HHS Public Access

Author manuscript

Circ Genom Precis Med. Author manuscript; available in PMC 2019 July 01.

Published in final edited form as:

Circ Genom Precis Med. 2018 July ; 11(7): e002099. doi:10.1161/CIRCGEN.117.002099.

Rho Guanine Nucleotide Exchange Factor *ARHGEF17* Is a Risk Gene for Intracranial Aneurysms

Xinyu Yang, MD[#1], **Jiani Li, PhD**[#2], **Yabo Fang, MS**[#3,4], **Zhen Zhang, MD**[1], **Daqing Jin, PhD**[3,4], **Xingdong Chen, PhD**[3], **Yan Zhao, MD**[1], **Mengqi Li, BS**[1], **Linchun Huan, MD**[1,5], **Thomas A. Kent, MD**[6], **Jing-fei Dong, PhD**[7,8], **Rongcai Jiang, MD**[1], **Shuyuan Yang, MD**[1], **Li Jin, PhD**[3,4], **Jianning Zhang, MD**[1], **Tao P. Zhong, PhD**[3,4], and **Fuli Yu, PhD**[1,2]

[1]Department of Neurosurgery, Tianjin Neurological Institute, Tianjin Medical University General Hospital, Tianjin, China

[2]Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX

[3]State Key Laboratory of Genetic Engineering, School of Life Sciences, Zhong Shan Hospital, Fudan University

[4]Shanghai Key Laboratory of Regulatory Biology, Institute of Molecular Medicine, East China Normal University School of Life Sciences, Shanghai

[5]Department of Neurosurgery, Linyi People's Hospital, Shandong, China

[6]Engineering Medicine, Texas A&M Health Science Center and College of Engineering, Houston, TX

[7]Blood Works Northwest Research Institute

[8]Division of Hematology, Department of Medicine, University of Washington School of Medicine, Seattle, WA

[#] These authors contributed equally to this work.

Abstract

Background—Intracranial Aneurysm (IA) is usually a late-onset disease, affecting 1-3% of the general population and leading to life-threatening subarachnoid hemorrhage (SAH). Genetic susceptibility has been implicated in IAs but the causative genes remain elusive.

Methods—We performed next-generation sequencing (NGS) in a discovery cohort of 20 Chinese IA patients. Bioinformatics filters were exploited to search for candidate deleterious variants with rare and low allele frequency. We further examined the candidate variants in a multi-ethnic sample

Correspondence: Fuli Yu, PhD, Baylor College of Medicine, One Baylor Plaza, N1619, Houston, TX 77030-3498, Tel: 1-713-798-7676, Fax: 1-713-798-5741, fyu@bcm.edu; Tao P. Zhong, PhD, Life Sciences Building, 500 Dongchuan Road, Shanghai, China 200241, Tel: 86-21-54345021, Fax: 86-21-51630547, taozhong@fudan.edu.cn; Jianning Zhang, Tianjin Medical University, 154 Anshan Rd, Heping Qu, China, 300051, Tel: 86-22-60814400, Fax: 86-22-60814400, jianningzhang@hotmail.com.

Disclosures: Dr. Thomas A. Kent is a member and shareholder in Acelerox, LLC to commercialize the medical use of carbon nanomaterial.

collection of 86 whole-exome sequenced unsolved familial IA cases from three previously published studies.

Results—We identified that the low-frequency variant c.4394C>A_p.Ala1465Asp (rs2298808) of *ARHGEF17* was significantly associated with IA in our Chinese discovery cohort (P $=7.3\times10^{-4}$; OR=7.34). It was subsequently replicated in Japanese familial IA patients (P=0.039; OR=4.00; 95% CI = 0.832-14.8) and was associated with IA in the large Chinese sample collection comprising 832 sporadic IA-affected and 599 control individuals (P=0.041; OR =1.51; 95% CI = 1.02-Inf). When combining the sequencing data of all familial IA patients from four different ethnicities (i.e. Chinese, Japanese, European American, and French-Canadian), we identified a significantly increased mutation burden for *ARHGEF17* (21/106 versus 11/306; P = 8.1×10^{-7}; OR=6.6; 95% CI = 2.9-15.8) in cases as compared to controls. In zebrafish, *arhgef17* was highly expressed in the brain blood vessel. *arhgef17* knockdown caused blood extravasation in the brain region. Endothelial lesions were identified exclusively on cerebral blood vessels in the *arhgef17*-deficient zebrafish.

Conclusions—Our results provide compelling evidence that *ARHGEF17* is a risk gene for IA.

Keywords

Cerebral Aneurysm; Genetics; Basic Science Research; Computational Biology; Functional Genomics; intracranial aneurysm; genomics; genetics; animal models; ARHGEF17; human genetics; zebrafish model

Introduction

Intracranial Aneurysms (IAs) are weakened areas in the cerebral vascular that lead to abnormal dilation or ballooning[1]. Unruptured IAs are usually asymptomatic unless they reach sufficient size to influence surrounding structures. They are often found either through screening high-risk patients or as purely incidental findings during brain imaging for other neurological symptoms. The danger posed by IAs is rupture, leading to life-threatening subarachnoid hemorrhage (SAH)[2]. About 1%−3% of the general population has or will develop IAs[3,4]. And 0.5% to 1% of people with IAs may suffer from rupture annually[5]. Meta-analysis showed a wide range of SAH fatality varied from 8.3% to 66.7% that decreased 0.4% per year (95% CI = 0.5 to 1.2) coinciding with the introduction of improved management strategies of pre-symptomatic and symptomatic IAs[6].

The formation and rupture of IAs are likely related to multiple risk factors including smoking, hypertension, heavy alcohol consumption[7] and other vascular disease[8]. A small percentage of which are related to infections[9] and trauma[10]. Many studies have shown that there are underlying genetic risk factors for a large fraction of IA and SAH cases. It is known that 7% to 20% of patients have a positive family history[11]. First-degree relatives of individuals with aneurysmal SAH have a 4- to 7-fold higher risk of being affected than the general population[12].

The genetic studies on IAs have been carried out through genome-wide association studies (GWASs) and family studies. To date, several GWASs have examined sporadic IA cases and identified several candidate loci: 4q31.23[13], 2q33.1[14,15], 8q11.23, and 9p21.3[14–16]; 18q11.2,

13q13.1, 9p21.3, and 10q24.23[16]. Unfortunately, follow-up replication studies showed that these loci are unlikely to explain a large fraction of IAs[17]. In family studies, linkage analysis of single IA families or sibling pairs have mapped several candidate loci, but no IA-causing mutations has been identified. Recently, four studies[18–21] applied whole-exome sequencing (WES) on multiple IA-affected families to analyze the effect of the rare variant on IAs, three IA candidate genes (*ADAMTS15*, *RNF213*, and *THSD1*) were identified. However, many IA cases involved in those four studies cannot be explained by the mutations of these genes[18–21]. Therefore, new genetic causality for IA needs to be elucidated to unravel the cellular and molecular mechanisms underlying IAs pathophysiology.

One of the challenges common to GWASs and familial studies on IAs is the number of patients or families in each study. Firstly, unruptured IAs usually do not manifest any clinical phenotype and the ruptured IAs have high mortality. Accumulating a large sample size requires a long-term epidemiological screening program involving multiple hospitals or registries. Secondly, as a late-onset disease, some family members may have deceased or not manifested IA at the time, which poses challenges to ascertain a complete pedigree for segregation analyses. Finally, the incomplete penetrance needs to be taken into considerations in the analysis of IAs.

Because much of the genetic contribution to complex traits cannot be explained by GWASs that focus on the identification of common variants, we focused our attention on the identification of rare (MAF < 0.5%) and low-frequency (0.5% ⩽ MAF < 5%) variants and applied next-generation sequencing (NGS), which enables more complete assessments of low-frequency and rare genetic variants. It is plausible that analyses of rare and low-frequency variants could explain additional disease risk or trait variability[22] (Figure 1A). There is also recent empirical evidence that low-frequency and rare variants are associated with complex diseases[23–25].

It is possible that few genetic alleles with modest effect size and incomplete penetrance can segregate in familial fashion as well as increase the IA risk in the general population (Figure 1B). For example, previous studies on familial late-onset Alzheimer disease showed that a significant excess of rare coding variants in *APP*, *PSEN1*, and *PSEN2* was noted in probands. These variants did not actually co-segregate with the disease and some of those rare variants were also identified in sporadic Alzheimer cases[26,27], suggesting that those rare variants may serve to modulate the risk of disease both in a familial format and in a sporadic format. With this possibility in mind, to increase the potential genetic loads of IAs in our study samples, lower the burden of sample collection, and identify these risk alleles underlying both familial and sporadic IAs, we purposefully sampled both familial and sporadic IA cases and compared the alleles to the population controls (Figure 1B). This subject ascertaining strategy was also employed in a recent IA association study by Kurki *et al.*, in which they sampled 760 cases from both familial and sporadic IAs (40% of familial IAs) from Finnish population[28]. They successfully replicated the 9p21.3 and 13q13.1 IA association signals and identified four novel loci: 2q23.3, 5q31.1, 6q24.2, and 7p22.1[28].

We leveraged recent foundational developments in human genetics, such as the latest *in silico* predictors, large genetic variant databases (1000 Genomes Project (1000G)[29] and

Exome Aggregation Consortium (ExAC)[30]), and Genotype-Tissue Expression (GTEx) information[31] to prioritize the IA risk gene. We also incorporated unsolved IA cases (no known genetic risk factors identified) from previous studies[18–21] as the replication datasets to replicate the candidate variants and to assess the variant burden of the candidate genes (Figure 1C).

Methods

The sequencing data that support the findings of this study are available from the corresponding author upon reasonable request. Codes used in this paper were uploaded to github repository and links were included in the Supplementary Materials.

Ethics approval and consent to participate

The manuscript has ethics approval and consent to participate by the medical ethics committee of Tianjin Medical University General Hospital (20170035). Informed consent has been obtained and this report was processed according to the principles expressed in the Declaration of Helsinki.

Further details of the study participants, data analyses, and functional studies in zebrafish can be found in the Supplementary Materials.

Results

Genome sequencing identifies 30 candidate variants

Based on our conceptual design described in the Introduction and Figure 1, we developed a human-genetics analysis pipeline that includes three major stages: variant discovery, variant replication, and risk gene prioritization (Figure 2). In the discovery stage, we genome-sequenced IA patients and performed variant filtering to identify a list of candidate variants. In the variant replication stage, we performed a replication test on the candidate variants using the unsolved IA cases from the previous studies[18,19,21] to examine if any of the candidate variants can explain additional IA cases. Finally, we integrated all the sequencing data of IA samples and tested the variant enrichment in the genes to prioritize the risk gene of IAs (Figure 2). To further use the genetic information from familial cases, we performed the segregation analysis and Identity-by Descent (IBD) analysis to examine the association of candidate variants with IAs in families. We also examined the low-frequency risk variant in additional sporadic IA cases (Figure 2).

In our discovery stage, we ascertained 20 Chinese IA samples (Figure S1A, B) and interrogated their genetic variants using whole-exome sequencing or whole-genome sequencing technologies. We focused on the genetic variants in the coding regions by subsetting the WGS to the exome regions that captured by WES (that is, Agilent SureSelect Human All Exon 50MB). Using our ensemble variant calling pipeline (Methods in Supplementary Materials and Figure S), we identified 115,912 SNVs and 6,033 Indels in total. On average, we found 42,115 SNVs and 3,267 Indels per exome and 97.51% of those SNVs and 83.92% of indels were in dbSNP141. The variant set had a mean transition-to-transversion (Ti/Tv) ratio of 2.56 for all SNVs per exome (Table S6).

We first kept variants in our cases with allele frequency ≤ 5% in the East Asian population of both the 1000G and ExAC databases (Figure 2). Then we performed a case-control analysis to further ensure the putative risk allele was significantly enriched in the cases and was potentially associated with IA. We aggregated the 20 familial and sporadic cases and chose 208 Han Chinese subjects in the 1000G as controls after verifying the congruence of population structure by principal component analysis (PCA) (Figure S4). We applied the two-sided Fisher's exact test on each variant in cases and controls and kept variants (6,685 variants left) that were significantly enriched in cases with P-value < 8×10^{-3} after multiple hypotheses correction using Benjamini–Hochberg procedure at FDR = 5% (Figure S5 and Methods).

Then we kept the variants annotated as loss-of-function (LoF) variants. For missense variants, we adapted widely used bioinformatics tools including GERP++[32], CADD[33], SIFT[34], PolyPhen-2[35], LRT[36], and MutationTaster[37] to predict their functional effects[38] (Figure S5 and Methods in Supplementary Materials). We found 117 variants passed these bioinformatics filters. Finally, we required that at least two IA-affected families share a candidate variant, and 30 variants remained as candidates after this filter (Figure S5). Three of them are LoF variants that occurred at essential splicing sites and 27 of them are damaging missense variants predicted by multiple *in silico* algorithms (Methods and Table S7).

Replicate candidate variants in additional familial cases

We next examined the 30 candidate variants in additional unsolved IA cases from previously published studies[18,19,21]. There are 86 unsolved familial IAs (24 families) with WES data available. We named the Japanese cohort as "Replication 1", the European American cohort as "Replication 2", and the French-Canadian cohort as "Replication 3" (Methods and Table S1).

For Replication 1, we included 37 IA cases from 11 families (Figure S2A; Methods). The control samples for Replication 1 were from two databases that were used in the original publication by Yan *et al.*[19]: 104 1000G Japanese samples and 1208 Japanese samples from JGVCD[39] (Table S1; Methods). For Replication 2, we included exome sequencing data of 34 European Americans IA cases from seven families[18] (Figure S2B). The control samples were 99 Northern Europeans from Utah (CEU) subjects from 1000G. The ethnicities of cases and controls were examined by PCA in Figure S6. For Replication 3, we included 15 French-Canadians IA cases (Figure S2C). We also applied the 99 CEU subjects from 1000G as controls for Replication 3 since the ethnicities of the patients were confirmed in the original publication by Zhou *et al.*[21]

We applied our variant calling and prioritization methods to these replication datasets (Methods). Among the 30 candidate variants identified in the initial Chinese patients, six of them were observed in replication cohorts as well: five (rs115753757 of *ADAM15*, rs2298808 of *ARHGEF17*, rs35217482 of *AOX1*, rs114777682 of *AKAP13*, and rs150645471 of *ACSM5*) were found in Replication 1, and one (rs2397084 of *IL17F*) were found both in the Replication 2 and Replication 3 (Figure 3A; Table S8). After comparing the allele count of the variants to the population controls using two-sided Fisher's exact test,

we ranked these six variants based on their p-values by an ascending order and plotted it against the p-values of 30 variants in the discovery cohort (Figure 3A). Two out of the six variants (rs115753757 of *ADAM15* and rs2298808 of *ARHGEF17*) were top ranking with p-value < 0.05 (Fisher's exact test; Figure. 3A; Table S8), suggesting a suggestive association with IAs.

Gene-level variant enrichment analysis identified a significantly higher variant burden of *ARHGEF17* in IA cases than in controls

To assess the variant burden of the six genes with variants rediscovered in replication, particularly *ADAM15* and *ARHGEF17* with nominal significance in additional replications, we performed a 'gene-level variant enrichment test', a simplified gene-based association test: if a gene is associated with IA, we expect to identify more distinct rare deleterious variants at gene-level in cases than in controls, which can help us evaluate association for multiple variants in a gene[22] (Figure 2; Figure S7).

We combined all four datasets of IA patients (106 in total) and considered the entire gene as a testing unit. We used 103 Han Chinese in Beijing (CHB) subjects, 104 JPT subjects, and 99 CEU subjects from 1000G as control individuals (in total 306 control samples). We applied the same variant prioritization methods that we developed (Figure S8). For each gene, we categorized the damaging variants into three categories: exclusively identified in cases, exclusively identified in controls, and identified in both cases and controls (Figure 3B). We compared the number of distinct damaging variants that exclusively observed in cases to those exclusively observed in controls, and we found a higher number of distinct damaging variants of *ARHGEF17* in cases than in controls. Four distinct *ARHGEF17* rare deleterious variants were exclusively observed in IA-affected individuals, while only one was exclusively observed in the 1000G control individuals (Figure. 3B, Table 1). For the other five genes, we did not observe higher number of distinct damaging variants in cases than in controls (Figure. 3B; Table S9).

We next examined the frequency differences of these *ARHGEF17* deleterious variants in cases and controls. We found that 21 of the 106 cases (21%) and 11 of the 306 1000G controls (3.6%) have the *ARHGEF17* deleterious variants ($P = 8.1 \times 10^{-7}$ by two-sided Fisher's exact test; OR=6.6; 95% CI =2.9 −15.8; Table 1). When we consider the family as a unit by pruning the related subjects, nine out of 36 families (25%) carried the deleterious mutations of *ARHGEF17* ($P=4.3\times10^{-5}$ by two-sided Fisher's exact test; OR=8.8; 95% CI = 3.5-Inf; Table 1 and Table S10).

We also examined the expression profile of the six genes across different human tissues using the Genotype-Tissue Expression (GTEx) data[31]. We found that *ARHGEF17* is the only gene among the six that is highly expressed in blood vessels (Figure. 3C), which provided additional evidence that *ARHGEF17* could be a candidate gene of IAs and plays a role in human blood vessels. Based on the ExAC database, *ARHGEF17* is extremely intolerant to LoF variants (probability of being LoF intolerant (pLI) = 1.00), and ARHGEF17 is also intolerant to missense variants with Z-score equals 3.18.

Pedigree and IBD analyses on *ARHGEF17* variants

We went back to the pedigrees for segregation analysis. We found that three out of the five *ARHGEF17* variants (c.4394C>A_p.Ala1465Asp, c.5168G>A_p.Arg1723Gln and c.5327G>A_p.Cys1776Tyr) were co-segregated with IAs in at least one of the pedigrees (Figure. 3D and Table 1). The variant c.3511G>A_p.Ala1171Thr was a novel variant found in a Chinese sporadic IA patient. The variant c.5581C>T_p.Arg1861Cys was found in two patients from a French-Canadian family (Figure. 3D and Table.1). We performed IBD analysis on the families with *ARHGEF17* mutations. Among the four families (Discovery P1, Discovery P4, Replication1 P6, and Replication2: FamilyD) with *ARHGEF17* mutations co-segregated with IA, three (Discovery P1, Replication1 P6, and Replication2: FamilyD) had shared IBD haplotype that cover *ARHGEF17* gene region (Table S11).

Overall, in our Chinese cases, we found two *ARHGEF17* mutations (c.4394C>A_p.Ala1465Asp and c.3511G>A_p.Ala1171Thr). We screened the two mutations in the IA-unaffected Chinese family members that were available to our study. The variant c.3511G>A_p.Ala1171Thr was only found in IA-affected family members. We found that four unaffected individuals had the mutation c.4394C>A_p.Ala1465Asp and three of them are younger than the IA-affected family members (Figure 3D and Figure S1A). While not manifesting IA, the clinical history records of those four individuals showed that three of them had cerebrovascular abnormalities, including cerebral infarctions, cerebral artery stenosis, and bilateral fetal posterior cerebral artery. Two of them also had hypertension (Table S12).

Association analysis in sporadic IAs provided additional evidence supporting that *ARHGEF17* is a risk gene of IA

Since c.4394C>A_p.Ala1465Asp (rs2298808) is a low-frequency variant in East Asians and is carried by IA-unaffected individuals as well, we further examined if it was enriched in sporadic IAs cases. We independently screened this variant in additional 832 unrelated Chinese SAH patients caused by IA-rupture and 599 unrelated Chinese individuals with no evidence of IAs or other malformations (Table S2).

We found the alternative allele of rs2298808 showed nominal enrichment in IA-affected individuals compared to control individuals (AF_{Cases} = 3.73%; $AF_{Controls}$ = 2.50%; P=0.041 by one-sided Fisher's exact test, OR=1.51; 95% CI = 1.02-Inf; Table S12). In addition, the homozygous rs2298808 were only observed in five patients with ruptured IAs (Table S13). They provided additional lines of evidence that *ARHGEF17* is a risk gene of IAs.

Functional studies of *ARHGEF17* function and its mutations

We performed additional functional analyses on *ARHGEF17* using *in silico* methods. We first performed a literature search to examine if *ARHGEF17* is biologically related to IAs. In addition, we also performed computational modeling of *ARHGEF17* domains using structural homologies to examine effects of mutations on the domain structures.

In our literature search, we found that *ARHGEF17* encodes a Dbl family Rho Guanine Nucleotide Exchange Factor (GEF), which plays an important role in activating the Rho family GTPases[40–42]. The proteins of Rho subfamily (RhoA, B, and C) promote the formation of actin stress fibers that are essential for cellular processes, including cell shape, polarity, migration, cell-cell and cell-matrix interactions[40,43,44]. We also found that the depletion of *ARHGEF17* leads to defective human umbilical vein endothelial cell junctions[45], suggesting that it could play a role in maintaining the integrity of blood vessels.

The predicted domains of ARHGEF17 include the catalytic Dbl homology (DH) domain for Rho GTPase activation, pleckstrin homology (PH) domain for lipid interaction, and the β-propeller fold that may mediate protein-protein interaction[46,47]. The variants c.3511G>A_p.Ala1171Thr and c.4394C>A_p.Ala1465Asp map to the DH domain and the PH domain respectively. Two variants (c.5327G>A_p.Cys1776Tyr and c.5581C>T_p.Arg1861Cys) map to the β-propeller domain (Figure 4A). All five mutations affect evolutionarily conserved residues (Figure 4B). The variant c.5168G>A_p.Arg1723Gln is the first amino acid of the predicted β-propeller domain that connects to the linker sequence between two domains (Figure S11A).

Computational modeling of *ARHGEF17* important domains using protein homology showed that four variants (c.4394C>A_p.Ala1465Asp, c.5168G>A_p.Arg1723Gln, c.5327G>A_p.Cys1776Tyr, and c.C5581T_p.Arg1861Cys) are exposed on the surface of the protein domain, suggesting that these variants might alter the enzyme in its interactions with substrate or cofactors and/or in its structural stability. Thr at position 1171 can introduce a bulky side chain compared to the wild-type Ala and might result in clashes in the domain structure (Figure. S11A, B).

arhgef17-deficient zebrafishes display intracranial hemorrhage

We found that zebrafish Arhgef17 is 67% identical to the human ARHGEF17 with a high degree of conservation in the DH domain, PH domain and Beta-propeller[48] (Figure S12). Whole-mount *in situ* hybridization displayed abundant expression of *arhgef17* in the zebrafish head region, including the forebrain, the midbrain, and the hindbrain during development (Figure 5A-C). Using ARHGEF17 antibody, we detected that Arhgef17 is expressed in the brain blood vessels at 72 hpf (Figure 5D, 5DI, 5DI-I, and 5DI-II).

To validate the functional impact of ARHGEF17 *in vivo*, we developed a zebrafish model with *arhgef17* deficiency (Methods). We designed two splice-blocking morpholinos, I1E2-MO and I9E10-MO to knockdown *arhgef17* by skipping exon 2 (DH domain) and exon 10 (PH domain), respectively (Figure. S13A). RT-PCR analyses of embryos injected with I1E2-MO and I9E10-MO determined the knockdown efficiency (Figure S13B-D). *arhgef17* transcripts skipping exon 2 or partial exon 2 were detected in embryos with microinjection of I1E2-MO (Figure S12B,E), confirmed by DNA sequencing (Figure S13F). Microinjection of I9E10-MO caused a frameshift, resulting in nonsense-mediated mRNA decay (NMD) of *arhgef17*, as the new electrophoretic band was absent, and the PCR product of unedited sequence was reduced (Figure S13C, D).

arhgef17-deficient embryos appeared to be morphologically indistinguishable from their wild-type (WT) siblings up to 42 hours-post-fertilization (hpf) (Figure. 5E,F). From 42 hpf to 80 hpf, embryos injected with I1E2-MO or I9E10-MO displayed intracranial hemorrhage in multiple areas in the zebrafish head region, including the forebrain, the midbrain, and the hindbrain, in comparison to control embryos (Figure 5E,F**;** Figure S14A,B). Whole-mount o-Dianisidine analyses for hemoglobin staining[49] confirmed erythrocyte extravasation in *arhgef17*-deficient embryos (Figure 5G; Figure S14C,D), suggesting the presence of defects in vessel integrity or development. To identify endothelial lesions that cause intracranial hemorrhage, we employed double transgenic zebrafish Tg (*kdrl:EGFP; gata1:DsRed*), in which the endothelium and the erythroid are marked by enhanced green fluorescent protein (EGFP) and *Discosoma* sp. red fluorescent protein (DsRed), respectively (Figure 5H). We observed erythrocyte extravasation marked by DsRed in the midbrain and hindbrain region (Figure 5I; Figure 6A,B), as well as the forebrain (Figure. 5J and6B) in *arhgef17*-deficient embryos. These intracranial bleeding sites are adjacent to the prosencephalic artery (PrA) (Figure. 5J) and the hindbrain vascular network, comprising the basilar artery (BA), primordial hindbrain channels (PHBCs) and central arteries (CtAs)[50] (Figure 5H**)**. However, overall blood vessel morphology and development appeared to be normal (Figure 5G, H, I; 5A, B**)**. Furthermore, confocal imaging analyses identified multiple endothelial lesions, localizing specifically to a CtA (Figure 6AI, AI-I), ACeV (Figure. 6BI, BI-I), mesencephalic vein (MsV) (Figure. 5BII, BII-I) and posterior cerebral vein (PCeV) (Figure 6BIII, BIII-I). When we co-injected wild-type human *ARHGEF17* mRNA into zebrafish zygotes with *arhgef17* morpholinos, we observed a statistically significant reduction of intracranial hemorrhages (Figure 6C). However, co-injection with rs2298808 variant human mRNA can only partially reduce hemorrhage phenotypes in *arhgef17*-deficiency embryos (Figure 6C). These results reveal that endothelial leakage in multiple cranial vessels causes erythrocyte extravasation in *arhgef17* deficient embryos and *ARHGEF17* is a risk gene of IAs.

Discussion

IA is usually a late-onset complex disease with both genetic and environmental risk factors[17]. Similar to other common complex diseases such as Alzheimer's disease, breast cancer, and diabetes, it is also genetically heterogeneous, so a large proportion of the patients does not follow a canonical Mendelian inheritance pattern[26,51,52]. Many affected families may include both inherited and sporadic components of the disease. Meanwhile, the sporadic cases can include inherited components as well, likely due to the incomplete penetrance of genetic risk factors.

In this study, we pooled familial and sporadic IA cases together and performed a case-control analysis in candidate variants filtering. We think this method may help (1) increase the genetic loads in cases; (2) lose the burden of collecting large families with multiple family affected; (3) identify the risk factors underlying the general population.

Aggregation of the unsolved IA cases from previous studies provided us the opportunity to study the genetic risk factors in multi-ethnic familial IA samples. We found *ARHGEF17* had a significantly high mutation burden in the IA cases compared to control individuals (21/106

vs. 11/306; $P=7.4\times10^{-12}$), suggesting that *ARHGEF17* mutations may be IA risk factors in both Asian and European populations.

We found four IA unaffected individuals in familial samples had the *ARHGEF17* damaging mutation (rs2298808, MAF = 3.24% in EAS from ExAC, Table S12), which is likely due to the incomplete penetrance of this variant. We also found that the four heterozygous carriers of rs2298808 from two families are IA negative and relatively younger. One family had two heterozygous IA patients and the mean diagnose age is 50, while the mean age of heterozygous carriers is 43 (Figure S1A; family P4). The other family contained one 58-year old heterozygous IA patient, while the mean age of the two heterozygous carriers is 53.5 (Figure S1A; family P6). Therefore, at this point we cannot entirely exclude the possibility that some of unaffected heterozygotes will develop IA when they grow older. We further examined rs2298808 in sporadic Chinese IA cases. We found that the alternative allele frequency of rs2298808 is 3.73% in 832 unrelated cases, while 2.50% in 599 controls. This difference achieved nominal statistical significance (P=0.041 one-sided Fisher's exact test; OR=1.51, Table S12), suggesting that it is a genetic risk factor of IAs. More interestingly, we only observed the homozygous rs2298808 in IA-affected individuals but not in the control individuals, and all of them have ruptured IAs (Table S13). Our rescue experiments using rs2298808 variant mRNA and found that rs2298808 mRNA can only partially rescue the hemorrhage phenotype in *Arhgef17*-deficiency embryos, suggesting that s2298808 mutation cannot disrupt the *ARHGEF17* function completely in zebrafish. Nevertheless, this result does not contrast our conclusion that *ARHGEF17* variant (s2298808) is associated with human IA with incomplete penetrance.

ARHGEF17 encodes one of the Dbl family guanine-nucleotide-exchange factors (GEFs) and many Dbl family GEFs have been shown as oncogenes. Similarly, *ARHGEF17* was identified originally as a gene whose expression was upregulated in endothelial cells during tumor cell-induced angiogenesis[53]. This is for the first time that *ARHGEF17* is implicated in a cerebrovascular disease. Recent studies on *ARHGEF17* showed that (1) ARHGEF17 is localized to the actin cytoskeleton via sequences in the N-terminus and activates RhoA at the intercellular junctions[46]; (2) the depletion of *ARHGEF17* leads to defective human umbilical vein endothelial cell junctions[45]. Importantly, it has been shown that the specific activation of RhoA at intercellular junctions is critical for proper endothelial junction formation and integrity[54]. Other known risk loci in IAs are also implicated in the endothelial cell function and integrity. The recent study identified *THSD1*, which plays an important role in the endothelial cell focal adhesion and attachment[20]; the other two IA candidate genes *ADAMTS15* and *RNF213* have antiangiogenic activity and might also be involved in endothelial cell-cell junction stabilization[19,21]. The loss of endothelial-specific expression of Sox17 promoted IAs in an elastase-treated mouse model[55]; the destruction of tight junctions may facilitate macrophage migration and cerebral aneurysm formation in rats[56]. Therefore, it is plausible that *ARHGEF17* mutations disrupt the RhoA activation that consequently leads to the endothelial junction destabilization in blood vessels, conferring risks for an IA formation and rupture.

We performed functional studies in the zebrafish model and showed that *arhgef17* is abundantly expressed in the head region of zebrafish, including brain blood vessels. We have

followed the technical guidelines for morpholino use in zebrafish[57,58], and showed injection of two splicing-blocking MOs at low dosages (~3 ng) causes intracranial hemorrhage without overall embryonic abnormalities. The morpholinos-based knockdown leads to the perturbation of DH domain or results in nonsense-mediated mRNA decay. Importantly, the intracranial hemorrhage could be rescued by co-injection of human wild-type Arhgef17 mRNA. Thus, the prominent phenotypes of intracranial hemorrhage are caused by *arhgef17* deficiency rather than morpholino toxicities and/or off-target effects. Human mRNA with rs2298808 variant can partially rescue the hemorrhage phenotype in Arhgef17-deficient embryos, suggesting that rs2298808 variant cannot completely disrupt the *ARHGEF17* function in zebrafish and could be a risk factor for IAs.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments:

We thank Sau Wai Cheung, Graeme Mardon, Chad Shaw and Christian Schaaf for constructive suggestions, and Yufeng Shen, Richard A. Gibbs, and Lee-Jun Wong for critical reading of the manuscript. We are grateful to all of the patients and their families for participating in this study.

Sources of Funding: This study was supported by grants from the US National Institutes of Health and National Human Genome Research Institute (R01HG008115 to FY), the National Natural Science Foundation of China (31530044 and 31471357 to TPZ; 81571144 to XY and JZ; 81330029 to JZ), the Tianjin Science and Technology Commission (16JCZDJC35700 to XY and JZ, 15ZXLCSY00060 to JZ), the US National Institutes of Health-National Heart, Lung, and Blood Institute (5R01HL125957 to FY and JFD), and the National Institute of Neurological Disorders and Stroke (R01NS094535 to TK).

References:

1. Vlak MH, et al. Prevalence of unruptured intracranial aneurysms, with emphasis on sex, age, comorbidity, country, and time period: a systematic review and meta-analysis. Lancet Neurol. 2011;10:626–636. doi:10.1016/S1474-4422(11)70109-0. [PubMed: 21641282]
2. van Gijn J, et al. Subarachnoid haemorrhage. Lancet. 2007;369:306–318. doi:10.1016/S0140-6736(07)60153-6. [PubMed: 17258671]
3. Rinkel GJE. Natural history, epidemiology and screening of unruptured intracranial aneurysms. J Neuroradiol. 2008;35:99–103. doi:10.1016/j.neurad.2007.11.004. [PubMed: 18242707]
4. Vernooij MW, et al. Incidental Findings on Brain MRI in the General Population. N Engl J Med. 2007;357:1821–1828. doi:10.1056/NEJMoa070972. [PubMed: 17978290]
5. Rinkel GJ, et al. Prevalence and risk of rupture of intracranial aneurysms: a systematic review. Stroke. 1998;29:251–256. http://www.ncbi.nlm.nih.gov/pubmed/9445359 Accessed April 23, 2017. [PubMed: 9445359]
6. Nieuwkamp DJ, et al. Changes in case fatality of aneurysmal subarachnoid haemorrhage over time, according to age, sex, and region: a meta-analysis. Lancet Neurol. 2009;8:635–642. doi:10.1016/S1474-4422(09)70126-7. [PubMed: 19501022]
7. Teunissen LL, et al. Risk factors for subarachnoid hemorrhage: a systematic review. Stroke. 1996;27:544–549. [PubMed: 8610327]
8. Brown RD, et al. Unruptured intracranial aneurysms and arteriovenous malformations: frequency of intracranial hemorrhage and relationship of lesions. J Neurosurg. 1990;73:859–863. doi:10.3171/jns.1990.73.6.0859. [PubMed: 2230969]
9. Clare CE, et al. Infectious intracranial aneurysms. Neurosurg Clin N Am. 1992;3:551–566. [PubMed: 1633479]

10. Benoit BG, et al. Traumatic cerebral aneurysms. Clinical features and natural history. J Neurol Neurosurg Psychiatry. 1973;36:127–138. http://www.ncbi.nlm.nih.gov/pubmed/4691685 Accessed April 22, 2017. [PubMed: 4691685]
11. Caranci F, et al. Epidemiology and genetics of intracranial aneurysms. Eur J Radiol. 2013;82:1598–1605. doi:10.1016/j.ejrad.2012.12.026. [PubMed: 23399038]
12. Krischek B, et al. The genetics of intracranial aneurysms. J Hum Genet. 2006;51:587–594. doi: 10.1007/s10038-006-0407-4. [PubMed: 16736093]
13. Low S- K, et al. Genome-wide association study for intracranial aneurysm in the Japanese population identifies three candidate susceptible loci and a functional genetic variant at EDNRA. Hum Mol Genet. 2012;21:2102–2110. doi:10.1093/hmg/dds020. [PubMed: 22286173]
14. Foroud T, et al. Genome-wide association study of intracranial aneurysms confirms role of Anril and SOX17 in disease risk. Stroke. 2012;43:2846–2852. doi:10.1161/STROKEAHA.112.656397. [PubMed: 22961961]
15. Bilguvar K, et al. Susceptibility loci for intracranial aneurysm in European and Japanese populations. Nat Genet. 2008;40:1472–1477. doi:10.1038/ng.240. [PubMed: 18997786]
16. Yasuno K, et al. Genome-wide association study of intracranial aneurysm identifies three new risk loci. Nat Genet. 2010;42:420–425. doi:10.1038/ng.563. [PubMed: 20364137]
17. Tromp G, et al. Molecular basis and genetic predisposition to intracranial aneurysm. Ann Med. 2014;46:597–606. doi:10.3109/07853890.2014.949299. [PubMed: 25117779]
18. Farlow JL, et al. Lessons learned from whole exome sequencing in multiplex families affected by a complex genetic disorder, intracranial aneurysm. PLoS One. 2015;10:e0121104. doi:10.1371/journal.pone.0121104. [PubMed: 25803036]
19. Yan J, et al. Genetic study of intracranial aneurysms. Stroke. 2015;46:620–626. doi:10.1161/STROKEAHA.114.007286. [PubMed: 25649796]
20. Santiago-Sim T, et al. THSD1 (Thrombospondin Type 1 Domain Containing Protein 1) Mutation in the Pathogenesis of Intracranial Aneurysm and Subarachnoid Hemorrhage. Stroke. 2016;47.
21. Zhou S, et al. RNF213 Is Associated with Intracranial Aneurysms in the French-Canadian Population. Am J Hum Genet. 2016;99:1072–1085. doi:10.1016/j.ajhg.2016.09.001. [PubMed: 27745834]
22. Lee S, et al. Rare-variant association analysis: study designs and statistical tests. Am J Hum Genet. 2014;95:5–23. doi:10.1016/j.ajhg.2014.06.009. [PubMed: 24995866]
23. Rivas MA, et al. Deep resequencing of GWAS loci identifies independent rare variants associated with inflammatory bowel disease. Nat Genet. 2011;43:1066–1073. doi:10.1038/ng.952. [PubMed: 21983784]
24. Jonsson T, et al. A mutation in APP protects against Alzheimer's disease and age-related cognitive decline. Nature. 2012;488:96–99. doi:10.1038/nature11283. [PubMed: 22801501]
25. Gudmundsson J, et al. A study based on whole-genome sequencing yields a rare variant at 8q24 associated with prostate cancer. Nat Genet. 2012;44:1326–1329. doi:10.1038/ng.2437. [PubMed: 23104005]
26. Cruchaga C, et al. Rare Variants in APP, PSEN1 and PSEN2 Increase Risk for AD in Late-Onset Alzheimer's Disease Families Toft M, ed. PLoS One. 2012;7:e31039. doi:10.1371/journal.pone.0031039. [PubMed: 22312439]
27. Cooper DN, et al. Where genotype is not predictive of phenotype: towards an understanding of the molecular basis of reduced penetrance in human inherited disease. Hum Genet. 2013;132:1077–1130. doi:10.1007/s00439-013-1331-2. [PubMed: 23820649]
28. Kurki MI, et al. High risk population isolate reveals low frequency variants predisposing to intracranial aneurysms. PLoS Genet. 2014;10:e1004134. doi:10.1371/journal.pgen.1004134. [PubMed: 24497844]
29. Auton A, et al. A global reference for human genetic variation. Nature. 2015;526:68–74. doi: 10.1038/nature15393. [PubMed: 26432245]
30. Lek M, et al. Analysis of protein-coding genetic variation in 60,706 humans. Nature. 2016;536:285–291. doi:10.1038/nature19057. [PubMed: 27535533]
31. Melé M, et al. The human transcriptome across tissues and individuals. Science (80-). 2015;348.

32. Davydov EV, et al. Identifying a high fraction of the human genome to be under selective constraint using GERP++. PLoS Comput Biol. 2010;6. doi:10.1371/journal.pcbi.1001025.
33. Kircher M, et al. A general framework for estimating the relative pathogenicity of human genetic variants. Nat Genet. 2014;46:310–315. doi:10.1038/ng.2892. [PubMed: 24487276]
34. Ng PC, et al. SIFT: Predicting amino acid changes that affect protein function. Nucleic Acids Res. 2003;31:3812–3814. [PubMed: 12824425]
35. Adzhubei IA, et al. A method and server for predicting damaging missense mutations. Nat Methods. 2010;7:248–249. doi:10.1038/nmeth0410-248. [PubMed: 20354512]
36. Chun S, et al. Identification of deleterious mutations within three human genomes. Genome Res. 2009;19:1553–1561. doi:10.1101/gr.092619.109. [PubMed: 19602639]
37. Schwarz JM, et al. MutationTaster2: mutation prediction for the deep-sequencing age. Nat Methods. 2014;11:361–362. doi:10.1038/nmeth.2890. [PubMed: 24681721]
38. Richards S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med. 2015;17:405–424. doi:10.1038/gim.2015.30. [PubMed: 25741868]
39. Yamaguchi-Kabata Y, et al. iJGVD: an integrative Japanese genome variation database based on whole-genome sequencing. Hum Genome Var. 2015;2:15050. doi:10.1038/hgv.2015.50. [PubMed: 27081555]
40. Hall A, et al. Rho GTPases: molecular switches that control the organization and dynamics of the actin cytoskeleton. Philos Trans R Soc London B Biol Sci. 2000;355.
41. Aktories K, et al. Clostridial Rho-inhibiting protein toxins. Curr Top Microbiol Immunol. 2005;291:113–145. http://www.ncbi.nlm.nih.gov/pubmed/15981462 Accessed November 5, 2016. [PubMed: 15981462]
42. Ridley AJ. Rho GTPases and actin dynamics in membrane protrusions and vesicle trafficking. Trends Cell Biol. 2006;16:522–529. doi:10.1016/j.tcb.2006.08.006. [PubMed: 16949823]
43. Raftopoulou M, Hall A. Cell migration: Rho GTPases lead the way. Dev Biol. 2004;265:23–32. doi:10.1016/j.ydbio.2003.06.003. [PubMed: 14697350]
44. Ridley AJ, et al. Cell Migration: Integrating Signals from Front to Back. Science (80-). 2003;302.
45. Ngok SP, et al. TEM4 is a junctional Rho GEF required for cell-cell adhesion, monolayer integrity and barrier function. J Cell Sci. 2013;126:3271–3277. doi:10.1242/jcs.123869. [PubMed: 23729734]
46. Mitin N, et al. Identification of a novel actin-binding domain within the Rho guanine nucleotide exchange factor TEM4. PLoS One. 2012;7:e41876. doi:10.1371/journal.pone.0041876. [PubMed: 22911862]
47. Rümenapp U, et al. A mammalian Rho-specific guanine-nucleotide exchange factor (p164-RhoGEF) without a pleckstrin homology domain. Biochem J. 2002;366:721–728. doi:10.1042/BJ20020654. [PubMed: 12071859]
48. Lutz S, et al. RhoGEF17, a Rho-specific guanine nucleotide exchange factor activated by phosphorylation via cyclic GMP-dependent kinase Iα. Cell Signal. 2013;25:630–638. doi:10.1016/j.cellsig.2012.11.016. [PubMed: 23195829]
49. Paik EJ, et al. Hematopoietic development in the zebrafish. Int J Dev Biol. 2010;54:1127–1137. doi:10.1387/ijdb.093042ep. [PubMed: 20711990]
50. Isogai S, et al. The Vascular Anatomy of the Developing Zebrafish: An Atlas of Embryonic and Early Larval Development. Dev Biol. 2001;230:278–301. doi:10.1006/dbio.2000.9995. [PubMed: 11161578]
51. Steinthorsdottir V, et al. Identification of low-frequency and rare sequence variants associated with elevated or reduced risk of type 2 diabetes. Nat Genet. 2014;46:294–298. doi:10.1038/ng.2882. [PubMed: 24464100]
52. King M-C. “The Race” to Clone BRCA1. Science (80-). 2014;343:1462–1465. doi:10.1126/science.1251900.
53. Cook DR, et al. Rho guanine nucleotide exchange factors: regulators of Rho GTPase activity in development and disease. Oncogene. 2014;33:4021–4035. doi:10.1038/onc.2013.362. [PubMed: 24037532]

54. Terry SJ, et al. Spatially restricted activation of RhoA signalling at epithelial junctions by p114RhoGEF drives junction formation and morphogenesis. Nat Cell Biol. 2011;13:159–166. doi: 10.1038/ncb2156. [PubMed: 21258369]

55. Lee S, et al. Deficiency of Endothelium-Specific Transcription Factor Sox17 Induces Intracranial Aneurysm. Circulation. 2015;131:995–1005. doi:10.1161/CIRCULATIONAHA.114.012568. [PubMed: 25596186]

56. Tada Y, et al. Reduction of endothelial tight junction proteins is related to cerebral aneurysm formation in rats. J Hypertens. 2010;28:1883–1891. doi:10.1097/HJH;0b013e32833c2273. [PubMed: 20577123]

57. Rasouli SJ, et al. Regulation of cardiomyocyte behavior in zebrafish trabeculation by Neuregulin 2a signaling. Nat Commun. 2017;8:15281. doi:10.1038/ncomms15281. [PubMed: 28485381]

58. Moulton JD. Using Morpholinos to Control Gene Expression In: Current Protocols in Nucleic Acid Chemistry. Vol 68 Hoboken, NJ, USA: John Wiley & Sons, Inc.; 2017:4.30.1–4.30.29. doi: 10.1002/cpnc.21.

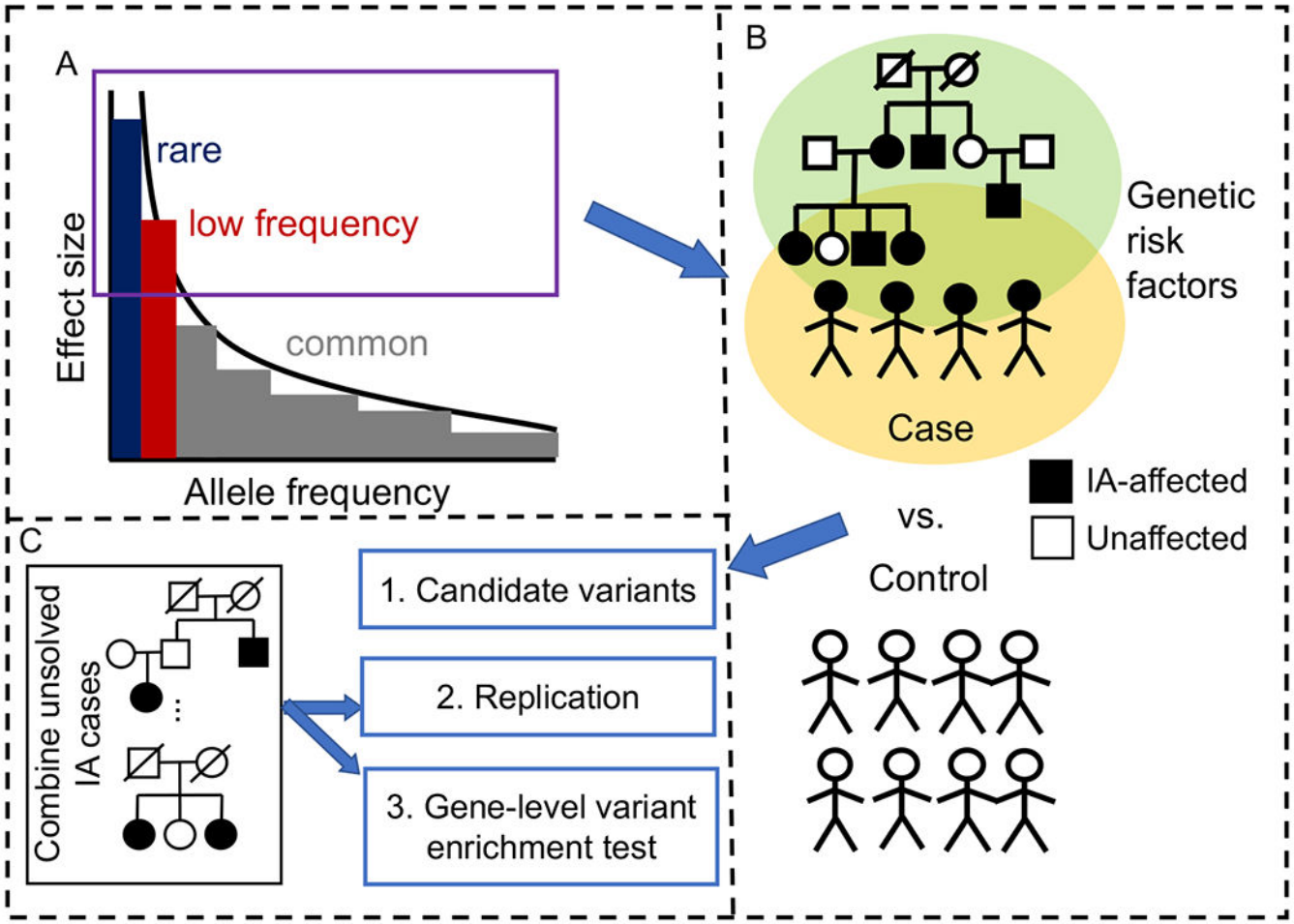


Figure 1. The rationale and experimental design of our study.

(A) For a late-onset complex disease, a suite of common variants will have modest effects on IAs and contribute in an interactive manner to confer susceptibility. Meanwhile, some rare or low-frequency variants will have larger effects on IAs and contribute individually and/or in combination to the risk of IAs. **(B)** Those rare or low-frequency alleles with larger effect sizes and/or low-penetrance can contribute to a late on-set disease like IA familial form and increase IA risk in general population. To increase the potential genetic load of IAs, we can aggregate the familial IAs and sporadic IAs to perform the variant association test to find the candidate variants. **(C)** We can integrate the sequencing data of unsolved IA cases from previous studies to replicate the candidate variants followed by collapsing the variants to gene level to perform gene-level variant enrichment test to prioritize the risk gene of IAs.

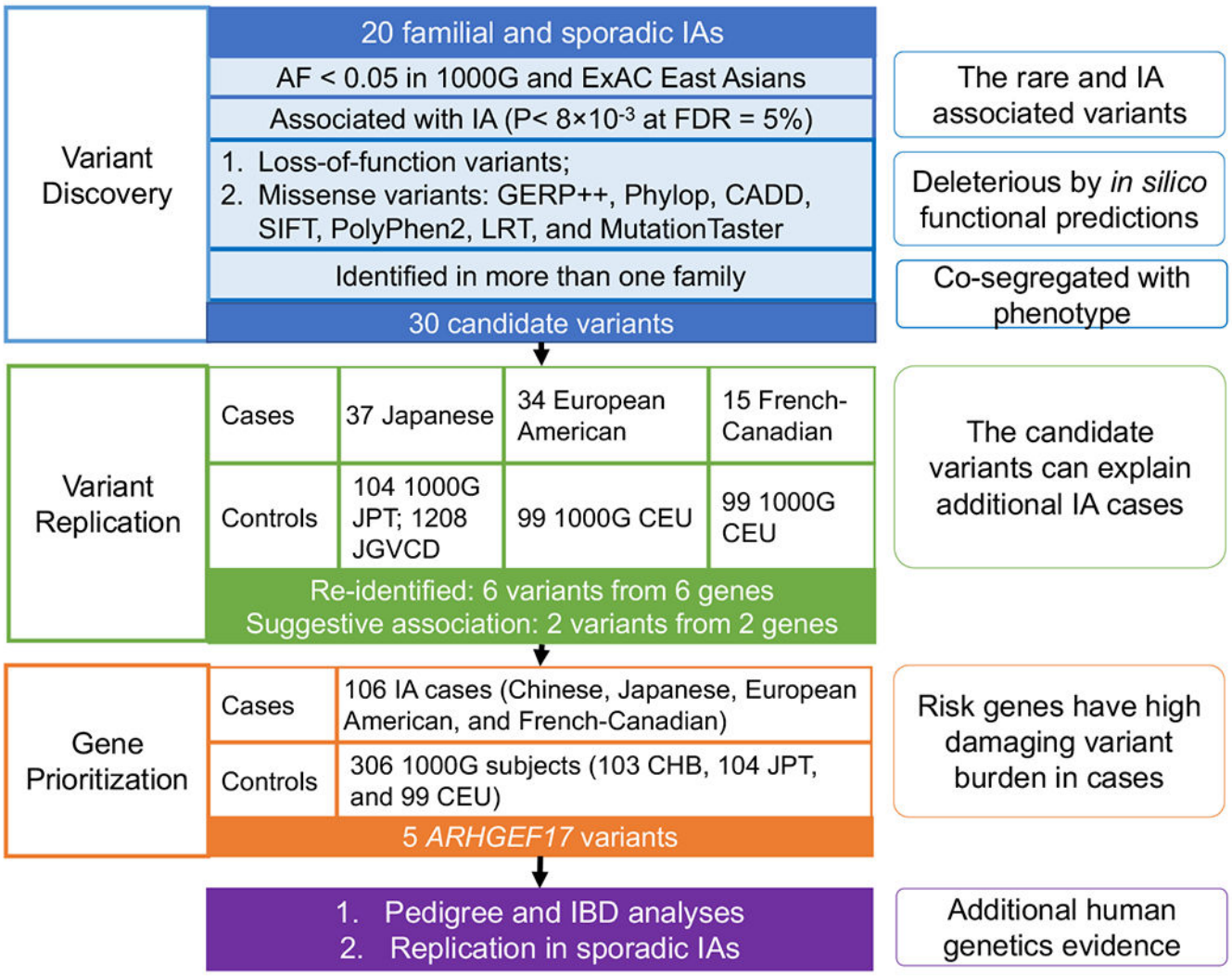
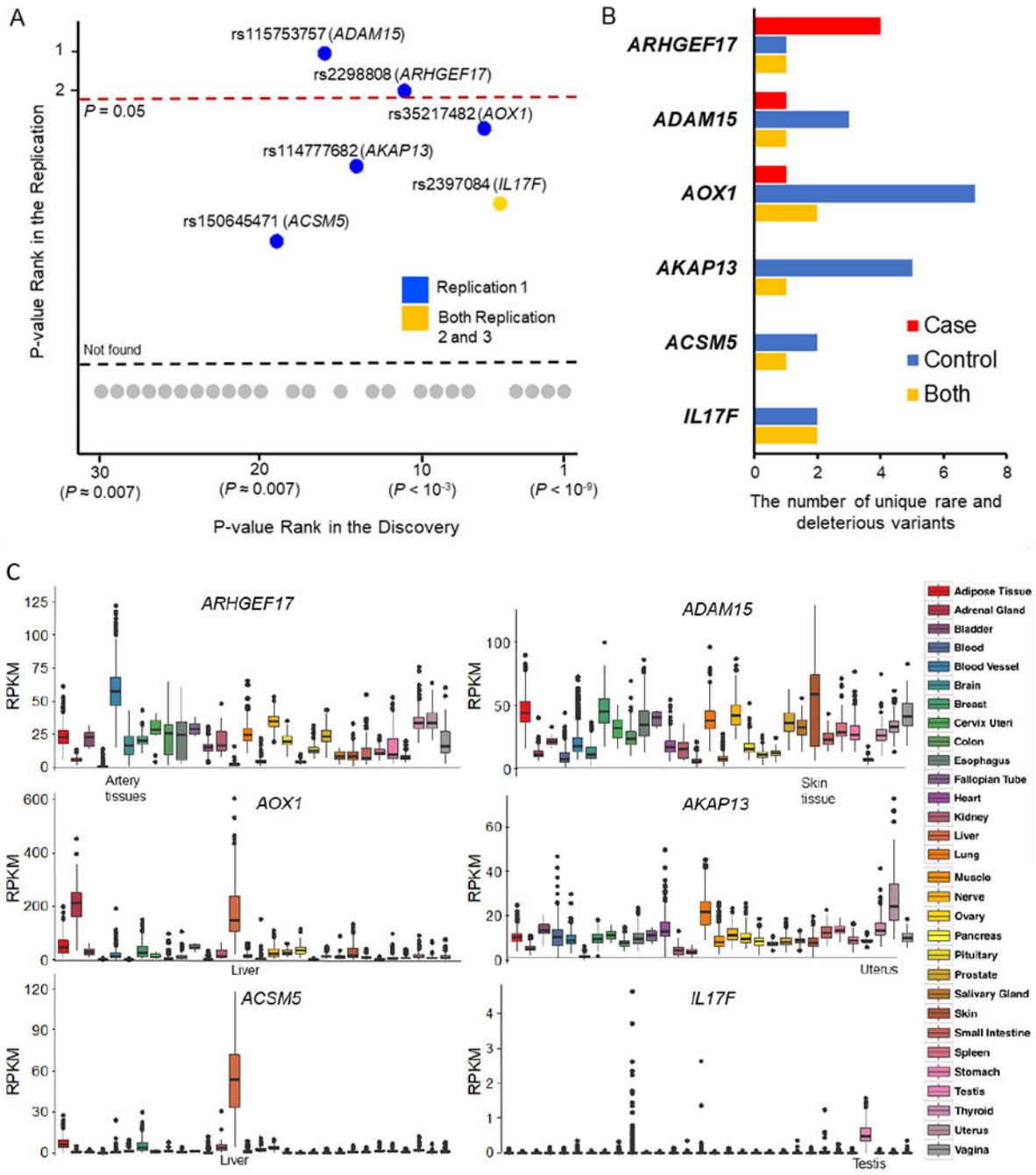


Figure 2. The flowchart of human genetics analyses.

We genome sequenced 20 Chinese IA cases. We performed an initial variant filtering analysis and discovered 30 candidate variants. This was followed by a replication of those 30 candidate variants using three independent cohorts including samples from Japanese (Replication 1), European American (Replication 2), and French-Canadian (Replication 3) cohorts. Finally, a gene-level variant burden test was performed on the genes with variant rediscovered in the replication samples. For the low-frequency damaging variant in the IA risk gene, we performed additional association study by comparing the variant count in 1031 unrelated Chinese sporadic IA cases to 599 Chinese controls.



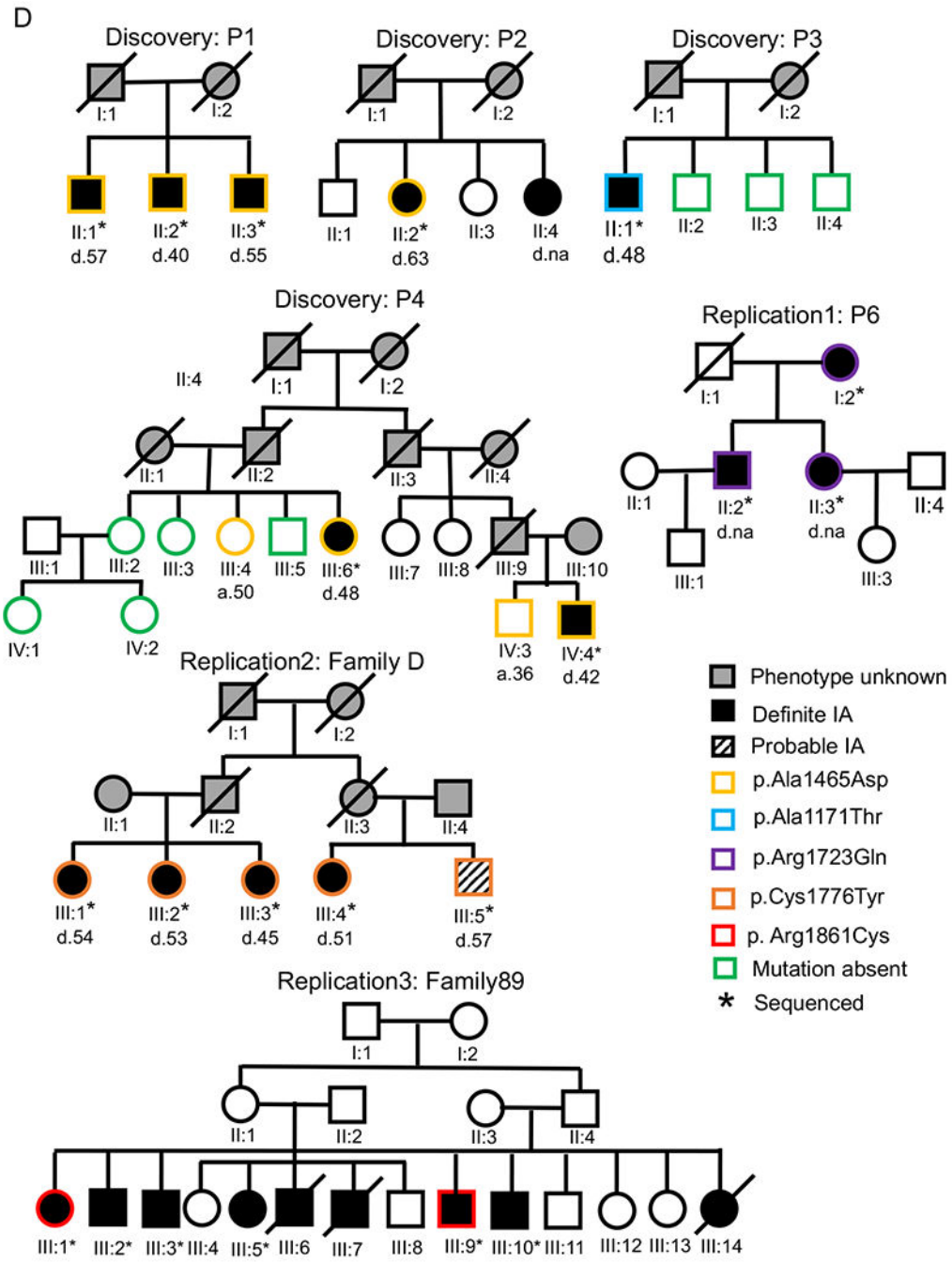


Figure 3. *ARHGEF17* is a candidate gene of IA

(A) The rank-order of 30 candidate variants in variant discovery and replication based on the p-values. The variants were ranked by P-values generated through two-sided Fisher's exact test in discovery samples and replication samples, respectively. **(B)** The number of unique rare and deleterious variants of the six genes with variants also found in replication samples. We collapsed the variant onto gene level and count how many distinct variants found in cases and controls, or in both. The allele frequency of the variant in cases and controls were not considered in this plot. **(C)** The expression profile of the six genes in different tissues.

The expression data were obtained from GTEx consortium. **(D)** Pedigrees of the affected patients with *ARHGEF17* mutations (two partially carried families Discovery P6 and Replication1 P1 were shown in the Figure S2A and Figure S3A). The age at diagnosis of IA (d.) is shown in years, and "a." indicates the age of the individual. The labels for the pedigrees in Replication1, Replication2, and Replication3 were the same with their labels previous publications as shown in Figure. S3. The unaffected samples were screened for the *ARHGEF17* mutations (c.4394C>A_p.Ala1465Asp and c.3511G>A_p.Ala1171Thr) by Sanger Sequencing.

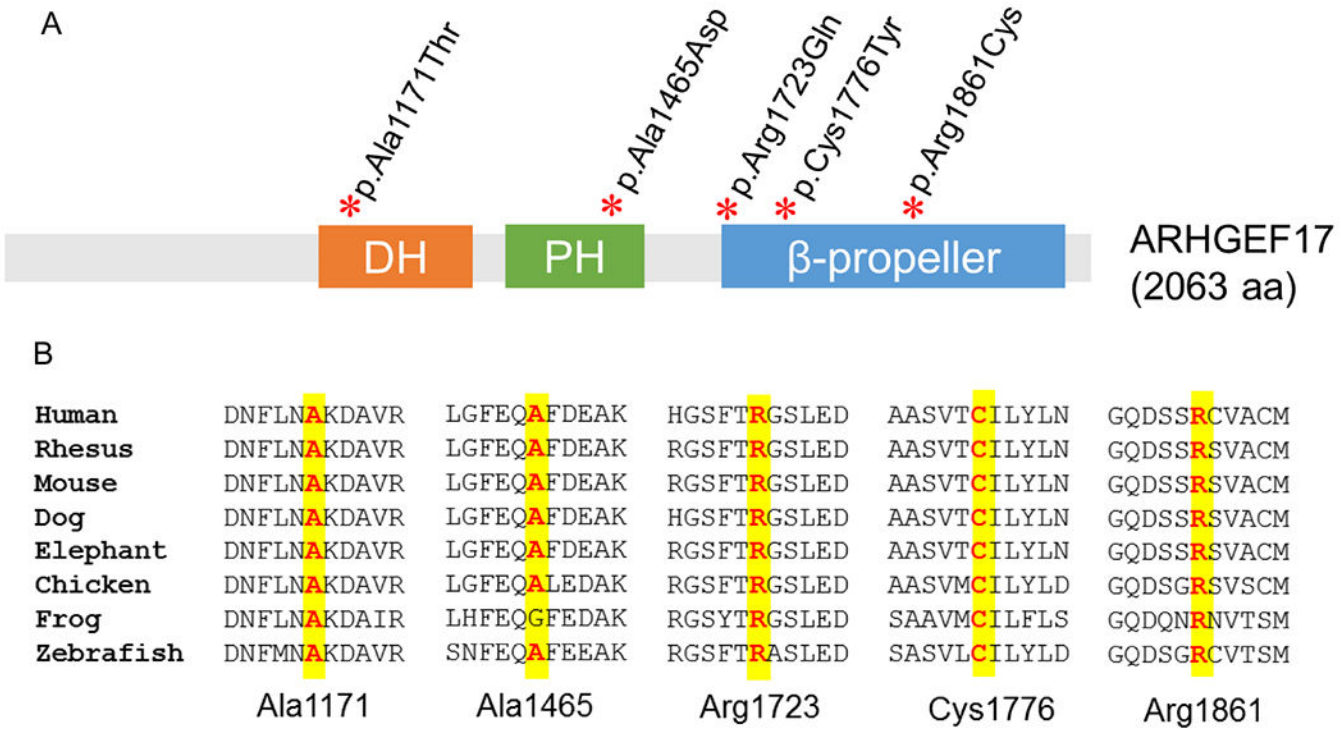


Figure 4. A schematic depicting the protein domain structure of ARHGEF17 (2,063 amino acids).

(A) ARHGEF17 encompasses the Dbl-homology domain (DH: 1,066–1,254; in orange), Pleckstrin homology domain (PH: 1,286–1,476; in green) and Beta-propeller (1,723–2,063; in blue). A red asterisk is used to indicate the position of each pathogenic variant identified in the nine families. **(B)** Conservation alignment indicating that the affected amino acids of ARHGEF17 are conserved across different vertebrates.

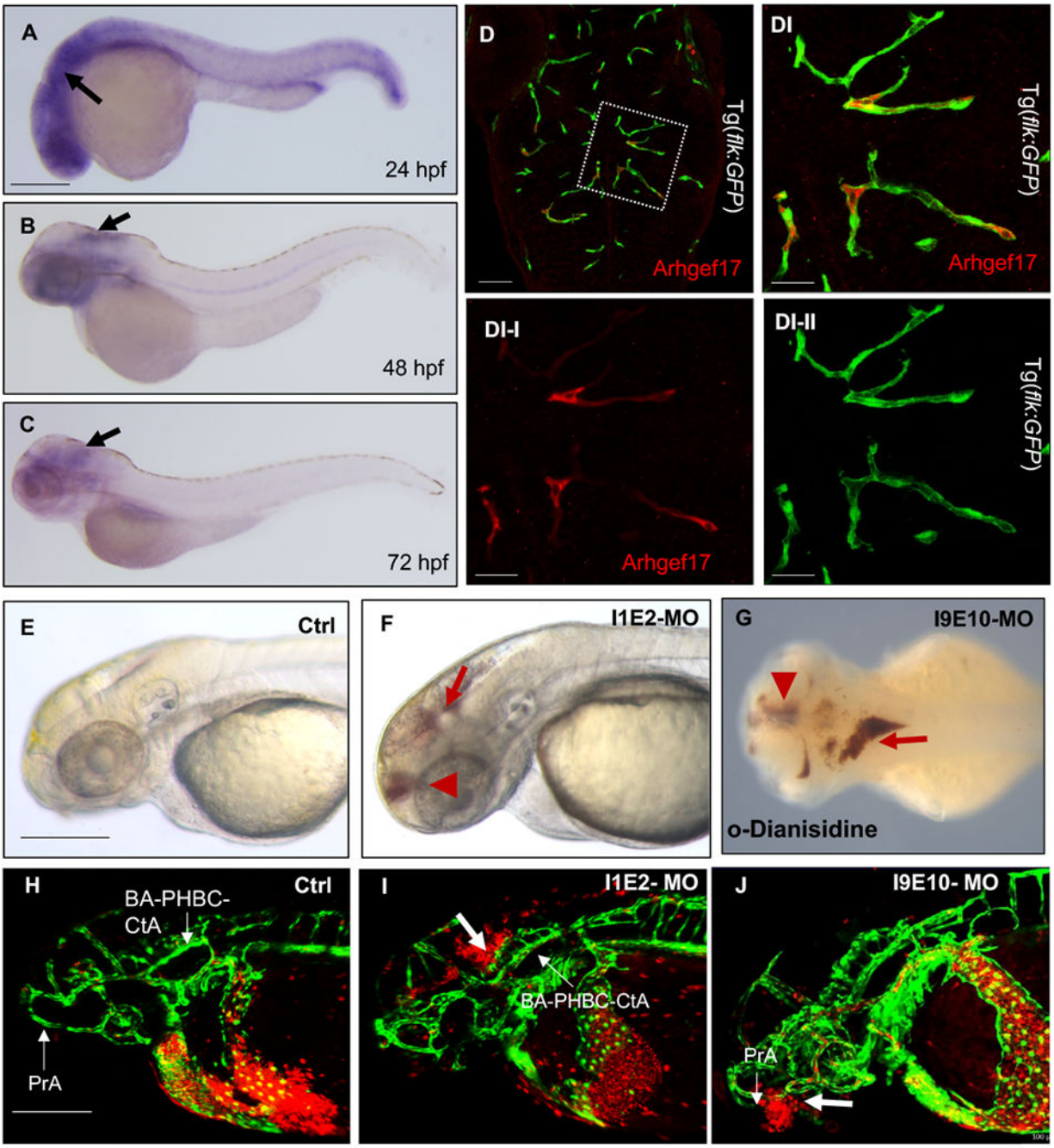


Figure 5. *arhgef17* deficiency causes intracranial hemorrhage in zebrafish

(A-C) Whole-mount *in situ* hybridization displaying *arhgef17* expression in the zebrafish brain (arrow) during development at 24 hpf, 48 hpf and 72 hpf. (**D**) Immunostaining indicates the expression of Arhgef17 (red, D, DI, DI-I) in zebrafish brain vessels (green, D, DI, DI-II) at 72 hpf. (**E, F**) Lateral view of zebrafish embryos showing intracranial hemorrhage in the forebrain (arrowhead) and the hindbrain (arrow) with I1E2-MO injection (F), compared to controls (E). (**G**) o-Dianisidine analyses showing erythrocyte extravasation in the forebrain (arrowhead) and the hindbrain (arrow) of I9E10-MO injected embryos. (**H**)

Confocal micrograph analyses of Tg*(kdrl:EGFP; gata1:DsRed)* embryos. Endothelial cells are marked by EGFP, and erythrocytes are indicated by DsRed. PrA: Prosencephalic artery; BA-PHBC-CtA: Vascular network comprising the basilar artery, primordial hindbrain channel and central arteries. (**I, J**) Intracranial bleeding (arrow) was detectable in the hindbrain adjacent to the BA-PHBC-CtA network in I1E2-MO-injected embryos (I), as well as the forebrain closed to PrA in I9E10-MO injected embryos (J). I1E2-MO or I9E10-MO: 3 ng per embryo injection; Ctrl: embryos injected by 3 ng *arhgef17* morpholinos with five nucleotide substitutes. 48 hpf (E-J). Scale bar: 200 μm (A-C; E-J); 50 μm (D); 20 μm (DI, DI-I, DI-II).

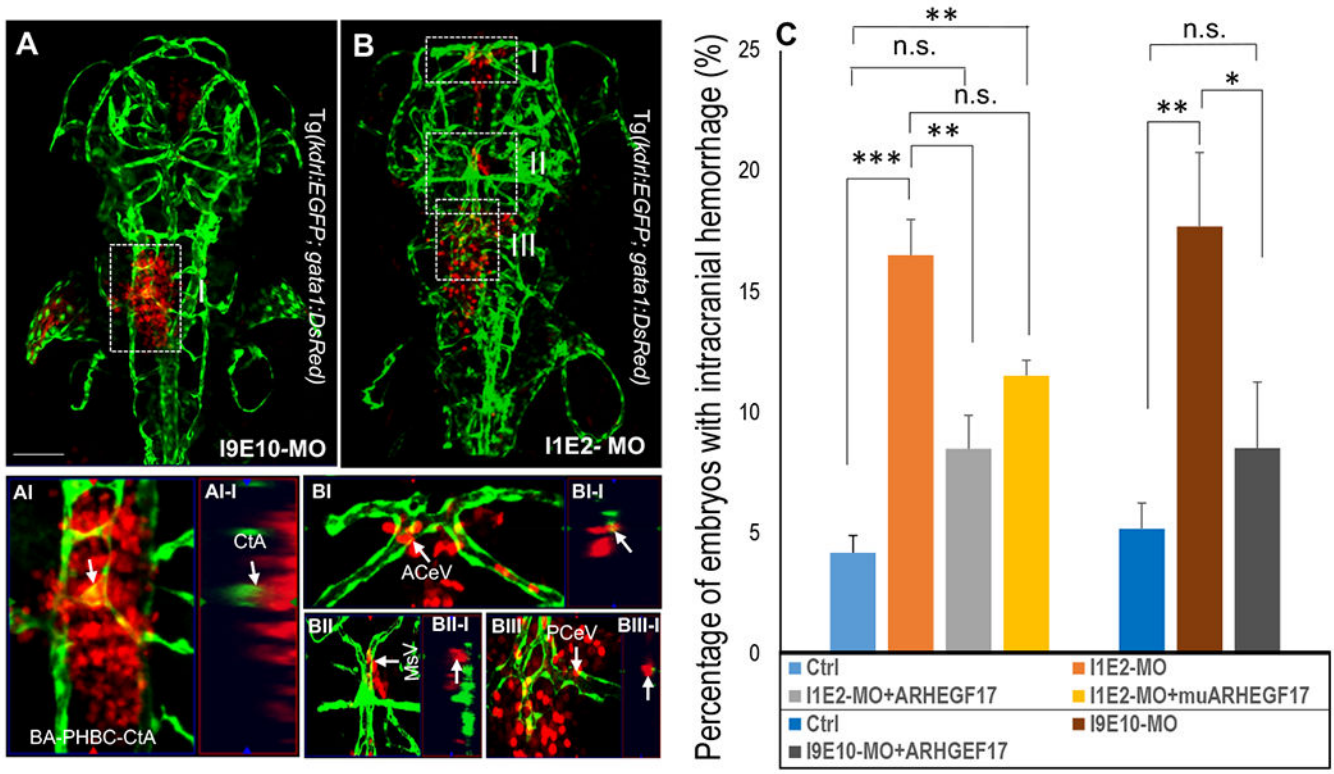


Figure 6. The co-injection of human *ARHGEF17* mRNA rescues intracranial hemorrhage in *arhgef17* deficient zebrafish.

(A) Confocal micrograph analyses revealing intracranial bleeding in the hindbrain in I9E10-MO-injected embryos. Extravasation of blood cells was detectable around the BA-PHPC-CtA network (arrow in AI). The endothelial leakage site was localized on a specific CtA (arrow in AI-I). (**B**) Intracranial hemorrhage was detectable in the forebrain (I), midbrain (II) and hindbrain (III) in embryos injected by I1E2-MO. Endothelial leakage sites were localized on ACeV (arrow in BI, BI-I), MsV (arrow in BII, BII-I) and PCeV (arrow in BIII, BIII-I). AI and BI-BIII show high-magnification of box regions in A and B, respectively. AI-I, BI-I, BII-I and BIII-I are confocal sections perpendicular to the Y-axis. Tg*(kdrl:EGFP; gata1:DsRed)*: Endothelium is marked by EGFP, and blood cells are labeled by DsRed; ACeV : anterior cerebral vein; MsV: mesencephalic vein; PCeV: posterior cerebral vein; 72 hpf. (**C**) Bar chart depicting percentages of embryos with intracranial hemorrhage in embryos injected by control-MO (n=688), I1E2-MO (n=533), I9E10-MO (n=233), I1E2-MO + human *ARHGEF17* mRNA (n=108), I1E2-MO + mu*ARHGEF17* mRNA (n=124) or I9E10-MO + human *ARHGEF17* mRNA (n=117). I1E2-MO or I9E10-MO: 3 ng per embryo injection. *ARHGEF17* mRNA: 160 pg per embryo injection. Experiments were repeated three times for each group. Numerical data are presented as mean in all panels (error bars show s.e.m.); statistical significance was calculated using ANOVA, *p<0.05, **p<0.01, ***p<0.001. Scale bar: 100 μm (A, B)

Table 1.

d deleterious variants of *ARHGEF17* prioritized from gene-level enrichment test in both cases and controls.

Variants	Allele count						Deleterious prediction[†] CADD/ SIFT/PP2/LRT/ MT)	AF in ExAC	Segregated Family[‡]	Partially Carried Family[‡]
	Discovery Chinese (N = 20)	Replication1 Japanese (N=37)	Replication2 European American (N=34)	Replication3 French Canadian (N=15)	Total[*] (N=106)	1000G subjects (N=306)				
p.Ala1171Thr	1	0	0	0	1	0	24.7/0.03/D/D/B	0	P3 (Discovery)	…
p.Ala1465Asp	3	0	0	0	10	10	24.2/0.08/D/D/D	2.6×10^{-3}	P1, P4 (Discovery)	P1-II2, II5, II9 (Replication1); P2-II2, P6-III6 (Discovery)
p.Arg1723Gln	3	0	0	0	3	0	32/0.04/D/D/D	0	P6 (Replication1)	…
p.Cys1776Tyr	0	5	0	0	5	0	18.47/0.01/B/D/D	3.1×10^{-4}	Family D (Replication2)	…
p.Arg1861Cys	0	0	0	2	2	0	33/0.004/D/D/D	6.6×10^{-5}	…	Family 89 - III1, III9 (Replication3)
p.Arg1329His ed in controls)	0	0	0	0	0	1	29.6/0.05/D/D/D	2.5×10^{-5}	…	…
otal	6	5	2	21	11					

$\times 10^{-7}$ (by Fisher's exact test; 21/106 versus 11/306; OR=6.6; 95% CI = 2.9–15.8);

bined Annotation Dependent Depletion; PP2, PolyPhen2 HDIV; LRT, likelihood ratio test; MT, MutationTaster; D, Damaging; B, Benign

ering the family as a unit, P = 4.3×10^{-5} (by Fisher's exact test; 9/36 versus 11/306; OR=8.8; 95% CI = 3.5-Inf)

Circ Genom Precis Med. Author manuscript; available in PMC 2019 July 01.