feldman, G.J., Fenstermacher, D.A., Cho, T.J., Choi, I.H., Connor, J.M., Delai, P., Glaser, D.L., LeMerrer, M.*, et al.* (2006). A recurrent mutation in the BMP type I receptor ACVR1 causes inherited and sporadic fibrodysplasia ossificans progressiva. Nature genetics 38, 525-527.

Zhang, J., and Hughes, S. (2006). Role of the ephrin and Eph receptor tyrosine kinase families in angiogenesis and development of the cardiovascular system. The Journal of pathology 208, 453-461.

Brouillard, P., and Vikkula, M. (2007). Genetic causes of vascular malformations. Human molecular genetics 16 Spec No. 2, R140-149.

Gonzalez-Mariscal, L., Lechuga, S., and Garay, E. (2007). Role of tight junctions in cell proliferation and cancer. Prog Histochem Cytochem 42, 1-57.

Holbert, M.A., Sikorski, T., Carten, J., Snowflack, D., Hodawadekar, S., and Marmorstein, R. (2007). The human monocytic leukemia zinc finger histone acetyltransferase domain contains DNA-binding activity implicated in chromatin targeting. The Journal of biological chemistry 282, 36603-36613.

Potente, M., Ghaeni, L., Baldessari, D., Mostoslavsky, R., Rossig, L., Dequiedt, F., Haendeler, J., Mione, M., Dejana, E., Alt, F.W.*, et al.* (2007). SIRT1 controls endothelial angiogenic functions during vascular growth. Genes Dev 21, 2644-2658.

Wiederstein, M., and Sippl, M.J. (2007). ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res 35, W407-410.

Griffin, C.T., Brennan, J., and Magnuson, T. (2008). The chromatin-remodeling enzyme BRG1 plays an essential role in primitive erythropoiesis and vascular development. Development 135, 493-500.

Makita, T., Sucov, H.M., Gariepy, C.E., Yanagisawa, M., and Ginty, D.D. (2008). Endothelins are vascular-derived axonal guidance cues for developing sympathetic neurons. Nature 452, 759-763.

Pagenstecher, A., Stahl, S., Sure, U., and Felbor, U. (2009). A two-hit mechanism causes cerebral cavernous malformations: complete inactivation of CCM1, CCM2 or CCM3 in affected endothelial cells. Human molecular genetics 18, 911-918.

Zhang, Y.E. (2009). Non-Smad pathways in TGF-beta signaling. Cell Res 19, 128-139.
Adams, R.H., and Eichmann, A. (2010). Axon guidance molecules in vascular patterning. Cold Spring Harb Perspect Biol 2, a001875.

Bilguvar, K., Ozturk, A.K., Louvi, A., Kwan, K.Y., Choi, M., Tatli, B., Yalnizoglu, D., Tuysuz, B., Caglayan, A.O., Gokben, S.*, et al.* (2010). Whole-exome sequencing identifies recessive WDR62 mutations in severe brain malformations. Nature 467, 207-210.

Fischbach, G.D., and Lord, C. (2010). The Simons Simplex Collection: a resource for identification of autism genetic risk factors. Neuron 68, 192-195.

Garcia-Alai, M.M., Allen, M.D., Joerger, A.C., and Bycroft, M. (2010). The structure of the FYR domain of transforming growth factor beta regulator 1. Protein Sci 19, 1432-1438.

Manichaikul, A., Mychaleckyj, J.C., Rich, S.S., Daly, K., Sale, M., and Chen, W.M. (2010). Robust relationship inference in genome-wide association studies. Bioinformatics 26, 2867-2873.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., and DePristo, M.A. (2010). The Genome Analysis Toolkit: a

MapReduce framework for analyzing next-generation DNA sequencing data. Genome research 20, 1297-1303.

Mosch, B., Reissenweber, B., Neuber, C., and Pietzsch, J. (2010). Eph receptors and ephrin ligands: important players in angiogenesis and tumor angiogenesis. J Oncol 2010, 135285.
Salaita, K., and Groves, J.T. (2010). Roles of the cytoskeleton in regulating EphA2 signals. Commun Integr Biol 3, 454-457.

Wang, K., Li, M., and Hakonarson, H. (2010a). ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res 38, e164.

Wang, Y., Nakayama, M., Pitulescu, M.E., Schmidt, T.S., Bochenek, M.L., Sakakibara, A., Adams, S., Davy, A., Deutsch, U., Luthi, U.*, et al.* (2010b). Ephrin-B2 controls VEGF-induced angiogenesis and lymphangiogenesis. Nature 465, 483-486.

Wolfe, J.M., and Myers, L. (2010). Fur in the midst of the waters: visual search for material type is inefficient. J Vis 10, 8.

Xu, D.S., Usman, A.A., Hurley, M.C., Eddleman, C.S., and Bendok, B.R. (2010). Adult presentation of a familial-associated vein of galen aneurysmal malformation: case report. Neurosurgery 67, E1845-1851; discussion 1851.

Barak, T., Kwan, K.Y., Louvi, A., Demirbilek, V., Saygi, S., Tuysuz, B., Choi, M., Boyaci, H., Doerschner, K., Zhu, Y.*, et al.* (2011). Recessive LAMC3 mutations cause malformations of occipital cortical development. Nature genetics 43, 590-594.

Browning, B.L., and Browning, S.R. (2011). A fast, powerful method for detecting identity by descent. American journal of human genetics 88, 173-182.

Hannibal, M.C., Buckingham, K.J., Ng, S.B., Ming, J.E., Beck, A.E., McMillin, M.J., Gildersleeve, H.I., Bigham, A.W., Tabor, H.K., Mefford, H.C.*, et al.* (2011). Spectrum of MLL2 (ALR) mutations in 110 cases of Kabuki syndrome. American journal of medical genetics Part A 155A, 1511-1516.
Ogawa, T., Wakai, C., Saito, T., Murayama, A., Mimura, Y., Youfu, S., Nakamachi, T., Kuwagata, M., Satoh, K., and Shioda, S. (2011). Distribution of the longevity gene product, SIRT1, in developing mouse organs. Congenit Anom (Kyoto) 51, 70-79.

Tsutsumi, Y., Kosaki, R., Itoh, Y., Tsukamoto, K., Matsuoka, R., Shintani, M., Nosaka, S., Masaki, H., and Iizuka, Y. (2011). Vein of Galen aneurysmal malformation associated with an endoglin gene mutation. Pediatrics 128, e1307-1310.

Walker, E.J., Su, H., Shen, F., Choi, E.J., Oh, S.P., Chen, G., Lawton, M.T., Kim, H., Chen, Y., Chen, W., and Young, W.L. (2011). Arteriovenous malformation in the adult mouse brain resembling the human disease. Annals of neurology 69, 954-962.

Chaikuad, A., Alfano, I., Kerr, G., Sanvitale, C.E., Boergermann, J.H., Triffitt, J.T., von Delft, F., Knapp, S., Knaus, P., and Bullock, A.N. (2012). Structure of the bone morphogenetic protein receptor ALK2 and implications for fibrodysplasia ossificans progressiva. The Journal of biological chemistry 287, 36990-36998.

Deloison, B., Chalouhi, G.E., Sonigo, P., Zerah, M., Millischer, A.E., Dumez, Y., Brunelle, F., Ville, Y., and Salomon, L.J. (2012). Hidden mortality of prenatally diagnosed vein of Galen aneurysmal malformation: retrospective study and review of the literature. Ultrasound Obstet Gynecol 40, 652-658.

Guo, S., Zhou, Y., Xing, C., Lok, J., Som, A.T., Ning, M., Ji, X., and Lo, E.H. (2012). The vasculome of the mouse brain. PloS one 7, e52665.

Kim, J.H., Peacock, M.R., George, S.C., and Hughes, C.C. (2012). BMP9 induces EphrinB2 expression in endothelial cells through an Alk1-BMPRII/ActRII-ID1/ID3-dependent pathway: implications for hereditary hemorrhagic telangiectasia type II. Angiogenesis 15, 497-509.

Mura, M., Cappato, S., Giacopelli, F., Ravazzolo, R., and Bocciardi, R. (2012). The role of the 3'UTR region in the regulation of the ACVR1/Alk-2 gene expression. PloS one 7, e50958.

Neale, B.M., Kou, Y., Liu, L., Ma'ayan, A., Samocha, K.E., Sabo, A., Lin, C.F., Stevens, C., Wang, L.S., Makarov, V.*, et al.* (2012). Patterns and rates of exonic de novo mutations in autism spectrum disorders. Nature 485, 242-245.

Recinos, P.F., Rahmathulla, G., Pearl, M., Recinos, V.R., Jallo, G.I., Gailloud, P., and Ahn, E.S. (2012). Vein of Galen malformations: epidemiology, clinical presentations, management. Neurosurg Clin N Am 23, 165-177.

Sanchez-Carpintero, R., Herranz, A., Reynoso, C., and Zubieta, J.L. (2012). Dilated vein of Galen in Kabuki syndrome. Brain Dev 34, 76-79.

Sanders, S.J., Murtha, M.T., Gupta, A.R., Murdoch, J.D., Raubeson, M.J., Willsey, A.J., Ercan-Sencicek, A.G., DiLullo, N.M., Parikshak, N.N., Stein, J.L.*, et al.* (2012). De novo mutations revealed by whole-exome sequencing are strongly associated with autism. Nature 485, 237-241.

Schneider, C.A., Rasband, W.S., and Eliceiri, K.W. (2012). NIH Image to ImageJ: 25 years of image analysis. Nature methods 9, 671-675.

Van Houdt, J.K., Nowakowska, B.A., Sousa, S.B., van Schaik, B.D., Seuntjens, E., Avonce, N., Sifrim, A., Abdul-Rahman, O.A., van den Boogaard, M.J., Bottani, A.*, et al.* (2012). Heterozygous missense mutations in SMARCA2 cause Nicolaides-Baraitser syndrome. Nature genetics 44, 445-449, S441.

Xiao, Z., Carrasco, R., Kinneer, K., Sabol, D., Jallal, B., Coats, S., and Tice, D.A. (2012). EphB4 promotes or suppresses Ras/MEK/ERK pathway in a context-dependent manner: Implications for EphB4 as a cancer target. Cancer Biol Ther 13, 630-637.

Chida, A., Shintani, M., Wakamatsu, H., Tsutsumi, Y., Iizuka, Y., Kawaguchi, N., Furutani, Y., Inai, K., Nonoyama, S., and Nakanishi, T. (2013). ACVRL1 gene variant in a patient with vein of Galen aneurysmal malformation. J Pediatr Genet 2, 181-189.

Davis, R.B., Curtis, C.D., and Griffin, C.T. (2013). BRG1 promotes COUP-TFII expression and venous specification during embryonic vascular development. Development 140, 1272-1281.
Lisabeth, E.M., Falivelli, G., and Pasquale, E.B. (2013). Eph receptor signaling and ephrins. Cold Spring Harb Perspect Biol 5.

Revencu, N., Boon, L.M., Mendola, A., Cordisco, M.R., Dubois, J., Clapuyt, P., Hammer, F., Amor, D.J., Irvine, A.D., Baselga, E*., et al.* (2013). RASA1 mutations and associated phenotypes in 68 families with capillary malformation-arteriovenous malformation. Hum Mutat 34, 1632-1641.

Van der Auwera, G.A., Carneiro, M.O., Hartl, C., Poplin, R., Del Angel, G., Levy-Moonshine, A., Jordan, T., Shakir, K., Roazen, D., Thibault, J*., et al.* (2013). From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. Curr Protoc Bioinformatics 43, 11 10 11-33.

Van Itallie, C.M., and Anderson, J.M. (2013). Claudin interactions in and out of the tight junction. Tissue Barriers 1, e25247.

Zaidi, S., Choi, M., Wakimoto, H., Ma, L., Jiang, J., Overton, J.D., Romano-Adesman, A., Bjornson, R.D., Breitbart, R.E., Brown, K.K*., et al.* (2013). De novo mutations in histone-modifying genes in congenital heart disease. Nature 498, 220-223.

Altschul, D., Paramasivam, S., Ortega-Gutierrez, S., Fifi, J.T., and Berenstein, A. (2014). Safety and efficacy using a detachable tip microcatheter in the embolization of pediatric arteriovenous malformations. Childs Nerv Syst 30, 1099-1107.

De Rubeis, S., He, X., Goldberg, A.P., Poultney, C.S., Samocha, K., Cicek, A.E., Kou, Y., Liu, L., Fromer, M., Walker, S*., et al.* (2014). Synaptic, transcriptional and chromatin genes disrupted in autism. Nature 515, 209-215.

Iossifov, I., O'Roak, B.J., Sanders, S.J., Ronemus, M., Krumm, N., Levy, D., Stessman, H.A., Witherspoon, K.T., Vives, L., Patterson, K.E*., et al.* (2014). The contribution of de novo coding mutations to autism spectrum disorder. Nature 515, 216-221.

Kawasaki, J., Aegerter, S., Fevurly, R.D., Mammoto, A., Mammoto, T., Sahin, M., Mably, J.D., Fishman, S.J., and Chan, J. (2014). RASA1 functions in EPHB4 signaling pathway to suppress endothelial mTORC1 activity. J Clin Invest 124, 2774-2784.

Kramer, A., Green, J., Pollard, J., Jr., and Tugendreich, S. (2014). Causal analysis approaches in Ingenuity Pathway Analysis. Bioinformatics 30, 523-530.

Li, H. (2014). Toward better understanding of artifacts in variant calling from high-coverage samples. Bioinformatics 30, 2843-2851.

McCarthy, S.E., Gillis, J., Kramer, M., Lihm, J., Yoon, S., Berstein, Y., Mistry, M., Pavlidis, P., Solomon, R., Ghiban, E.*, et al.* (2014). De novo mutations in schizophrenia implicate chromatin remodeling and support a genetic overlap with autism and intellectual disability. Molecular psychiatry 19, 652-658.

Mishra-Gorur, K., Caglayan, A.O., Schaffer, A.E., Chabu, C., Henegariu, O., Vonhoff, F., Akgumus, G.T., Nishimura, S., Han, W., Tu, S.*, et al.* (2014). Mutations in KATNB1 cause complex cerebral malformations by disrupting asymmetrically dividing neural progenitors. Neuron 84, 1226-1239.
Samocha, K.E., Robinson, E.B., Sanders, S.J., Stevens, C., Sabo, A., McGrath, L.M., Kosmicki, J.A., Rehnstrom, K., Mallick, S., Kirby, A.*, et al.* (2014). A framework for the interpretation of de novo mutation in human disease. Nature genetics 46, 944-950.

Son, E.Y., and Crabtree, G.R. (2014). The role of BAF (mSWI/SNF) complexes in mammalian neural development. Am J Med Genet C Semin Med Genet 166C, 333-349.

Venugopal, A. (2014). Disseminated intravascular coagulation. Indian J Anaesth 58, 603-608.
Arboleda, V.A., Lee, H., Dorrani, N., Zadeh, N., Willis, M., Macmurdo, C.F., Manning, M.A., Kwan, A., Hudgins, L., Barthelemy, F.*, et al.* (2015). De novo nonsense mutations in KAT6A, a lysine acetyl-transferase gene, cause a syndrome including microcephaly and global developmental delay. American journal of human genetics 96, 498-506.

Dentici, M.L., Di Pede, A., Lepri, F.R., Gnazzo, M., Lombardi, M.H., Auriti, C., Petrocchi, S., Pisaneschi, E., Bellacchio, E., Capolino, R.*, et al.* (2015). Kabuki syndrome: clinical and molecular diagnosis in the first year of life. Arch Dis Child 100, 158-164.

Dong, C., Wei, P., Jian, X., Gibbs, R., Boerwinkle, E., Wang, K., and Liu, X. (2015). Comparison and integration of deleteriousness prediction methods for nonsynonymous SNVs in whole exome sequencing studies. Human molecular genetics 24, 2125-2137.

Ferguson, B.D., Tan, Y.H., Kanteti, R.S., Liu, R., Gayed, M.J., Vokes, E.E., Ferguson, M.K., Iafrate, A.J., Gill, P.S., and Salgia, R. (2015). Novel EPHB4 Receptor Tyrosine Kinase Mutations and Kinomic Pathway Analysis in Lung Cancer. Sci Rep 5, 10641.

Fish, J.E., and Wythe, J.D. (2015). The molecular regulation of arteriovenous specification and maintenance. Developmental dynamics : an official publication of the American Association of Anatomists 244, 391-409.

Homsy, J., Zaidi, S., Shen, Y., Ware, J.S., Samocha, K.E., Karczewski, K.J., DePalma, S.R., McKean, D., Wakimoto, H., Gorham, J.*, et al.* (2015). De novo mutations in congenital heart disease with neurodevelopmental and other congenital anomalies. Science 350, 1262-1266.

Krumm, N., Turner, T.N., Baker, C., Vives, L., Mohajeri, K., Witherspoon, K., Raja, A., Coe, B.P., Stessman, H.A., He, Z.X.*, et al.* (2015). Excess of rare, inherited truncating mutations in autism. Nature genetics 47, 582-588.

Pollen, A.A., Nowakowski, T.J., Chen, J., Retallack, H., Sandoval-Espinosa, C., Nicholas, C.R., Shuga, J., Liu, S.J., Oldham, M.C., Diaz, A.*, et al.* (2015). Molecular identity of human outer radial glia during cortical development. Cell 163, 55-67.

Stuart, B.D., Choi, J., Zaidi, S., Xing, C., Holohan, B., Chen, R., Choi, M., Dharwadkar, P., Torres, F., Girod, C.E.*, et al.* (2015). Exome sequencing links mutations in PARN and RTEL1 with familial pulmonary fibrosis and telomere shortening. Nature genetics 47, 512-517.

Szklarczyk, D., Franceschini, A., Wyder, S., Forslund, K., Heller, D., Huerta-Cepas, J., Simonovic, M., Roth, A., Santos, A., Tsafou, K.P.*, et al.* (2015). STRING v10: protein-protein interaction networks, integrated over the tree of life. Nucleic Acids Res 43, D447-452.

Tham, E., Lindstrand, A., Santani, A., Malmgren, H., Nesbitt, A., Dubbs, H.A., Zackai, E.H., Parker, M.J., Millan, F., Rosenbaum, K.*, et al.* (2015). Dominant mutations in KAT6A cause intellectual disability with recognizable syndromic features. American journal of human genetics 96, 507-513.

Van Laarhoven, P.M., Neitzel, L.R., Quintana, A.M., Geiger, E.A., Zackai, E.H., Clouthier, D.E., Artinger, K.B., Ming, J.E., and Shaikh, T.H. (2015). Kabuki syndrome genes KMT2D and KDM6A: functional analyses demonstrate critical roles in craniofacial, heart and brain development. Human molecular genetics 24, 4443-4453.

Vanyai, H.K., Thomas, T., and Voss, A.K. (2015). Mesodermal expression of Moz is necessary for cardiac septum development. Developmental biology 403, 22-29.

Ware, J.S., Samocha, K.E., Homsy, J., and Daly, M.J. (2015). Interpreting de novo Variation in Human Disease Using denovolyzeR. Current protocols in human genetics / editorial board, Jonathan L Haines [et al] 87, 7 25 21-15.

Wei, Q., Zhan, X., Zhong, X., Liu, Y., Han, Y., Chen, W., and Li, B. (2015). A Bayesian framework for de novo mutation calling in parents-offspring trios. Bioinformatics 31, 1375-1381.

Wiley, M.M., Muthukumar, V., Griffin, T.M., and Griffin, C.T. (2015). SWI/SNF chromatin-remodeling enzymes Brahma-related gene 1 (BRG1) and Brahma (BRM) are dispensable in multiple models of postnatal angiogenesis but are required for vascular integrity in infant mice. J Am Heart Assoc 4.

Zhang, S., Zhou, B., Kang, Y., Cui, X., Liu, A., Deleris, A., Greenberg, M.V., Cui, X., Qiu, Q., Lu, F., *et al.* (2015). C-terminal domains of a histone demethylase interact with a pair of transcription factors and mediate specific chromatin association. Cell Discov 1.

Gerrard, D.T., Berry, A.A., Jennings, R.E., Piper Hanley, K., Bobola, N., and Hanley, N.A. (2016). An integrative transcriptomic atlas of organogenesis in human embryos. Elife 5.

Lek, M., Karczewski, K.J., Minikel, E.V., Samocha, K.E., Banks, E., Fennell, T., O'Donnell-Luria, A.H., Ware, J.S., Hill, A.J., Cummings, B.B., *et al.* (2016). Analysis of protein-coding genetic variation in 60,706 humans. Nature 536, 285-291.

Roth Flach, R.J., Guo, C.A., Danai, L.V., Yawe, J.C., Gujja, S., Edwards, Y.J., and Czech, M.P. (2016). Endothelial Mitogen-Activated Protein Kinase Kinase Kinase Kinase 4 Is Critical for Lymphatic Vascular Development and Function. Mol Cell Biol 36, 1740-1749.

Timberlake, A.T., Choi, J., Zaidi, S., Lu, Q., Nelson-Williams, C., Brooks, E.D., Bilguvar, K., Tikhonova, I., Mane, S., Yang, J.F., *et al.* (2016). Two locus inheritance of non-syndromic midline craniosynostosis via rare SMAD6 and common BMP2 alleles. Elife 5.

Amyere, M., Revencu, N., Helaers, R., Pairet, E., Baselga, E., Cordisco, M., Chung, W., Dubois, J., Lacour, J.P., Martorell, L., *et al.* (2017). Germline Loss-of-Function Mutations in EPHB4 Cause a Second Form of Capillary Malformation-Arteriovenous Malformation (CM-AVM2) Deregulating RAS-MAPK Signaling. Circulation 136, 1037-1048.

Besse, W., Dong, K., Choi, J., Punia, S., Fedeles, S.V., Choi, M., Gallagher, A.R., Huang, E.B., Gulati, A., Knight, J., *et al.* (2017). Isolated polycystic liver disease genes define effectors of polycystin-1 function. J Clin Invest 127, 1772-1785.

Jin, S.C., Homsy, J., Zaidi, S., Lu, Q., Morton, S., DePalma, S.R., Zeng, X., Qi, H., Chang, W., Sierant, M.C., *et al.* (2017). Contribution of rare inherited and de novo variants in 2,871 congenital heart disease probands. Nature genetics 49, 1593-1601.

Protack, C.D., Foster, T.R., Hashimoto, T., Yamamoto, K., Lee, M.Y., Kraehling, J.R., Bai, H., Hu, H., Isaji, T., Santana, J.M., *et al.* (2017). Eph-B4 regulates adaptive venous remodeling to improve arteriovenous fistula patency. Sci Rep 7, 15386.

Roman, B.L., and Hinck, A.P. (2017). ALK1 signaling in development and disease: new paradigms. Cell Mol Life Sci 74, 4539-4560.

Timberlake, A.T., Furey, C.G., Choi, J., Nelson-Williams, C., Yale Center for Genome, A., Loring, E., Galm, A., Kahle, K.T., Steinbacher, D.M., Larysz, D., *et al.* (2017). De novo mutations in inhibitors of Wnt, BMP, and Ras/ERK signaling pathways in non-syndromic midline craniosynostosis. Proceedings of the National Academy of Sciences of the United States of America 114, E7341-E7347.

Willsey, A.J., Fernandez, T.V., Yu, D., King, R.A., Dietrich, A., Xing, J., Sanders, S.J., Mandell, J.D., Huang, A.Y., Richer, P.*, et al.* (2017). De Novo Coding Variants Are Strongly Associated with Tourette Disorder. Neuron 94, 486-499 e489.

Zhang, J., Schwartz, M.P., Hou, Z., Bai, Y., Ardalani, H., Swanson, S., Steill, J., Ruotti, V., Elwell, A., Nguyen, B.K.*, et al.* (2017). A Genome-wide Analysis of Human Pluripotent Stem Cell-Derived Endothelial Cells in 2D or 3D Culture. Stem Cell Reports 8, 907-918.

Duran, D., Karschnia, P., Gaillard, J.R., Karimy, J.K., Youngblood, M.W., DiLuna, M.L., Matouk, C.C., Aagaard-Kienitz, B., Smith, E.R., Orbach, D.B.*, et al.* (2018). Human genetics and molecular mechanisms of vein of Galen malformation. J Neurosurg Pediatr, 1-8.

Moon, J.E., Lee, S.J., and Ko, C.W. (2018). A de novo KMT2D mutation in a girl with Kabuki syndrome associated with endocrine symptoms: a case report. BMC Med Genet 19, 102.

Nikolaev, S.I., Vetiska, S., Bonilla, X., Boudreau, E., Jauhiainen, S., Rezai Jahromi, B., Khyzha, N., DiStefano, P.V., Suutarinen, S., Kiehl, T.R.*, et al.* (2018). Somatic Activating KRAS Mutations in Arteriovenous Malformations of the Brain. The New England journal of medicine 378, 250-261.

Vivanti, A., Ozanne, A., Grondin, C., Saliou, G., Quevarec, L., Maurey, H., Aubourg, P., Benachi, A., Gut, M., Gut, I.*, et al.* (2018). Loss of function mutations in EPHB4 are responsible for vein of Galen aneurysmal malformation. Brain : a journal of neurology 141, 979-988.

Table

# MAIN TABLES

## Table 1. Enrichment of damaging mutations in chromatin modifiers intolerant to LoF mutation in VOGM

| Table 1. Enrichment of damaging mutations in chromatin modifiers intolerant to LoF mutation in VOGM | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Cases, N = 52 | | | | | | Controls, N=1,789 | | | | | |
| | Observed | | Expected | | Enrichment | *p* | | Observed | | Expected | | Enrichment | *p* |
| | N | Rate | N | Rate | | | | N | Rate | N | Rate | | |
| **LoF-intolerant genes (n = 3,230)** | | | | | | | **LoF-intolerant genes (n = 3,230)** | | | | | | |
| Total | 16 | 0.31 | 15.60 | 0.30 | 1.03 | 0.49 | Total | 504 | 0.28 | 525.10 | 0.29 | 0.96 | 0.83 |
| Syn | 2 | 0.04 | 4.30 | 0.08 | 0.46 | 0.93 | Syn | 127 | 0.07 | 145.00 | 0.08 | 0.88 | 0.94 |
| Mis | 13 | 0.25 | 9.80 | 0.19 | 1.33 | 0.19 | Mis | 339 | 0.19 | 330.50 | 0.18 | 1.03 | 0.33 |
| D-Mis | 6 | 0.12 | 2.10 | 0.04 | 2.82 | 0.02 | D-Mis | 64 | 0.04 | 72.50 | 0.04 | 0.88 | 0.86 |
| LoF | 1 | 0.02 | 1.50 | 0.03 | 0.69 | 0.77 | LoF | 38 | 0.02 | 49.50 | 0.03 | 0.77 | 0.96 |
| Damaging | 7 | 0.13 | 3.60 | 0.07 | 1.96 | 0.07 | Damaging | 102 | 0.06 | 122.10 | 0.07 | 0.84 | 0.97 |
| **All chromatin modifier genes (n = 547)** | | | | | | | **All chromatin modifier genes (n = 547)** | | | | | | |
| Total | 4 | 0.08 | 2.30 | 0.04 | 1.73 | 0.20 | Total | 65 | 0.04 | 78.30 | 0.04 | 0.83 | 0.94 |
| Syn | 0 | 0.00 | NA | NA | NA | NA | Syn | 12 | 0.01 | 21.30 | 0.01 | 0.56 | 0.99 |
| Mis | 3 | 0.06 | 1.50 | 0.03 | 2.06 | 0.18 | Mis | 47 | 0.03 | 49.40 | 0.03 | 0.95 | 0.65 |
| D-Mis | 3 | 0.06 | 0.30 | 0.01 | 8.80 | $5.1x10^{-3}$ | D-Mis | 12 | 0.01 | 11.60 | 0.01 | 1.03 | 0.50 |
| LoF | 1 | 0.02 | 0.20 | 0.00 | 4.53 | 0.2 | LoF | 6 | 0.00 | 7.60 | 0.00 | 0.80 | 0.76 |
| Damaging | 4 | 0.08 | 0.60 | 0.01 | 7.12 | $2.7x10^{-3}$ | Damaging | 18 | 0.01 | 19.20 | 0.01 | 0.94 | 0.64 |
| **LoF-Intolerant chromatin modifier genes (n = 272)** | | | | | | | **LoF-Intolerant chromatin modifier genes (n = 272)** | | | | | | |
| Total | 4 | 0.08 | 1.60 | 0.03 | 2.56 | 0.07 | Total | 47 | 0.03 | 52.90 | 0.03 | 0.89 | 0.81 |
| Syn | 0 | 0.00 | NA | NA | NA | NA | Syn | 9 | 0.01 | 14.30 | 0.01 | 0.63 | 0.95 |
| Mis | 3 | 0.06 | 1.00 | 0.02 | 3.04 | 0.08 | Mis | 35 | 0.02 | 33.40 | 0.02 | 1.05 | 0.42 |
| D-Mis | 3 | 0.06 | 0.30 | 0.01 | 11.40 | $2.5x10^{-3}$ | D-Mis | 7 | 0.00 | 9.00 | 0.01 | 0.78 | 0.79 |
| LoF | 1 | 0.02 | 0.20 | 0.00 | 6.55 | 0.14 | LoF | 3 | 0.00 | 5.20 | 0.00 | 0.57 | 0.89 |
| Damaging | 4 | 0.08 | 0.40 | 0.01 | 9.63 | $8.9x10^{-4}$ | Damaging | 10 | 0.01 | 14.20 | 0.01 | 0.70 | 0.90 |

LoF: loss-of-function; N: the number of de novo mutations; Rate: the number of de novo mutations divided by the number of individuals in the cohort; Enrichment: ratio of observed to expected numbers of mutations; Intolerant genes: Genes with a pLI ≥ 0.90; D-Mis: Damaging missense mutations as predicted by MetaSVM; Damaging: D-Mis + LoF. Chromatin genes used for analysis were extracted from the Biomart database using GO:0016569 as the input.

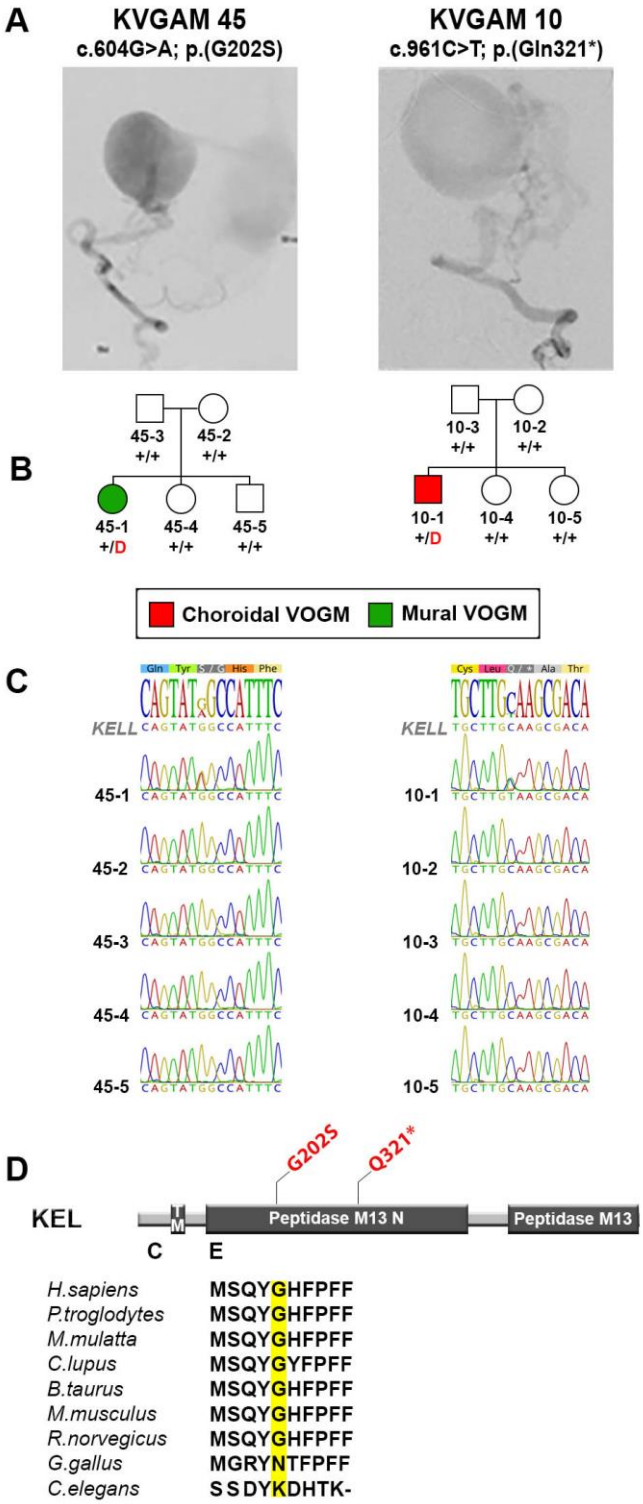
## Table 2. Transmitted mutations in *EPHB4* and *CLDN14* in VOGM

| Table 2. Transmitted mutations in *EPHB4* and *CLDN14* in VOGM | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Family** | **Type of VOGM** | **Ethnicity** | **Gene** | **Mutation** | **Domain affected** | **ExAC MAF** | **gnomAD MAF** | **pLI** | **MetaSVM** | **CADD** |
| VGAM115 | Choroidal | European | *EPHB4* | p.(Glu432fs1) | N/A | novel | novel | 0.99 | N/A | N/A |
| KVGAM25 | Choroidal | European | *EPHB4* | p.(Ala509Gly) | Fibronectin III | 3.30E-05 | 1.13E-05 | 0.99 | D | 25.0 |
| KVGAM33 | Choroidal | Mexican | *EPHB4* | p.(Lys650Asn) | Tyrosine kinase | novel | novel | 0.99 | D | 29.9 |
| KVGAM18 | Choroidal | European | *EPHB4* | p.(Phe867Leu) | Tyrosine kinase | novel | novel | 0.99 | D | 31.0 |
| VGAM100 | Choroidal | Mexican | *CLDN14* | p.(Ala113Pro) | Second intracellular segment | novel | 8.28E-06 | 0.00 | D | 24.2 |
| KVGAM20 | Choroidal | European | *CLDN14* | p.(Val143Met) | Second extracellular segment | 9.43E-06 | 1.24E-05 | 0.00 | D | 31.0 |
| KVGAM51 | Mural | European | *CLDN14* | p.(Val143Met) | Second extracellular segment | 9.43E-06 | 1.24E-05 | 0.00 | D | 31.0 |

Figure

## MAIN FIGURES

### Figure 1. Damaging *de novo KEL* mutations in choroidal Vein of Galen malformation



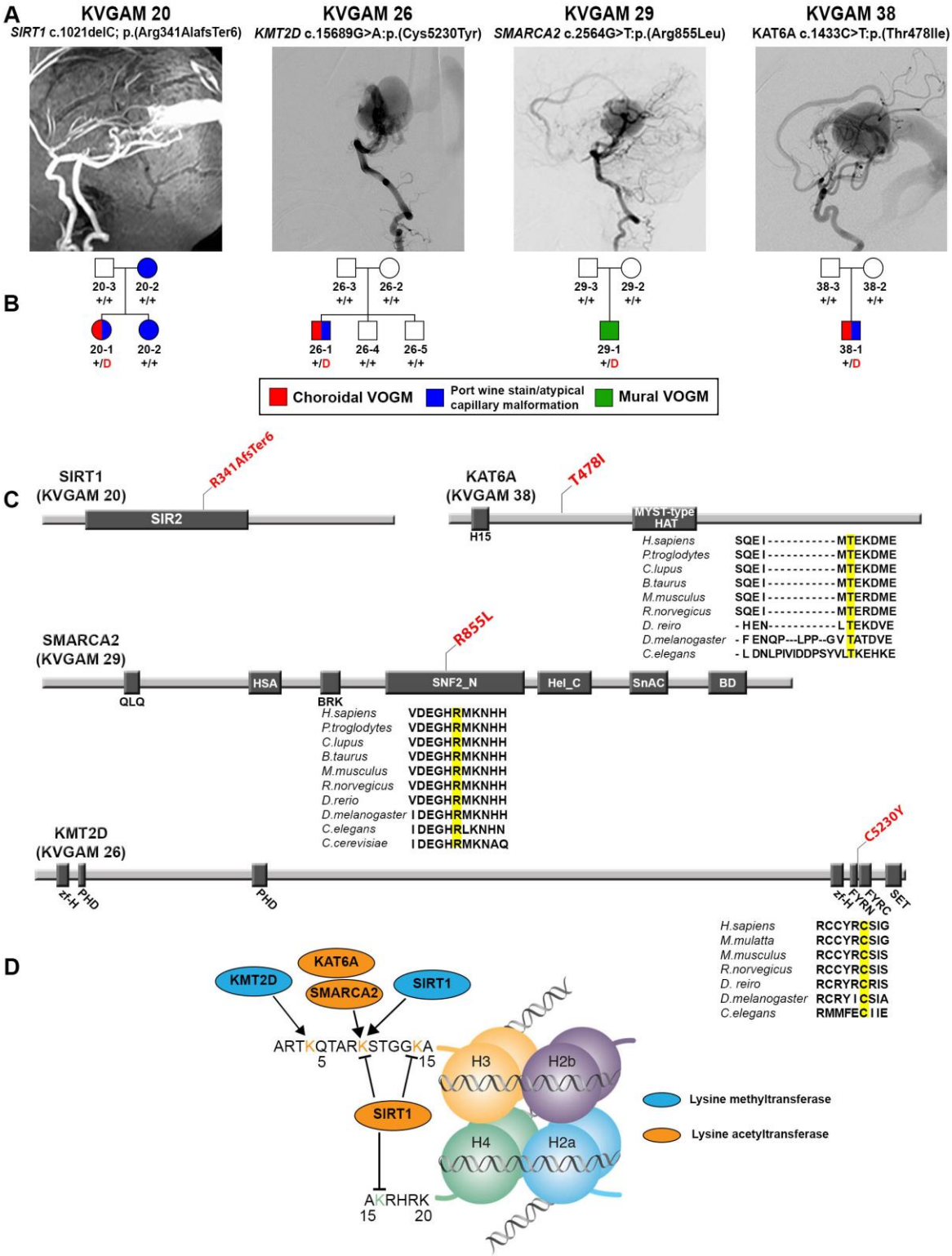

**(A)** Representative magnetic resonance angiography and digital subtraction angiography images demonstrating mural and choroidal Vein of Galen malformations in KVGAM45-1 and KVGAM10-1, respectively.

**(B)** Pedigrees depicting kindred structures. Note that probands carrying *de novo* mutations in *KEL* are the only family member in families KVGAM45 and KVGAM10 that demonstrates Vein of Galen malformations; none of the family members in these two families have any disease phenotypes or cutaneous manifestations.

**(C)** Mutations identified by exome sequencing were confirmed by direct PCR amplification with custom primers followed by Sanger sequencing.

**(D)** Linear representation of the KEL molecule, with functional domains represented as dark rectangles. Amino acid modifications are mapped (in red) on the protein structure. Conservation of the wild-type amino acid substituted by the missense mutations is depicted below. TM = Transmembrane domain.

## Figure 2. Damaging *de novo* mutations in chromatin modifiers in Vein of Galen malformation



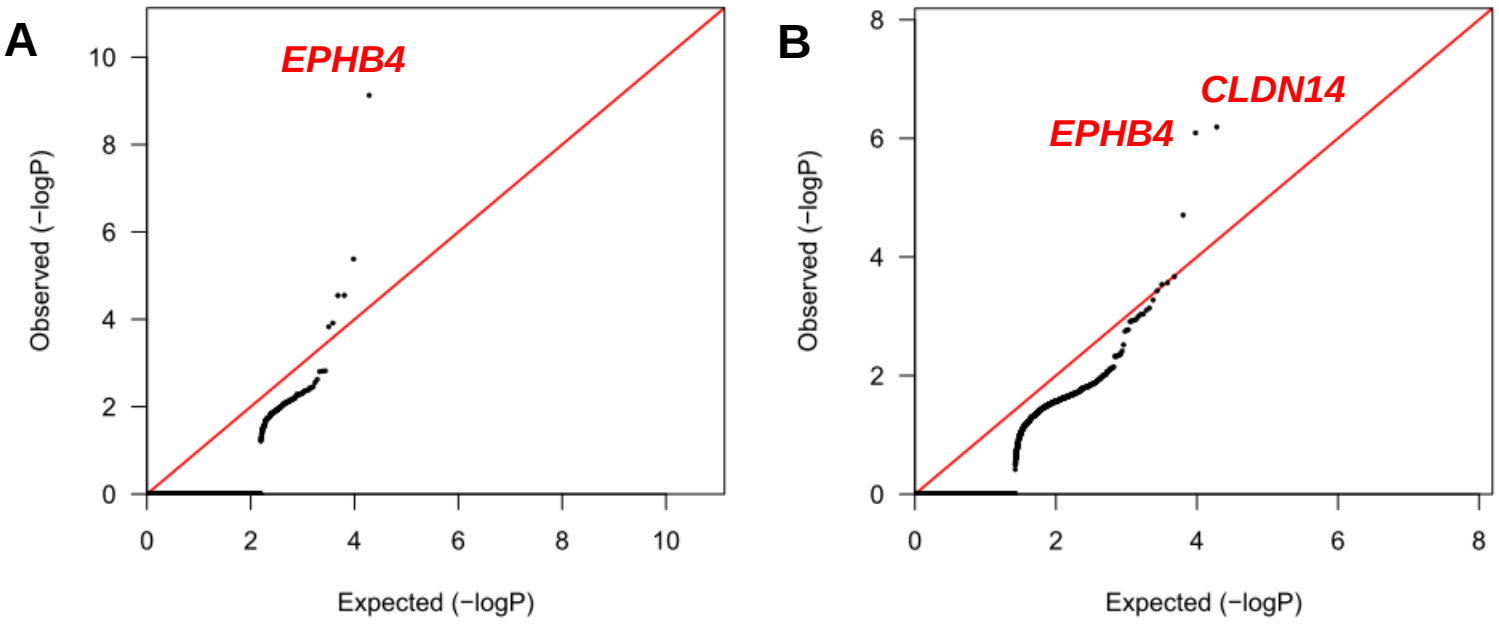

**(A)** Vascular imaging of probands. Magnetic resonance angiographies and a digital subtraction angiography reconstruction demonstrating choroidal Vein of Galen malformations are shown.

**(B)** Pedigree structures of VOGM kindreds. For each kindred, the gene and mutation, the angiographic image, and pedigree structure are shown. Subjects with atypical capillary malformations are denoted by blue symbols. Red 'D' denotes damaging mutation, '+' denotes wild type sequence.

**(C)** Linear representation of functional domains of SIRT1, KMT2D, SMARCA2, and KAT6A with location of VOGM mutations. Functional domains are represented by dark rectangles. Amino acid changes are located on the protein structure in red lettering. For missense mutations, conservation of the wild-type amino acid across phylogeny is shown. The mutated amino acid is highlighted in yellow. SIR2 = Sirtuin catalytic domain, SIR2 Domain; PHD = Zinc Finger PHD type; MOZ_SAS = Histone acetyltransferase domain, MYST-type; zf-H = PHD-like zinc binding domain; FYRN = F/Y-rich domain - F/Y-rich N-terminus motif; FYRC = F/Y-rich domain - F/Y-rich C-terminus motif; SET = Su(var)3-9, Enhancer-of-zeste and Trithorax; QLQ = Glutamine-Leucine-Glutamine domain; HAS = Helicase-SANT associated domain; BRK = BRK domain; SNF2_N = SNF2-related, N-terminal domain; Hel_C = Helicase C-terminal domain; SnAC = Snf2-ATP coupling, chromatin remodeling complex; BD = Bromodomain.

**(D)** Schematic representation of modifications on histone marks exerted by SIRT1, KMT2D, SMARCA2, and KAT6A.

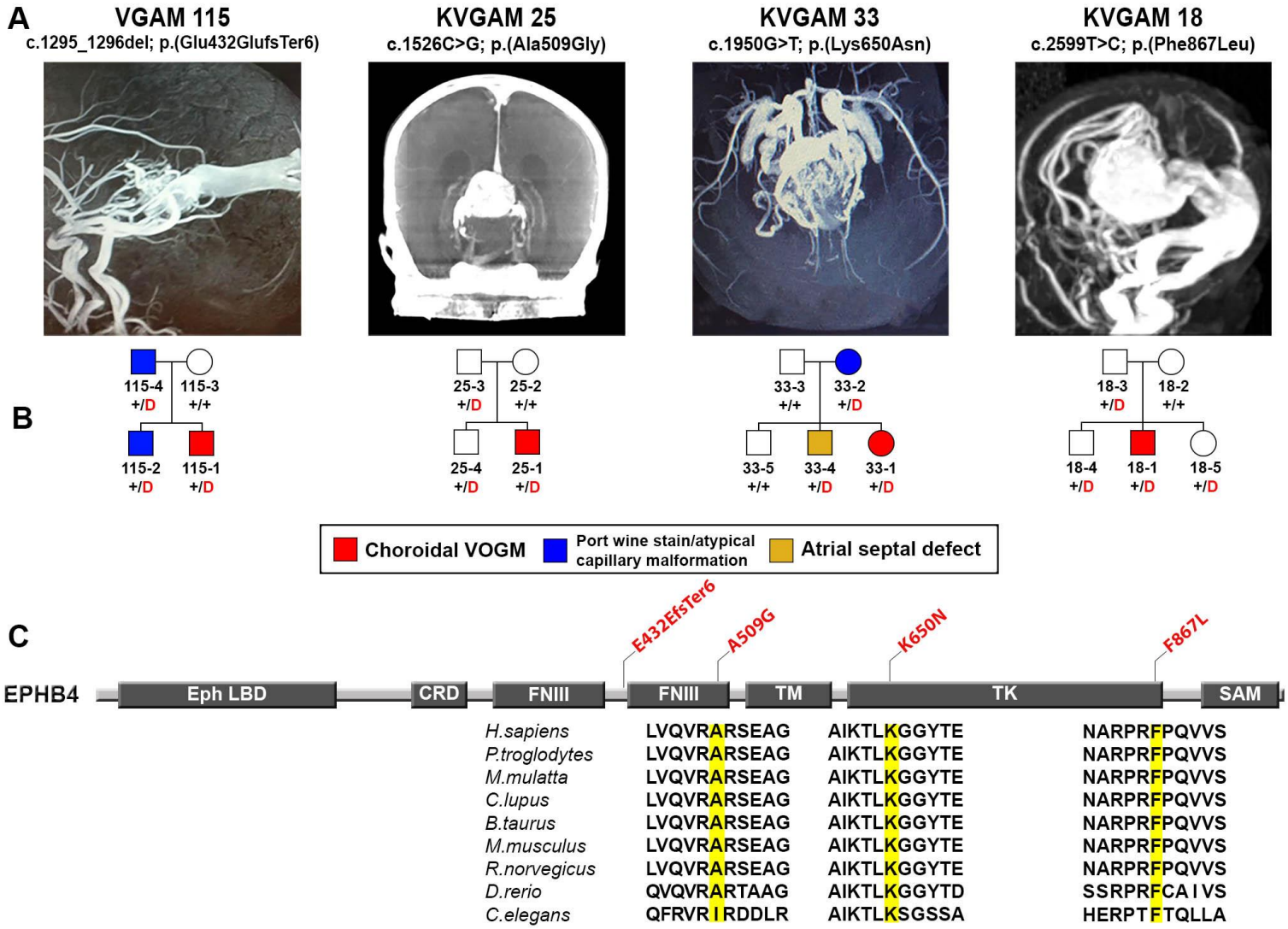
**Figure 3. Exome-wide significant enrichment of rare damaging transmitted mutations in *EPHB4* and *CLDN14***




**(A)** Quantile-quantile plots of observed vs. expected binomial test p-values for rare damaging (D-mis+LoF) variants with MAF < $2x10^{-5}$ in the Genome Aggregation database (gnomAD) in each gene that is intolerant to LoF (pLI > 0.9).

**(B)** Quantile-quantile plots of observed vs. expected binomial test p-values for rare damaging (D-mis+LoF) variants with MAF < $2x10^{-5}$ in gnomAD for all genes. MAF = Minor allele frequency; D-mis = Missense mutations predicted to be deleterious per MetaSVM; LoF = Canonical loss-of-function mutations (stop-gains, frameshift insertions or deletions, canonical splice site mutations).

## Figure 4. Damaging *EPHB4* mutations in choroidal Vein of Galen malformation




**(A)** Vascular imaging of probands from digital subtraction angiographies and magnetic resonance angiography demonstrating Vein of Galen malformations.

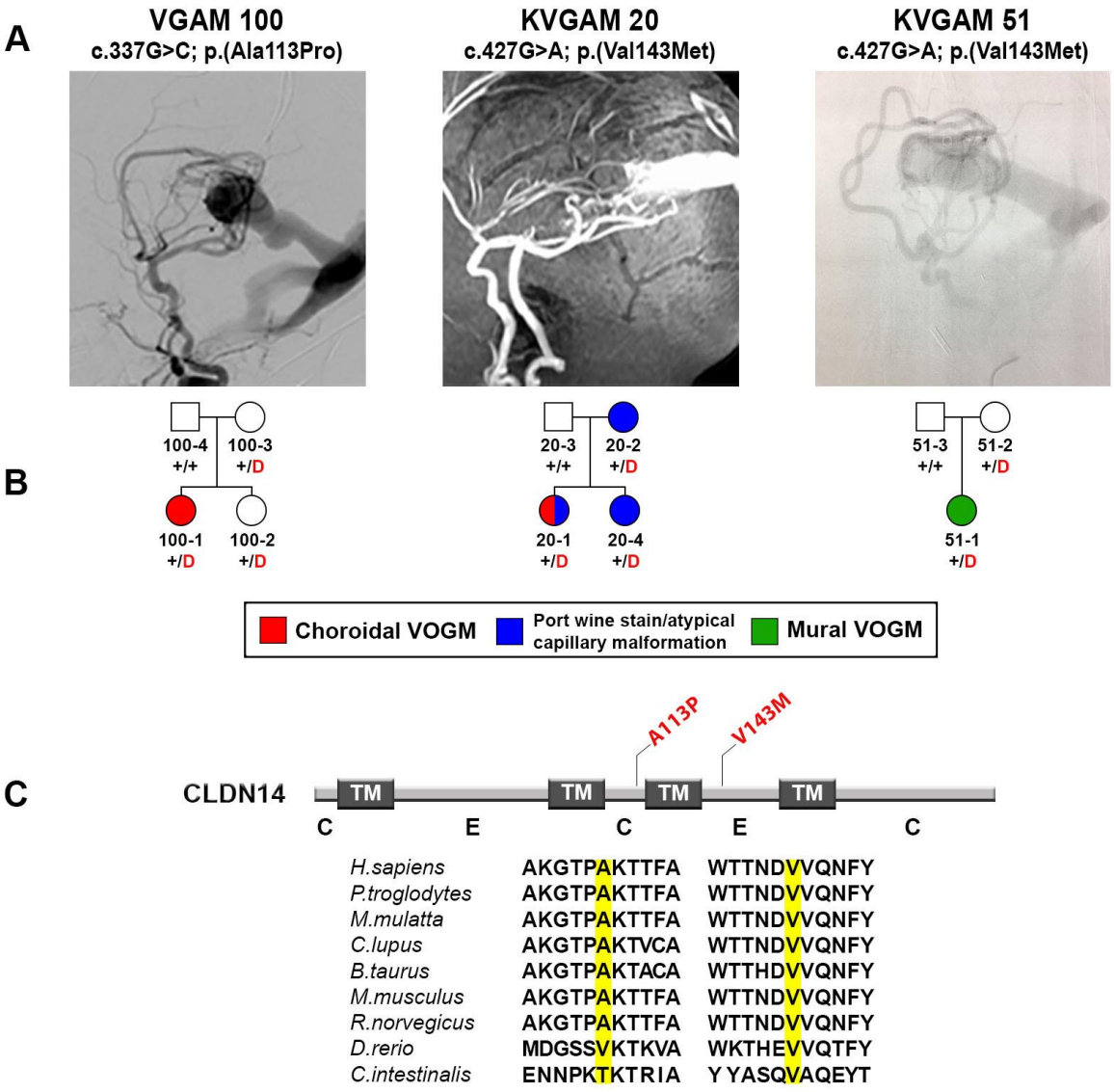
**(B)** Pedigree structures of the kindreds. For each kindred, the gene and mutation, the angiographic image, and pedigree structure are shown. A carrier harboring an atrial septal defect is represented by a yellow symbol in family KVGAM 33.

**(C)** Linear representation of Eph-B4 functional domains (dark rectangles) with location of VOGM mutations (in red). Conservation across phylogeny of the wild-type amino acid (in yellow) at each mutated position is shown. LB = ligand binding domain; CRD = Cysteine-rich domain; FNIII = Fibronectin III domain; TM = Transmembrane domain; TK = Tyrosine kinase domain; SAM = Sterile alpha motif.

**(D)** Representative immunoblots showing the effect of Ala509Gly, Lys650Asn, and Phe867Leu mutations on resting state Eph-B4 tyrosine phosphorylation in HEK 293T cells. Receptor phosphorylation was analyzed by immunoprecipitation (IP) with HA-Tag antibody and Immunoblotting (IB) using anti–p-tyrosine (p-Tyr) and HA-Tag antibodies. Blot demonstrates drastically reduced phosphotyrosine content in Lys650Asn mutant, and no detectable phosphotyrosine in the Phe867Leu mutant.

**(E)** Representative immunoblot demonstrating binding of Eph-B4 constructs by Ras-GAP. Ras-GAP protein was co-immunoprecipitated with Eph-B4 mutants. Ras-GAP binding was markedly reduced in Lys650Asn mutants and abrogated in Phe867Leu mutants.

## Figure 5. Damaging *CLDN14* mutations in Vein of Galen malformation



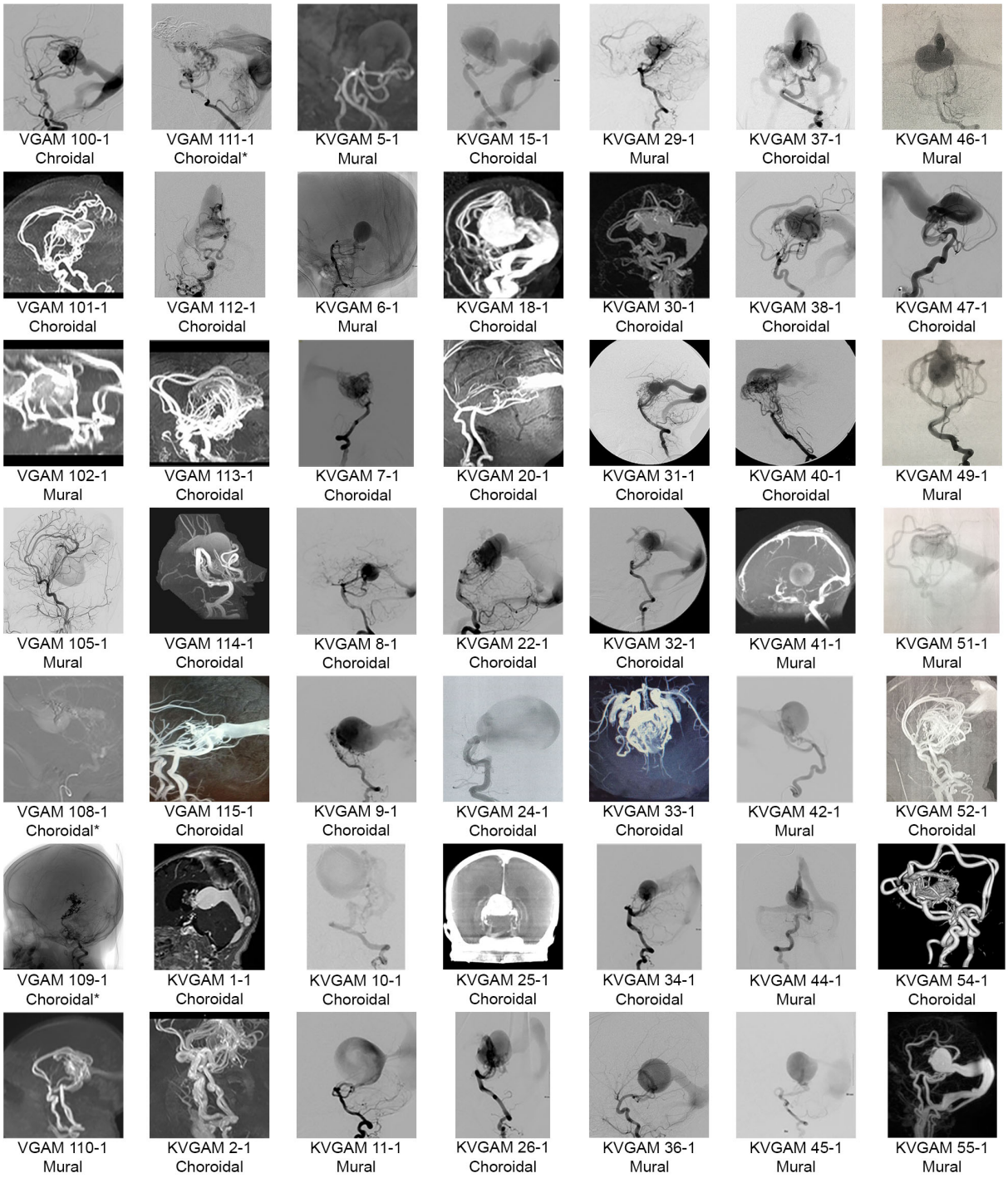

**(A)** Representative images from digital subtraction angiographies and magnetic resonance angiography demonstrating Vein of Galen malformations of the probands.

**(B)** Pedigree structures indicating genotypes and phenotypes as described in Figure 1.

**(C)** Linear representation of Claudin-14 functional domains (dark rectangles) with location of VOGM mutations (in red). Conservation of the wild-type amino acid is shown. TM = Transmembrane domain; C = cytoplasmic loop; E = extracellular loop.

# SUPPLEMENTARY FIGURES

## Supplementary figure 1. VOGM type and representative imaging for probands

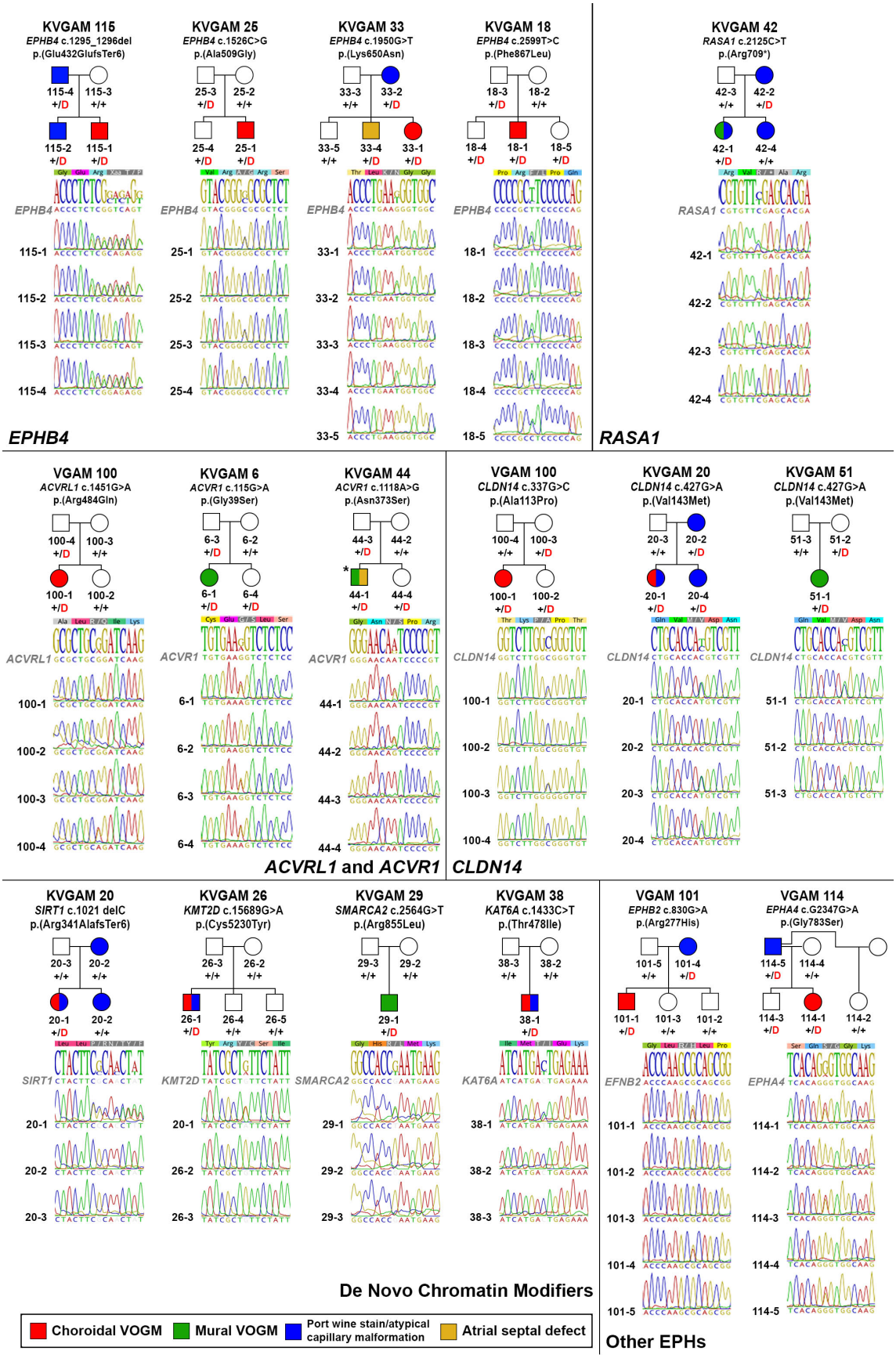



*** = Post Embolization**

**Per Detailed Angiogram Report:** VGAM 104-1: Choroidal; VGAM 107-1: Choroidal; KVGAM 3-1: Mural; KVGAM 4-1: Choroidal; KVGAM 35-1: Choroidal; KVGAM 43-1: Mural.

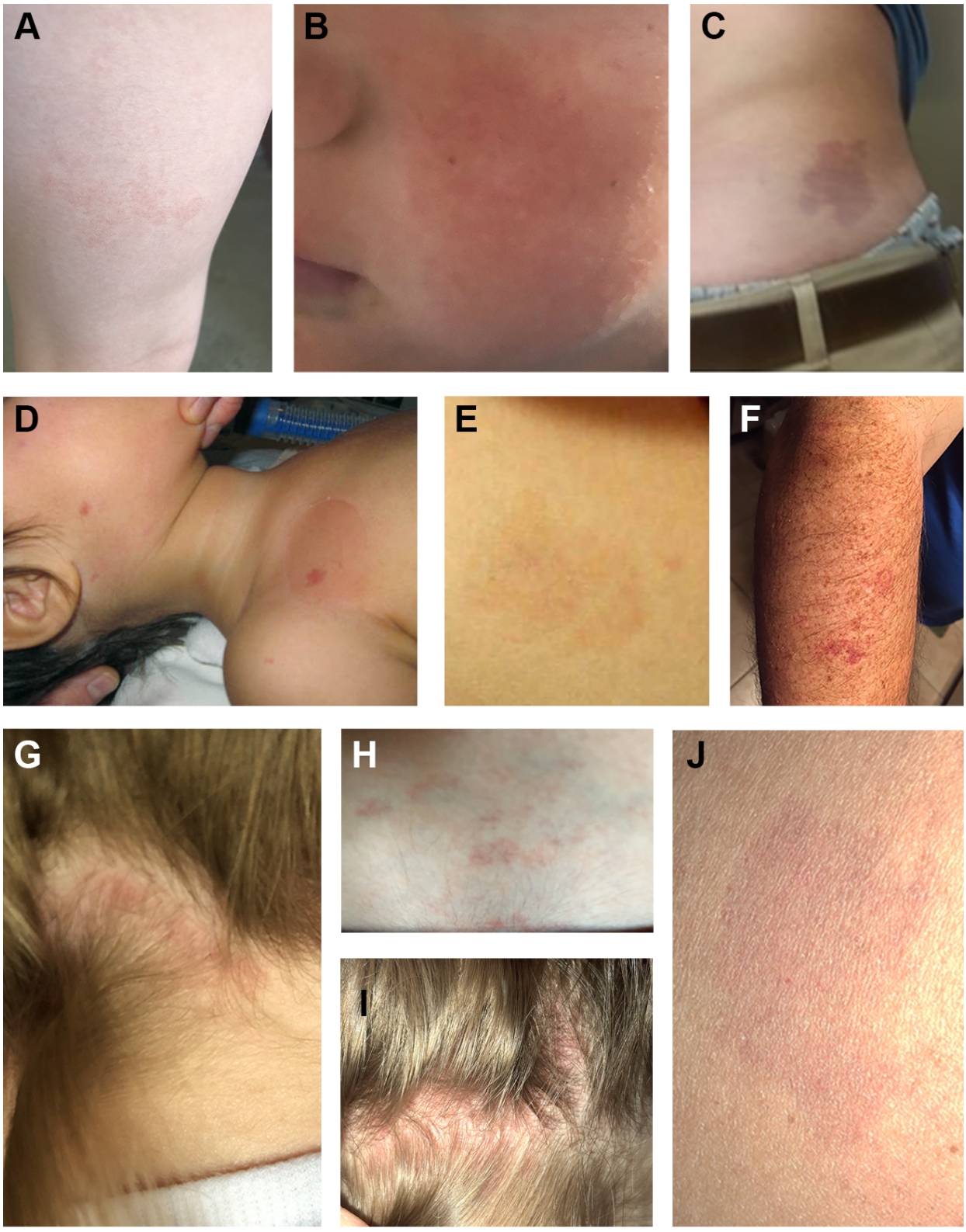
Representative images of 3-Tesla time-of-flight magnetic resonance angiography or digital subtraction angiography for all patients with available imaging, with patient codes and VOGM subtype. In cases for which imaging was not available, VOGM subtype was ascertained from detailed transcripts of angiographies during or before endovascular treatment of VOGM.

## Supplementary figure 2. Chromatogram alignments and segregation of select candidate variants



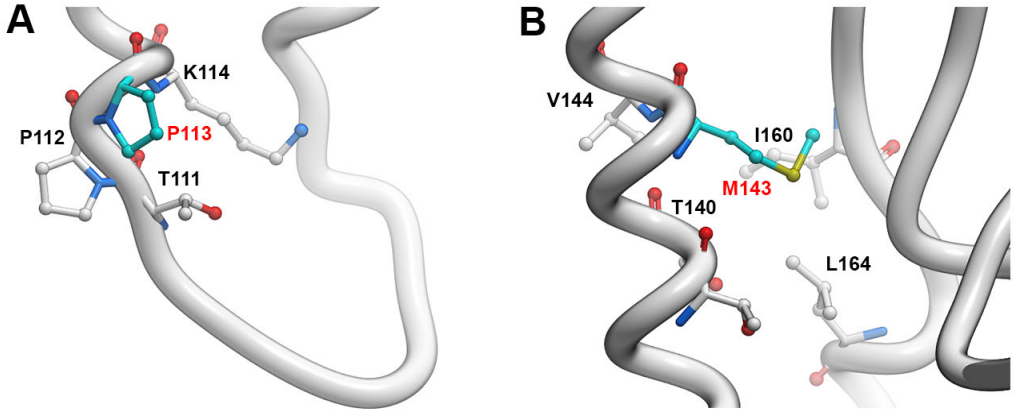

Mutations identified by exome sequencing were confirmed by direct PCR amplification with custom primers followed by Sanger sequencing.

**Supplementary figure 3. Cutaneous manifestations in VOGM probands and family members**




Atypical capillary malformation on the posterolateral aspect of the distal third of the thigh (A) and left cheek (B) of VGAM115-2. Panel C demonstrates an atypical capillary malformation on left flank of VGAM115-4 – both individuals depicted in panels A-C carry *EPHB4* c.1295_1296del; p.Glu432fs. Atypical capillary malformations on the angle of the jaw and superior right thorax of KVGAM42-1 (D) and KVGAM42-2 (E). Both individuals depicted in panels D-E carry *RASA1* c.2125C>T; p.Arg708*. (F). Atypical capillary malformation on the posterolateral aspect of the forearm in VGAM114-5, carrier of *EPHA4* c.2347G>A; p.Gly783Ser. (G). Nuccal atypical capillary malformation in a proband harboring *CLDN14* p.(Val143Met) (H). Lumbar atypical capillary malformation in the same proband depicted in panel G. (I). Nuccal atypical capillary malformation in the sister of the proband depicted in panels G and H, also harboring *CLDN14* p.Val143Met. (J). Atypical capillary malformation in the thigh of the mother of individuals depicted in panels G-I, carrying and transmitting *CLDN14* p.Val143Met.

## Supplementary figure 4. *In silico* modeling of *CLDN14* mutations



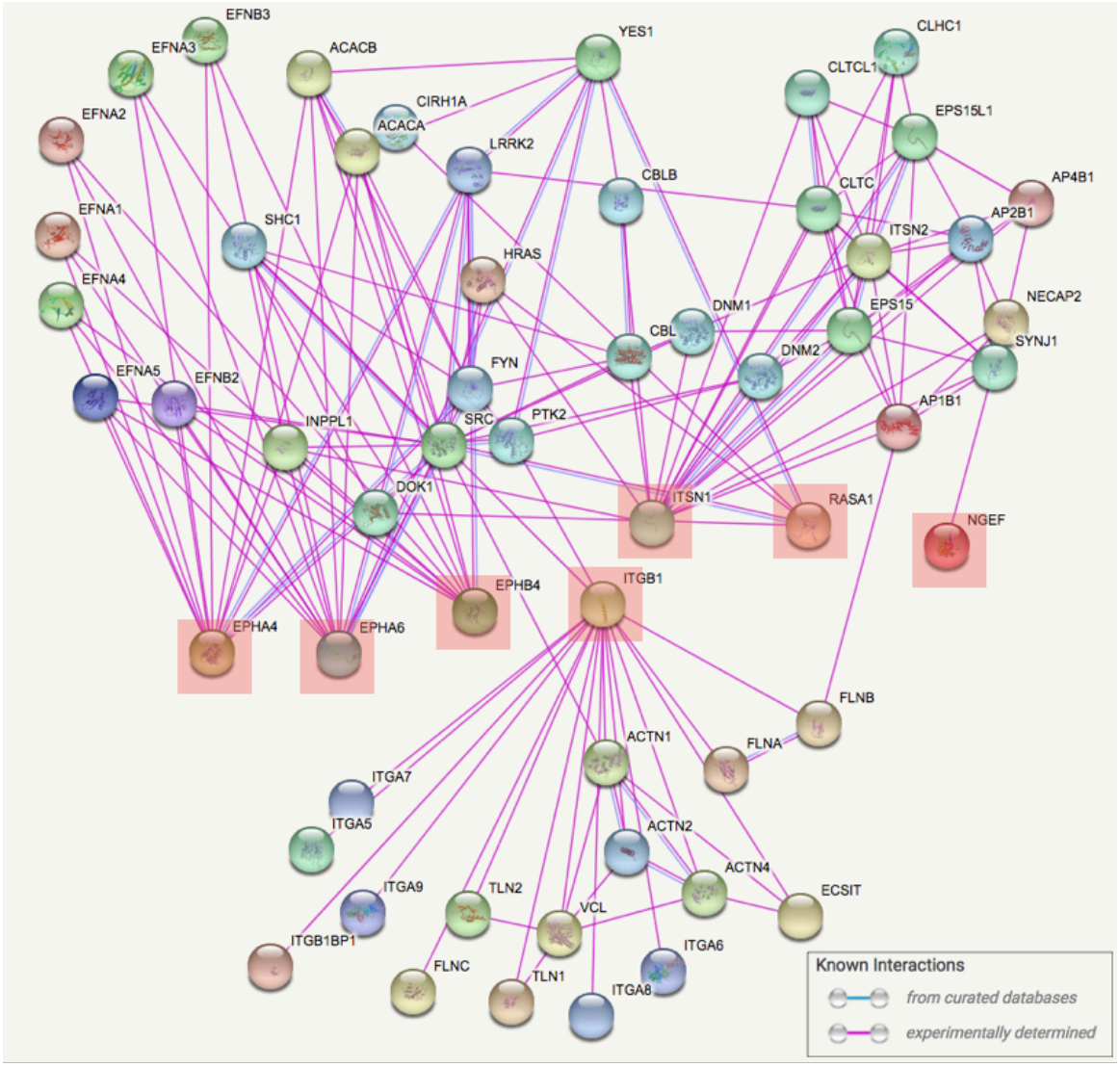

**A113P**
**ΔΔG 1.08 kcal/mol**

**V143M**
**ΔΔG 0.6 kcal/mol**

**(A)** *In silico* modeling of p.Ala113Pro. Ala113 is the first residue of Helix-3 and lies between P112 and K114 at the interface of Claudin-14 and the cytoplasmic side of the lipid bilayer. Just preceding helix-3 is the loop that connects helix-2 and -3, on which posttranslational modification sites are present including palmitoylation. The presence of Proline in an alpha helix can result in helix distortion since it lacks the backbone NH to establish hydrogen bond necessary to complete the H-bond chain. Additionally the steric or rotamer effects prevent Proline from adopting the helical geometry. Presence of two adjacent closed-cyclic side chains of Proline at 112 and 113 are presumed to structurally influence the conformation of the helix-2 and -3 connector loop and the destabilize the start of helix-3. The ΔΔG of A113P is calculated to be 1.08 Kcal/mol.

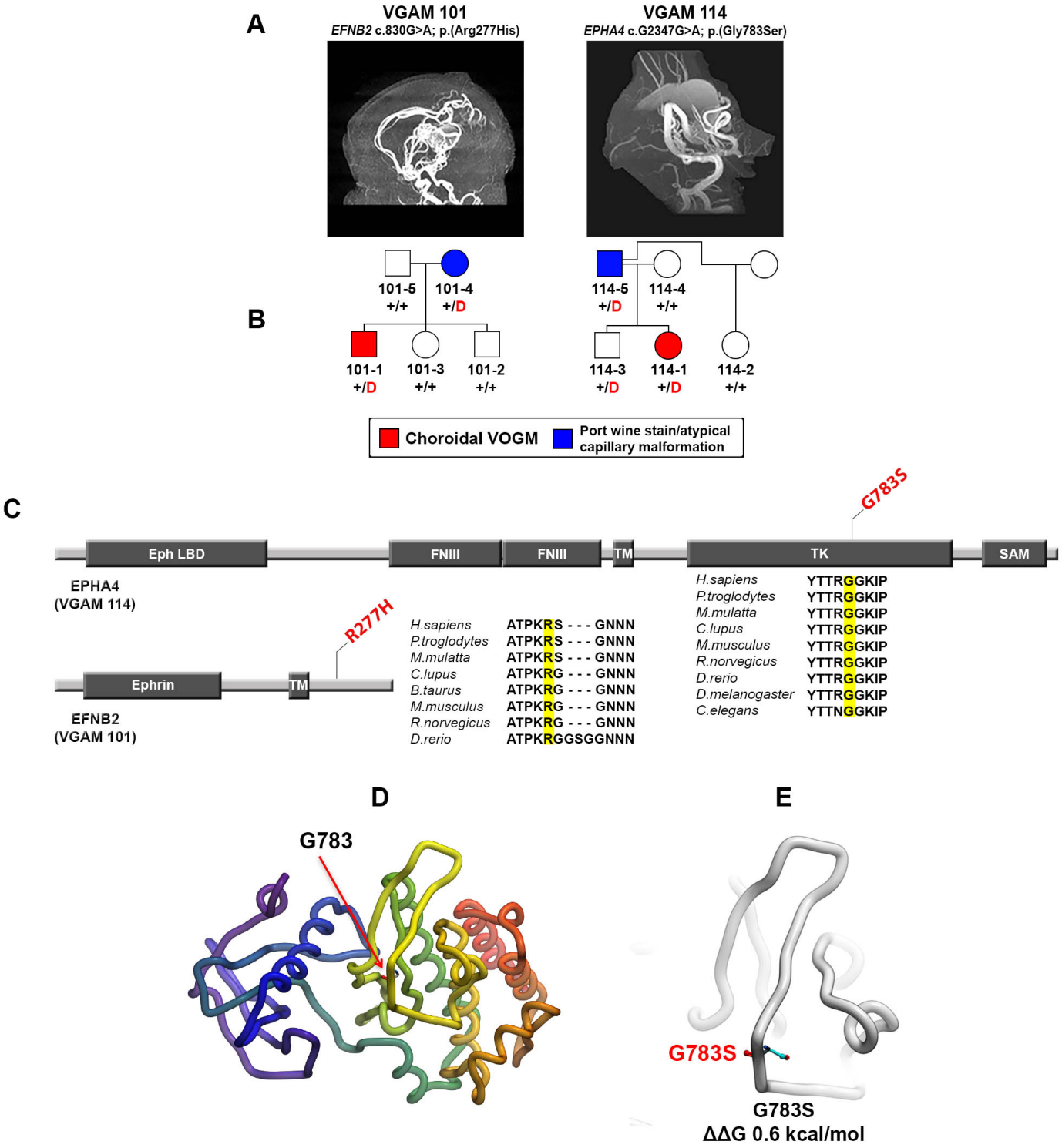
**(B)** *In silico* modeling of p.Val143Met. Val143 is located towards the end of helix-3 at the outer interface of the lipid bilayer. The side chain of Valine is tightly packed against beta sheets being surrounded by the hydrophobic side chains of L164, I160 and V144. The V143M mutation (ΔΔG = 0.6 Kcal/mol) results in the longer side chain of methionine having steric clashes with the beta sheets.

## Supplementary figure 5. Ephrin receptor signaling genes interactome




All 7 genes (highlighted) contributing to the significant result of the Ephrin receptor signaling pathway in pathway analysis using IPA was used as input in String. These 7 genes are mapped into a single experimentally supported STRING interactome.

**Supplementary figure 6. Mutations affecting other members of the Ephrin family in Vein of Galen malformation patients displaying CM-AVM-like cutaneous manifestations**




**(A)** Representative magnetic resonance angiography and digital subtraction angiography images demonstrating choroidal Vein of Galen malformations in probands.

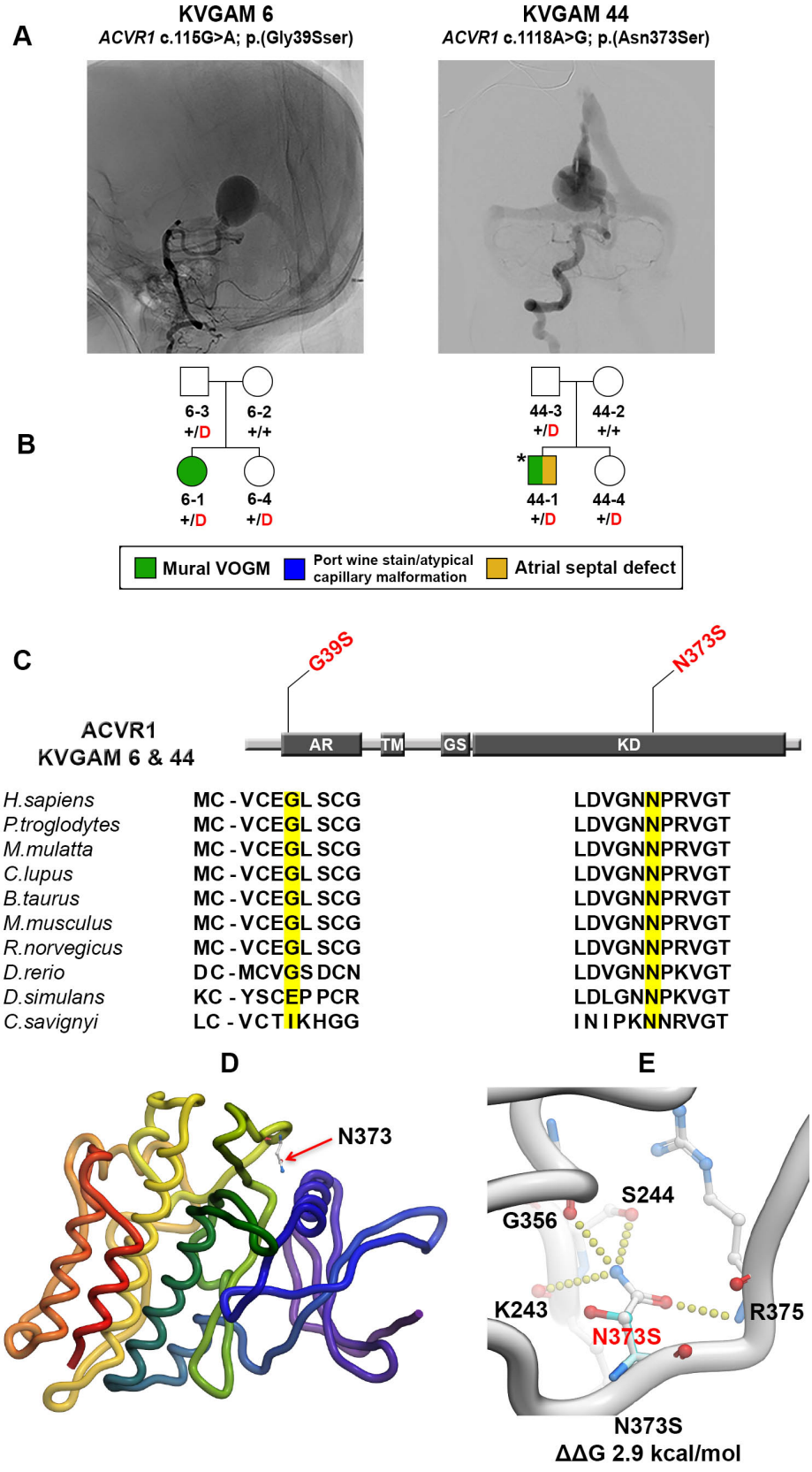
**(B)** Pedigrees depicting kindred structure and phenotypic variance. Note the presence of atypical capillary malformations represented by blue symbols in both parents transmitting *EFNB2* and *EPHB4* mutations in families VGAM101 and VGAM114.

**(C)** Linear representation of affected Ephrin molecules. Functional domains are represented by dark rectangles. Amino acid modifications are mapped on the protein structure in red lettering. Conservation of the wild-type amino acid substituted by the four deleterious missense mutations is depicted on the right of each molecule diagram. LB = ligand binding domain; CRD = Cysteine-rich domain; FNIII = Fibronectin III domain; TM = Transmembrane domain; TK = Tyrosine kinase domain; SAM = Sterile alpha motif.

**(D)** *In silico* modelling of the repercussion of Eph-A4 p.Gly783Ser. Gly783 lies at the base of a flexible loop in Eph-A4.

**(E)** The Gly-to-Ser substitution at this position (ΔΔG 0.6Kcal/mol) is predicted to affect loop stability.

## Supplementary figure 7. Damaging *ACVR1* mutations in mural Vein of Galen malformation




**(A)** Representative images from digital subtraction angiographies demonstrating mural Vein of Galen Malformations in probands.

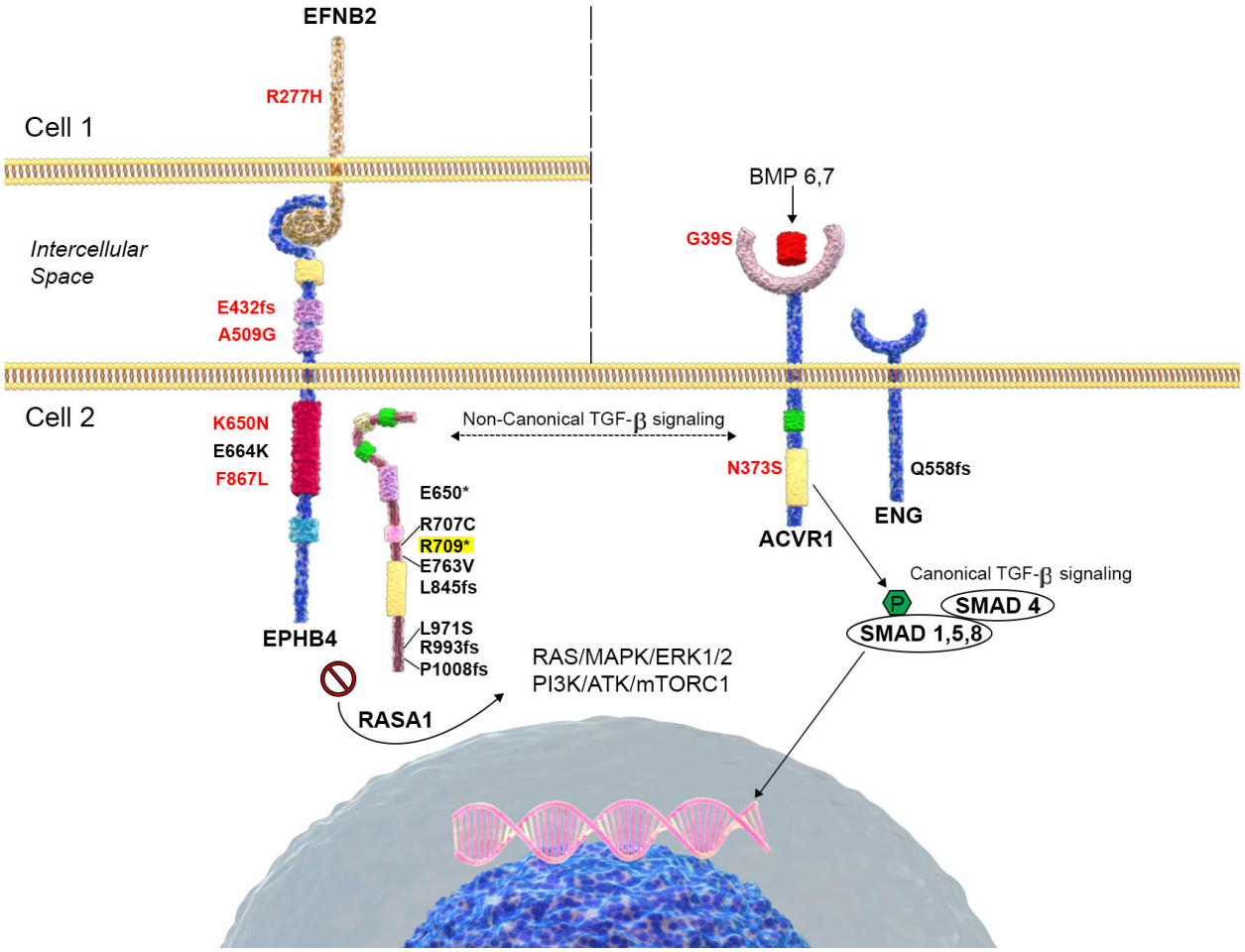
**(B)** Pedigrees depicting kindred structure and phenotypic variance. Blue symbols represent atypical capillary malformations in individuals harboring the c.1118A>G; p.Asn373Ser mutation. Yellow coloring of half the symbol denotes concomitant presence of an atrial septal defect and (asterisk) partial anomalous pulmonary venous return in the proband of family KVGAM44.

**(C)** Linear representation of the ACVR1 molecule, with functional domains represented as dark rectangles. Amino acid modifications are mapped (in red) on the protein structure. Conservation of the wild-type amino acid substituted by the two deleterious missense mutations is depicted below each. AR = activin receptor; TM = Transmembrane domain; KD = Kinase domain.

**(D)** *In silico* modelling of p.Asn373Ser. Asn373 in a 3D model of ACVR1.

**(E)** Predicted energetically costly (ΔΔG 2.9Kcal/mol) loss of hydrogen bond formation between the Asn373 side chain and backbone atoms of Gly356, Lys243, Arg375 and the side chain of Ser244 caused by Asn373 substitution to serine.

## Supplementary figure 8. The genomic landscape of VOGM




Schematic representation of signaling between two primitive endothelial cells showing contact-mediated interaction between the membrane-bound ligand Ephrin-B2 in the upper cell and its receptor, Eph-B4, in the lower cell. The figure depicts crosstalk between Eph-B4, its effector and binding partner RASA1, and activin type 1 receptors mediated by non-canonical TGF-β signaling. Mutations discovered in our cohort affecting *EPHB4*, *EFNB2*, and *ACVR1* are shown in red, and previously reported mutations in *RASA1*, *EPHB4*, and *ENG* in VOGM patients are indicated in black. *RASA1* p.Arg709* (highlighted in yellow) is the sole *RASA1* mutation detected in our cohort.

# SUPPLEMENTARY TABLES

**Supplementary Table 1. VOGM patient clinical and demographic characteristics**

| Variable | | | Type of VOGM | | |
|---|---|---|---|---|---|
| | | | *Choroidal (%)* | *Mural (%)* | Grand Total (%) |
| | | | **37 (67.27)** | **18 (32.73)** | **55 (100.00)** |
| **Gender** | | | | | |
| | Female | | 11 (29.73) | 8 (44.44) | **19 (34.55)** |
| | Male | | 26 (70.27) | 10 (55.56) | **36 (65.45)** |
| | | **Total** | 37 (100.00) | 18 (100.00) | **55 (100.00)** |
| **Age at diagnosis** | | | | | |
| | Prenatal | | 15 (40.54) | 7 (38.89) | **22 (40.00)** |
| | | Neonate (0-30 days) | 12 (32.43) | 0 (0.00) | **12 (21.82)** |
| | | Infant (1 month-2 years) | 9 (24.32) | 10 (55.56) | **19 (34.55)** |
| | Postnatal | Young child (3-6 years) | 0 (0.00) | 0 (0.00) | **0 (0.00)** |
| | | Child (7-12 years) | 1 (2.70) | 0 (0.00) | **1 (1.82)** |
| | | Adolescent (13-17 years) | 0 (0.00) | 0 (0.00) | **0 (0.00)** |
| | Unknown | | 0 (0.00) | 1 (5.56) | **1 (1.82)** |
| | | **Total** | 37 (100.00) | 18 (100.00) | **55 (100.00)** |
| **Term vs. preterm delivery** | | | | | |
| | Term (≥ 37 wks) | | 32 (86.49) | 13 (72.22) | **45 (81.82)** |
| | Preterm | | 4 (10.81) | 4 (22.22) | **8 (14.55)** |
| | Unknown | | 1 (2.70) | 1 (5.56) | **2 (3.64)** |
| | | **Total** | 37 (100.00) | 18 (100.00) | **55 (100.00)** |
| **Family history of cutaneous vascular abnormalities*** | | | | | |
| | Yes | | 23 (62.16) | 8 (44.44) | **31 (56.36)** |
| | No | | 7 (18.92) | 7 (38.89) | **14 (25.45)** |
| | Unknown | | 7 (18.92) | 3 (16.67) | **10 (18.18)** |
| | | **Total** | 37 (100.00) | 18 (100.00) | **55 (100.00)** |
| **Self-reported ethnicity** | | | | | |
| | European | | 29 (78.38) | 13 (72.22) | **42 (76.36)** |
| | | European/Asian | 3 (8.11) | 1 (5.56) | **4 (7.27)** |
| | Mixed | European/Hispanic | 1 (2.70) | 1 (5.56) | **2 (3.64)** |
| | | Latin American/African American | 1 (2.70) | 0 (0.00) | **1 (1.82)** |
| | Asian | | 1 (2.70) | 0 (0.00) | **1 (1.82)** |
| | African | | 0 (0.00) | 2 (11.11) | **2 (3.64)** |
| | Ashkenazi Jewish | | 0 (0.00) | 1 (5.56) | **1 (1.82)** |
| | Unknown | | 2 (5.41) | 0 (0.00) | **2 (3.64)** |
| | | **Total** | 37 (100.00) | 18 (100.00) | **55 (100.00)** |
| **Associated conditions** | | | | | **Total** |
| | High output cardiac failure | | 23 (62.16) | 1 (5.56) | **24 (43.64)** |
| | Progressive macrocephaly | | 22 (59.46) | 13 (72.22) | **35 (63.64)** |
| | Hydrocephalus | | 24 (64.86) | 9 (50.00) | **33 (60.00)** |
| | Prominent facial or scalp vasculature | | 18 (48.65) | 9 (50.00) | **27 (49.09)** |
| | Structural cardiac abnormalities† | | 4 (10.81) | 1 (5.56) | **5 (9.09)** |

*Includes first and second-degree relatives.

†Structural cardiac abnormalities include atrial septal defects, congenital valvular insufficiency, partial anomalous pulmonary venous return, patent foramen ovale, patent ductus arteriosus, and congenital pulmonary artery stenosis.

| Supplementary Table 2. Summary sequencing statistics for the VOGM cases and control cohorts | | | |
|---|---|---|---|
| Category | Cases* | Cases* | Controls |
| | (xGen; N=88) | (MedExome; N=71) | (Roche V2; N=5,367) |
| Read length (bp) | 99 | 99 | 50-94 |
| # of reads per sample (M) | 48.9 | 53.7 | 99.7 |
| Median coverage at each targeted base (X) | 54.4 | 37.2 | 67 |
| Mean coverage at each targeted base (X) | 59.4 | 43.8 | 79.1 |
| % of all reads that map to target | 60.70% | 48.30% | 57.60% |
| % of all bases that map to target | 46.10% | 39.10% | 49.00% |
| % of targeted bases read at least 8x | 98.40% | 96.30% | 94.60% |
| % of targeted bases read at least 10x | 98.00% | 95.00% | 93.40% |
| % of targeted bases read at least 15x | 96.30% | 89.80% | 89.90% |
| % Mean error rate | 0.30% | 0.30% | 0.40% |

*88 VOGM samples were sequenced using the xGen Exome Research Panel v1.0 capture reagent (IDT). The remaining cases were sequenced using the MedExome capture reagent. 8X, 10X and 15X were comparable across the platforms.

**Supplementary Table 3. Enrichment of *de novo* variant types in 52 VOGM and 1,789 control trios**

| | Cases, N = 52 | | | | | | Controls, N = 1,789 | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Observed | | Expected | | Enrichment | p-value | Observed | | Expected | | Enrichment | p-value |
| | N | Rate | N | Rate | | | N | Rate | N | Rate | | |
| **All Genes** | | | | | | | | | | | | |
| **Total** | 70 | 1.35 | 58.10 | 1.12 | 1.20 | 0.07 | 1830 | 1.02 | 1949.90 | 1.09 | 0.90 | 1.00 |
| **Syn** | 17 | 0.33 | 16.50 | 0.32 | 1.03 | 0.48 | 484 | 0.27 | 549.60 | 0.31 | 0.90 | 1.00 |
| **D-Mis** | 15 | 0.29 | 6.80 | 0.13 | 2.19 | $4.7x10^{-3}$ | 222 | 0.12 | 232.80 | 0.13 | 1.00 | 0.80 |
| **T-Mis** | 32 | 0.62 | 29.70 | 0.57 | 1.08 | 0.36 | 974 | 0.54 | 993.30 | 0.56 | 1.00 | 0.70 |
| **LoF** | 6 | 0.12 | 5.10 | 0.10 | 1.17 | 0.40 | 150 | 0.08 | 174.30 | 0.10 | 0.90 | 1.00 |
| **Damaging** | 21 | 0.40 | 12.00 | 0.23 | 1.76 | 0.01 | 372 | 0.21 | 407.10 | 0.23 | 0.90 | 1.00 |

Syn = Synonymous single nucleotide variant; D-Mis = Missense variants with "D" annotation per MetaSVM; T-Mis = Missense variants with "T" or "." annotation per MetaSVM; LoF = Predicted loss-of-function variants (canonical splice site, frameshift indels, and stop gains); Damaging = D-mis and/or LoF.

**Supplementary Table 4. *De novo* burden analysis comparing cases to controls**

| | # of de novo mutations | | # of de novo mutations per subject | | | |
|---|---|---|---|---|---|---|
| | Cases N=52 | Controls N=1,789 | Cases N=52 | Controls N=1,789 | Binomial P-value* | Odds ratio (95% CI)[&] |
| **LoF-intolerant genes (n = 3,230)** | | | | | | |
| Total | 16 | 504 | 0.31 | 0.28 | 0.58 | NA |
| Syn | 2 | 127 | 0.04 | 0.07 | 0.30 | NA |
| Mis | 13 | 339 | 0.25 | 0.19 | 1.00 | 2.45 (0.54 - 10.94) |
| D-Mis | 6 | 64 | 0.12 | 0.04 | 0.06 | 5.95 (1.17 - 30.33) |
| LoF | 1 | 38 | 0.02 | 0.02 | 1.00 | 1.67 (0.15 - 18.94) |
| Damaging | 7 | 102 | 0.14 | 0.06 | 0.19 | 4.36 (0.89 - 21.43) |
| **All chromatin genes (n = 547)** | | | | | | |
| Total | 4 | 65 | 0.08 | 0.04 | 0.52 | NA |
| Syn | 0 | 12 | 0.00 | 0.01 | 1.00 | NA |
| Mis | 3 | 47 | 0.06 | 0.03 | 0.60 | NA |
| D-Mis | 3 | 12 | 0.06 | 0.01 | $6.9x10^{-3}$ | NA |
| LoF | 1 | 6 | 0.02 | 0.00 | 0.62 | NA |
| Damaging | 4 | 18 | 0.08 | 0.01 | $2.1x10^{-3}$ | NA |
| **LoF-intolerant chromatin genes (n = 272)** | | | | | | |
| Total | 4 | 47 | 0.08 | 0.03 | 0.22 | NA |
| Syn | 0 | 9 | 0.00 | 0.01 | 1.00 | NA |
| Mis | 3 | 35 | 0.06 | 0.02 | 0.33 | NA |
| D-Mis | 3 | 7 | 0.06 | 0.00 | $3.0x10^{-4}$ | NA |
| LoF | 1 | 3 | 0.02 | 0.00 | 0.34 | NA |
| Damaging | 4 | 10 | 0.08 | 0.01 | $1.9x10^{-5}$ | NA |

* The p-values compare the number of de novo mutations between 52 case trios and 1,789 control trios using a two-tailed binomial exact test as shown previously (Sanders et al. 2012 Nature; Zaidi et al. 2013 Nature).

[&]The odds ratio calculates the proportion of de novo mutations in a specific category to synonymous de novo mutations and then compare these ratios in case trios versus control trios. NA, not applicable.

**Supplementary Table 5. Damaging *de novo* mutation in loss-of-function intolerant chromatin modifier genes in 55 cases**

| Salient genetic information | | | | | | | | | | | | Salient clinical and segregation information | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Proband code** | **Gene** | **Coordinate (hg19)** | **Ethnicity** | **Inheritance** | **AA Change** | **gnomAD combined MAF** | **pLI** | **Missense Z-score** | **OMIM phenotype** | **CADD** | **MetaSVM** | **Type of VOGM** | **Cutaneous manifestations or non CNS AVMs in kindred?** | **Vascular cutaneous manifestations on proband?** | **Vascular Birthmarks/ AVMs on mother** | **Vascular Birthmarks/AVMs on father** | **Concordance with mutation type** |
| KVGAM38 | *KAT6A* | 8:41832271:G:A | Mexican | *De novo* | p.Thr478Ile | novel | 1.00 | 2.14 | Mental retardation (AD) | 24.2 | D | Choroidal | y | y | n | n | y |
| KVGAM26 | *KMT2D* | 12:49420060:C:T | European | *De novo* | p.Cys5230Tyr | novel | 1.00 | 3.10 | Kabuki syndrome (AD) | 22.4 | D | Choroidal | y | y | n | n | y |
| KVGAM20* | *SIRT1* | 10:69666624:TC:T | European | *De novo* | p.Arg341fs5 | novel | 0.95 | 0.58 | None | . | N/A | Choroidal | y | y | y | n | n |
| KVGAM29 | *SMARCA2* | 9:2086866:G:T | European | *De novo* | p.Arg855Leu | novel | 1.00 | 5.57 | Nicolaides-Baraitser syndrome (AD) | 26.5 | D | Mural | n | n | n | n | N/A |

* This subject also carried a p.Val143Met mutation in CLDN14 transmitted from mother

**Supplementary Table 6a. Case-control analysis of gene burden of the damaging variants in all probands versus 3,578 Autism parental controls**

| | Case | | Control | | | |
|---|---|---|---|---|---|---|
| Gene | # mutant allele | # ref allele | # mutant allele | # ref allele | Fisher p value | Odds ratio |
| *EPHB4* | 4 | 106 | 3 | 7,153 | 1.68E-06 | 89.97 |
| *MYH7* | 3 | 107 | 0 | 7,156 | 3.38E-06 | Inf |
| *AOC2* | 3 | 107 | 0 | 7,156 | 3.38E-06 | Inf |
| *CLDN14* | 3 | 107 | 0 | 7,156 | 3.38E-06 | Inf |
| *OTOG* | 2 | 108 | 0 | 7,156 | 2.27E-04 | Inf |
| *NALCN* | 2 | 108 | 0 | 7,156 | 2.27E-04 | Inf |
| *LTBP2* | 2 | 108 | 0 | 7,156 | 2.27E-04 | Inf |
| *MLH3* | 2 | 108 | 0 | 7,156 | 2.27E-04 | Inf |
| *PKM* | 2 | 108 | 0 | 7,156 | 2.27E-04 | Inf |
| *TPSG1* | 2 | 108 | 0 | 7,156 | 2.27E-04 | Inf |

**Supplementary Table 6b. Case-control analysis of gene burden of the damaging variants in all probands versus ExAC controls**

| | Case | | Control | | | |
|---|---|---|---|---|---|---|
| Gene | # mutant allele | # ref allele | # mutant allele | # ref allele | Fisher p value | Odds ratio |
| *EPHB4* | 4 | 106 | 92 | 121,318 | 1.98E-06 | 49.76 |
| *AOC2* | 3 | 107 | 39 | 121,373 | 8.07E-06 | 87.26 |
| *CLDN14* | 3 | 107 | 50 | 120,946 | 1.65E-05 | 67.82 |
| *MPST* | 2 | 108 | 17 | 116,829 | 1.48E-04 | 127.26 |
| *DNAH12* | 2 | 108 | 37 | 121,367 | 5.89E-04 | 60.74 |
| *EMILIN3* | 2 | 108 | 40 | 121,192 | 6.85E-04 | 56.11 |
| *ABHD17A* | 2 | 108 | 42 | 118,070 | 7.91E-04 | 52.06 |
| *PRKDC* | 1 | 109 | 0 | 121,412 | 9.05E-04 | Inf |
| *SEPT1* | 1 | 109 | 0 | 121,412 | 9.05E-04 | Inf |
| *ADGRL4* | 2 | 108 | 50 | 120,618 | 1.06E-03 | 44.67 |

**Supplementary Table 7. Significant pathways in Ingenuity ® Pathway Analysis including intolerant genes harboring damaging de novo mutations, damaging rare transmitted mutations, and significantly mutated genes (*KEL*, *EPHB4*, *CLDN14*) (n = 128)**

| Rank | Ingenuity Canonical Pathways | P-value | Corrected P-value | Ratio | Genes |
|---|---|---|---|---|---|
| **1** | **Axonal Guidance Signaling** | **6.61E-08** | **1.33E-05** | **0.03** | ***SLIT3, ITGB1, SHH, EPHB4, NGEF, ITSN1, EPHA4, KEL, SLIT2, ADAMTS2, MET, EPHA6, SHANK2, RASA1, NRP1*** |
| 2 | nNOS Signaling in Skeletal Muscle Cells | 1.32E-07 | 2.64E-05 | 0.15 | *CACNA1G, CACNG2, CACNA1B, RYR2, CACNA2D3, CACNA1A* |
| 3 | GABA Receptor Signaling | 1.41E-06 | 2.81E-04 | 0.07 | *CACNA1G, CACNG2, CACNA1B, GABRB1, KCNH2, CACNA2D3, CACNA1A* |
| 4 | Netrin Signaling | 2.19E-06 | 4.34E-04 | 0.09 | *CACNA1G, CACNG2, CACNA1B, RYR2, CACNA2D3, CACNA1A* |
| 5 | FcγRIIB Signaling in B Lymphocytes | 6.92E-06 | 1.36E-03 | 0.08 | *CACNA1G, CACNG2, CD79B, CACNA1B, CACNA2D3, CACNA1A* |
| 6 | Corticotropin Releasing Hormone Signaling | 1.74E-05 | 3.41E-03 | 0.05 | *CACNA1G, CACNG2, SHH, CACNA1B, KRT1, CACNA2D3, CACNA1A* |
| 7 | G Beta Gamma Signaling | 5.62E-05 | 0.01 | 0.05 | *CACNA1G, CACNG2, CACNA1B, CACNA2D3, CACNA1A, DNM2* |
| 8 | Synaptic Long Term Depression | 7.41E-05 | 0.01 | 0.04 | *CACNA1G, CACNG2, CACNA1B, RYR2, PPP2R2B, CACNA2D3, CACNA1A* |
| **9** | **Ephrin Receptor Signaling** | **7.59E-05** | **0.01** | **0.04** | ***ITGB1, EPHA6, NGEF, EPHB4, ITSN1, EPHA4, RASA1*** |
| 10 | Androgen Signaling | 1.55E-04 | 0.03 | 0.04 | *CACNA1G, CACNG2, CACNA1B, POLR2H, CACNA2D3, CACNA1A* |
| 11 | CCR5 Signaling in Macrophages | 2.40E-04 | 0.05 | 0.05 | *CACNA1G,CACNG2,CACNA1B,CACNA2D3,CACNA1A* |

**Supplementary Table 8. Mutation burden in cases compared to ethnicity matched (European) controls in pathways significant from IPA**

| Control_Group | Pathway (# of intolerant genes) | Case_Carrier | Case_NonCarrier | CaseTotal | Control_Carrier* | Control_NonCarrier | ControlTotal | Enrichment | Confident_Interval | p-value |
|---|---|---|---|---|---|---|---|---|---|---|
| Autism | nNOS Signaling in Skeletal Muscle Cells (n = 18) | 4 | 41 | 45 | 196 | 2,690 | 2,886 | 1.34 | [0.43, Inf) | 0.37 |
| **Autism** | **Axonal Guidance Signaling (n = 161)** | **14** | **31** | **45** | **279** | **2,607** | **2,886** | **4.22** | **[2.29, Inf)** | **6.94E-05** |
| **Autism** | **Ephrin Receptor Signaling (n = 73)** | **8** | **37** | **45** | **127** | **2,759** | **2,886** | **4.69** | **[2.15, Inf)** | **8.36E-04** |
| Autism | Androgen Signaling (n = 55) | 4 | 41 | 45 | 193 | 2,693 | 2,886 | 1.36 | [0.44, Inf) | 0.36 |
| Autism | CCR5 Signaling in Macrophages (n = 31) | 3 | 42 | 45 | 138 | 2,748 | 2,886 | 1.42 | [0.37, Inf) | 0.37 |
| Autism | Corticotropin Releasing Hormone Signaling (n = 57) | 5 | 40 | 45 | 244 | 2,642 | 2,886 | 1.35 | [0.5, Inf) | 0.33 |
| Autism | FcγRIIB Signaling in B Lymphocytes (n = 36) | 3 | 42 | 45 | 149 | 2,737 | 2,886 | 1.31 | [0.34, Inf) | 0.42 |
| Autism | G Beta Gamma Signaling (n = 53) | 4 | 41 | 45 | 189 | 2,697 | 2,886 | 1.39 | [0.45, Inf) | 0.34 |
| Autism | GABA Receptor Signaling (n = 45) | 6 | 39 | 45 | 175 | 2,711 | 2,886 | 2.38 | [0.97, Inf) | 0.06 |
| Autism | Netrin Signaling (n = 32) | 4 | 41 | 45 | 212 | 2,674 | 2,886 | 1.23 | [0.4, Inf) | 0.43 |
| Autism | Synaptic Long Term Depression (n = 65) | 5 | 40 | 45 | 289 | 2,597 | 2,886 | 1.12 | [0.42, Inf) | 0.48 |
| ExAC | nNOS Signaling in Skeletal Muscle Cells (n = 18) | 4 | 41 | 45 | 3,145 | 33,532 | 36,677 | 1.04 | [0.34, Inf) | 0.55 |
| **ExAC** | **Axonal Guidance Signaling (n = 161)** | **14** | **31** | **45** | **3,878** | **32,799** | **36,677** | **3.82** | **[2.1, Inf)** | **1.52E-04** |
| **ExAC** | **Ephrin Receptor Signaling (n = 73)** | **8** | **37** | **45** | **1,621** | **35,056** | **36,677** | **4.68** | **[2.18, Inf)** | **7.33E-04** |
| ExAC | Androgen Signaling (n = 55) | 4 | 41 | 45 | 4,718 | 31,959 | 36,677 | 0.66 | [0.22, Inf) | 0.85 |
| ExAC | CCR5 Signaling in Macrophages (n = 31) | 3 | 42 | 45 | 2,138 | 34,539 | 36,677 | 1.15 | [0.3, Inf) | 0.49 |
| ExAC | Corticotropin Releasing Hormone Signaling (n = 57) | 5 | 40 | 45 | 3,466 | 33,211 | 36,677 | 1.20 | [0.45, Inf) | 0.42 |
| ExAC | FcγRIIB Signaling in B Lymphocytes (n = 36) | 3 | 42 | 45 | 2,321 | 34,356 | 36,677 | 1.06 | [0.28, Inf) | 0.55 |
| ExAC | G Beta Gamma Signaling (n = 53) | 4 | 41 | 45 | 2,726 | 33,951 | 36,677 | 1.22 | [0.4, Inf) | 0.43 |
| ExAC | GABA Receptor Signaling (n = 45) | 6 | 39 | 45 | 2,995 | 33,682 | 36,677 | 1.73 | [0.71, Inf) | 0.16 |
| ExAC | Netrin Signaling (n = 32) | 4 | 41 | 45 | 3,361 | 33,316 | 36,677 | 0.97 | [0.32, Inf) | 0.60 |
| ExAC | Synaptic Long Term Depression (n = 65) | 5 | 40 | 45 | 4,494 | 32,183 | 36,677 | 0.89 | [0.34, Inf) | 0.66 |

P values were calculated using the one-tailed Fisher's Exact Test. Values in red bold are P values exceeding the Bonferroni multiple-testing cutoff (0.05/(2x11) = 2.27E-03)

* ExAC database does not provide individual level information, number of damaging mutations in each gene set carried by each individual was assumed to be 1.

**Supplementary Table 9. Gene burden in cases compared to expectation for axonal guidance and Ephrin receptor signaling pathways**

| Axon guidance signaling | | | | | Ephrin receptor signaling | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **LoF-intolerant genes with damaging mutations in VOGM cases** | **Observed** | **Expected** | **Enrichment** | **p-value** | **LoF-intolerant genes with damaging mutations in VOGM cases** | **Observed** | **Expected** | **Enrichment** | **p-value** |
| | 18 | 7.14 | 2.52 | **3.02E-04** | | 10 | 2.90 | 3.45 | **7.28E-04** |
| **LoF-intolerant genes with synonymous mutations in VOGM cases** | **Observed** | **Expected** | **Enrichment** | **p-value** | **LoF-intolerant genes with synonymous mutations in VOGM cases** | **Observed** | **Expected** | **Enrichment** | **p-value** |
| | 21 | 20.12 | 1.04 | 0.45 | | 4 | 8.17 | 0.49 | 0.96 |
| **LoF-intolerant genes with damaging mutations in autism controls** | **Observed** | **Expected** | **Enrichment** | **p-value** | **LoF-intolerant genes with damaging mutations in autism controls** | **Observed** | **Expected** | **Enrichment** | **p-value** |
| | 372 | 348.77 | 1.07 | 0.11 | | 166 | 141.56 | 1.17 | 0.02 |
| **LoF-intolerant genes with damaging mutations in ExAC controls** | **Observed** | **Expected** | **Enrichment** | **p-value** | **LoF-intolerant genes with damaging mutations in ExAC controls** | **Observed** | **Expected** | **Enrichment** | **p-value** |
| | 3878 | 3967.08 | 0.98 | 0.93 | | 1621 | 1611.57 | 1.01 | 0.41 |

P values were calculated using the one-tailed binomial test comparing observed number of variants in LoF-intolerant genes that belong to statistically significant canonical pathways of interest to the expected number of mutations in each set (see Methods). Values in bold red are P values exceeding the Bonferroni multiple-testing cutoff (0.05/(2x3) = 0.008)

**Supplementary table 10a. Gene burden in cases compared to expectation for axonal guidance pathway after removing genes in the ephrin signaling pathway**

| LoF-intolerant genes with damaging mutations in VOGM cases | Observed | Expected | Enrichment | p-value |
|---|---|---|---|---|
| | 8 | 6.66 | 1.2 | 0.35 |

P values were calculated using the one-tailed binomial test comparing observed number of variants in LoF-intolerant genes that belong to statistically significant canonical pathways of interest to the expected number of mutations in each set (see Methods).

**Supplementary Table 10b. Mutation burden in cases compared to ethnicity matched (European) controls for axonal guidance pathway after removing genes in the ephrin signaling pathway**

| Control_Sample | Pathway | Case_Carrier | Case_NonCarrier | CaseTotal | Control_Carrier* | control_NonCarrier | ControlTotal | Enrichment | Confident_Interval | p-value |
|---|---|---|---|---|---|---|---|---|---|---|
| Autism | Axonal Guidance Signaling (n = 161) | 6 | 39 | 45 | 197 | 2,689 | 2,886 | 2.10 | [0.86,Inf) | 0.09 |
| ExAC | Axonal Guidance Signaling (n = 161) | 6 | 39 | 45 | 2,726 | 33,951 | 36,677 | 1.92 | [0.79,Inf) | 0.11 |

P values were calculated using the one-tailed Fisher's Exact Test. Bonferroni multiple-testing cutoff for this test is 0.05/2 = 0.025

* ExAC database does not provide individual level information, number of damaging mutations in each gene set carried by each individual was assumed to be 1.

**Supplementary Table 11. Mutations in genes in the Ephrin receptor signaling pathway**

| Proband code | Inheritance | Mutation Type | Gene | Coordinate (hg19) | Coding sequence variant | AA Modification | ExAC MAF | gnomAD combined MAF | pLI | Missense Z-score | OMIM phenotype | CADD | Meta SVM |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| VGAM105 | Transmitted | LoF | *ITGB1* | chr10:33196073 | N/A | Splice site | novel | novel | 0.91 | 3.47 | None | 20.9 | N/A |
| VGAM115 | Transmitted | LoF | *EPHB4* | chr7:100417179 | c.1295_1296del | p.Glu432fs1 | novel | novel | 0.99 | 2.84 | Non-immune hydrops fetails with/without atrial septal defect | . | D |
| KVGAM25 | Transmitted | D-mis | | chr7:100414876 | c.1526C>G | p.Ala509Gly | 3.30E-05 | 1.10E-05 | | | | 25 | D |
| KVGAM33 | Transmitted | D-mis | | chr7:100410537 | c.1950G>T | p.Lys650Asn | novel | novel | | | | 29.9 | D |
| KVGAM18 | Transmitted | D-mis | | chr7:100403202 | c.2599T>C | p.Phe867Leu | novel | novel | | | | 31 | D |
| KVGAM55 | Transmitted | D-mis | *NGEF* | chr2:233745889 | c.1909G>A | p.Asp637Asn | 2.83E-05 | 1.22E-05 | 0.95 | 2.18 | None | 34 | D |
| KVGAM7 | Transmitted | D-mis | *ITSN1* | chr21:35260471 | c.5033C>T | p.Thr1678Met | 1.89E-05 | 8.13E-06 | 1.00 | 3.44 | None | 34 | D |
| VGAM114 | Transmitted | D-mis | *EPHA4* | chr2:222299011 | c.G2347G>A | p.Gly783Ser | novel | novel | 1.00 | 3.15 | None | 34 | D |
| KVGAM45 | Transmitted | D-mis | *EPHA6* | chr3:96706498 | c.775C>T | p.Arg259Cys | novel | novel | 0.95 | 0.93 | None | 34 | D |
| KVGAM42 | Transmitted | LoF | *RASA1* | chr5:86672323 | c.2125C>T | p.Arg709* | novel | novel | 1.00 | 2.96 | CM-AVM1; Parkes Weber Syndrome; Somatic basal cell carcinoma | 40 | N/A |

**Supplementary Table 12. Transmitted VOGM-associated mutations in genes causing CM-AVM1, HHT, and paralog**

| Family | Type of VOGM | Ethnicity | Gene | Mutation | Domain affected | ExAC MAF | gnomAD MAF | pLI | MetaSVM | CADD |
|---|---|---|---|---|---|---|---|---|---|---|
| KVGAM6 | Mural | European | *ACVR1* | p.(G39S) | Activin receptor | novel | 3.24E-05 | 0.96 | D | 24.8 |
| KVGAM44 | Mural | European | *ACVR1* | p.(N373S) | Kinase domain | novel | novel | 0.96 | T | 24.1 |
| VGAM100 | Choroidal | Mexican | *ACVRL1* | p.(R484Q) | Kinase domain | novel | novel | 0.01 | D | 33 |

**Supplementary Table 13. Distribution of highly pathogenic mutations in VOGM patients with different subphenotypes**

| Proband Code | Gender | Ethnicity | Sub-Phenotype | Gene | Amino Acid Change | Type | Transmission | Proband with Capillary Malformation | Relatives with Capillary Malformation | Other phenotypes |
|---|---|---|---|---|---|---|---|---|---|---|
| KVGAM1 | M | European | Choroidal | | | | | n | n | Neurodevelopmental delay |
| KVGAM10 | M | European | Choroidal | *KEL* | p.Gln321* | stopgain | *De novo* | n | y | Pulmonary hypertension, neurodevelopmental delay |
| KVGAM15 | M | European | Choroidal | | | | | n | y | Patent ductus arteriosis, neurodevelopmental delay |
| KVGAM18 | M | European | Choroidal | *EPHB4* | p.Phe867Leu | Dmis | Transmitted | n | y | Cryptorchidism, strabismus, cerebral palsy |
| KVGAM2 | M | Undetermined | Choroidal | | | | | n | n | n |
| KVGAM20 | F | European | Choroidal | *CLDN14*<br>*SIRT1* | p.Val143Met<br>p.Arg341fs5 | Dmis<br>Frameshift deletion | Transmitted<br>*De novo* | y | y | n |
| KVGAM22 | F | European | Choroidal | | | | | n | n | n |
| KVGAM24 | M | European | Choroidal | | | | | n | n | Recurrent epistaxis |
| KVGAM25 | M | European | Choroidal | *EPHB4* | p.Ala509Gly | Dmis | Transmitted | n | n | Pectus excavatum |
| KVGAM26 | M | European | Choroidal | *KMT2D* | p.Cys5230Tyr | Dmis | *De novo* | n | n | Hypothyroidism, low set ears, long and everted palpebral fissures, neurodevelopmental delay, seizures |
| KVGAM30 | M | European | Choroidal | | | | | n | y | Neurodevelopmental delay, seizures |
| KVGAM31 | F | European | Choroidal | | | | | n | n | Neurodevelopmental delay |
| KVGAM32 | M | European | Choroidal | | | | | n | y | n |
| KVGAM33 | F | Mexican | Choroidal | *EPHB4* | p.Lys650Asn | Dmis | Transmitted | n | y | Neurodevelopmental delay |
| KVGAM34 | M | European | Choroidal | | | | | n | n | Neurodevelopmental delay, attention deficit hyperactivity disorder |
| KVGAM35 | M | European | Choroidal | | | | | n | n | n |
| KVGAM37 | M | European | Choroidal | | | | | n | y | Neurodevelopmental delay |
| KVGAM38 | M | Mexican | Choroidal | *KAT6A* | p.Thr478Ile | Dmis | *De novo* | y | n | Neurodevelopmental delay, seizures |
| KVGAM4 | M | Undetermined | Choroidal | | | | | n | n | n |
| KVGAM40 | M | European | Choroidal | | | | | n | y | Cerebral palsy, hip dysplasia, visual impairment, gastroesaphageal reflux disease, neurodevelopmental delay, seizures |
| KVGAM52 | F | European | Choroidal | | | | | n | n | n |
| KVGAM54 | F | European | Choroidal | | | | | n | y | n |
| KVGAM55 | F | European | Choroidal | *NGEF* | p.Asp637Asn | Dmis | Transmitted | n | n | n |
| KVGAM7 | F | European | Choroidal | *ITSN1* | p.Thr1678Met | Dmis | Transmitted | n | y | Pulmonary valve stenosis, neurodevelopmental delay |
| KVGAM8 | M | European | Choroidal | | | | | n | y | Neurodevelopmental delay, seizures |
| KVGAM9 | F | European | Choroidal | | | | | n | n | Patent ductus arteriosis, patent foramen ovale, seizures |
| VGAM100 | F | Mexican | Choroidal | *ACVRL1*<br>*CLDN14* | p.R484Q<br>p.Ala113Pro | Dmis<br>Dmis | Transmitted<br>Transmitted | n | n | Strabismus |
| VGAM101 | M | European | Choroidal | *EFNB2* | p.Arg277His | Dmis | Transmitted | y | y | Hip dysplasia, neurodevelopmental delay, seizures |
| VGAM104 | M | European | Choroidal | | | | | y | n | Cerebral palsy, neurodevelopmental delay |
| VGAM107 | F | European | Choroidal | | | | | n | n | Medulloblastoma, seizures |
| VGAM108 | M | European | Choroidal | | | | | n | y | Asthma |
| VGAM109 | M | Mexican | Choroidal | | | | | y | y | Cerebral palsy, neurodevelopmental delay, seizures |
| VGAM111 | M | European | Choroidal | | | | | n | n | Cerebral palsy, gastroesophageal reflux disease, renal agenesis, precocious puberty, mitral insufficiency, neurodevelopmental delay, seizures |
| VGAM112 | M | European | Choroidal | | | | | n | n | n |
| VGAM113 | M | European | Choroidal | | | | | n | n | Chiari Type I malformation |
| VGAM114 | F | European | Choroidal | *EPHA4* | p.Gly783Ser | Dmis | Transmitted | y | y | n |
| VGAM115 | M | European | Choroidal | *EPHB4* | p.Glu432fs1 | Frameshift deletion | Transmitted | n | y | Neurodevelopmental delay, seizures |
| VGAM47 | M | European | Choroidal | | | | | n | n | Hypospadias |
| KVGAM11 | M | European | Mural | | | | | n | y | n |
| KVGAM29 | M | European | Mural | *SMARCA2* | p.Arg855Leu | Dmis | *De novo* | n | n | Seizures |
| KVGAM3 | F | European | Mural | | | | | n | n | n |
| KVGAM36 | F | European | Mural | | | | | n | y | n |
| KVGAM41 | M | European | Mural | | | | | n | y | Major depressive disorder with psychotic features, attention deficit hyperactivity disorder, hydrocephalus, migraines |
| KVGAM42 | F | Mexican | Mural | *RASA1* | p.Arg709* | stopgain | Transmitted | y | y | Cerebral palsy, seizures |
| KVGAM43 | M | African | Mural | | | | | n | y | n |
| KVGAM44 | M | European | Mural | *ACVR1* | p.Asn373Ser | T-mis (high CADD = 24.1) | Transmitted | n | y | Atrial septal defect, partial anomalous pulmonary venous return |
| KVGAM45 | F | European | Mural | *EPHA6*<br>*KEL* | p.Arg259Cys<br>p.Gly202Ser | Dmis<br>T-mis (high CADD = 22.6) | Transmitted<br>*De novo* | n | n | Neurodevelopmental delay, seizures |
| KVGAM49 | M | Undetermined | Mural | | | | | n | n | n |
| KVGAM5 | M | Mexican | Mural | | | | | n | n | Cystic fibrosis |
| KVGAM51 | F | European | Mural | *CLDN14* | p.Val143Met | Dmis | Transmitted | n | n | n |
| KVGAM6 | F | European | Mural | *ACVR1* | p.Gly39Ser | Dmis | Transmitted | n | n | n |
| VGAM102 | M | European | Mural | | | | | n | n | Attention deficit hyperactivity disorder, recurrent epistaxis |
| VGAM105 | M | European | Mural | *ITGB1* | Splice site | splice site | Transmitted | n | n | Chiari Type I malformation |
| VGAM110 | F | European | Mural | | | | | y | n | n |
| VGAM46 | M | European | Mural | | | | | n | n | n |

Gender M = Male, F = Female; Dmis = Missense variants with 'D' annotation per MetaSVM; T-mis = Missense variants with 'T' or '.' annotation per MetaSVM; y = Yes; n = No