Genome-wide association analysis identifies new candidate risk loci for familial intracranial aneurysm in the French-Canadian population

Sirui Zhou1,2, Ziv Gan-Or1,3, Amirthagowri Ambalavanan1,3, Dongbing Lai4, Pingxing Xie1, Cynthia V. Bourassa1, Stephanie Strong1,3, Jay P. Ross1,3, Alexandre Dionne-Laporte1, Dan Spiegelman1, Nicolas Dupré5, Tatiana M Foroud4, Lan Xiong2,6, Patrick A. Dion1,7 & Guy A. Rouleau1,7

Intracranial Aneurysm (IA) is a common disease with a worldwide prevalence of 1–3%. In the French-Canadian (FC) population, where there is an important founder effect, the incidence of IA is higher and is frequently seen in families. In this study, we genotyped a cohort of 257 mostly familial FC IA patients and 1,992 FC controls using the Illumina NeuroX SNP-chip. The most strongly associated loci were tested in 34 Inuit IA families and in 32 FC IA patients and 106 FC controls that had been exome sequenced (WES). After imputation, one locus at 3p14.2 (FHIT, rs1554600, \( p = 4.66 \times 10^{-9} \)) reached a genome-wide significant level of association and a subsequent validation in Nunavik Inuit cohort further confirmed the significance of the FHIT variant association (rs780365, FBAT-O, \( p = 0.002839 \)). Additionally, among the other promising loci (\( p < 5 \times 10^{-6} \)), the one at 3q13.2 (rs78125721, \( p = 4.77 \times 10^{-7} \)), which encompasses CCDC80, also showed an increased mutation burden in the WES data (CCDC80, SKAT-O, \( p = 0.0005 \)). In this study, we identified two new potential IA loci in the FC population: FHIT, which is significantly associated with hypertensive IA, and CCDC80, which has potential genetic and functional relevance to IA pathogenesis, providing evidence on the additional risk loci for familial IA. We also replicated the previous IA GWAS risk locus 18q11.2, and suggested a potential locus at 8p23.1 that warrants further study.

Intracranial Aneurysm (IA) has a prevalence of 1–3% in the general population1,2. The rupture of an IA can lead to subarachnoid hemorrhages (SAH), which has devastating consequences. Environmental and genetic factors, such as hypertension and smoking, family history and ethnicity all contribute to the risk of IA. Because of the complexity of IA, genome-wide association studies (GWAS) have become the predominant strategy used to look for genetic factors associated with IA. These studies used several large cohorts with IA patients, mainly of Finnish, Japanese or European descent. Several risk loci were discovered in these GWAS studies: 8q11.23 (SOX17), 9p21.3-23.1 (CDKN2A-CDKN2AS)1,2 and 2q33.1 from the European and Japanese cohorts; 18q11.2 (RBBP8), 13q13.1 (STARD13-KL) and 10q24.32 from the Finnish and Japanese cohorts; 1q23.1, 3p25.2, 7p21.2, 9q31.3 and 4q31.22 (EDNRA)3 from two Japanese cohorts, and 7p21.1 (HDAC9)4 from a cohort with European ancestry. A more recent Finnish IA study revealed additional GWAS risk loci, including 2q23.3, 5q31.3 and 6q24.2, represented by low-frequency SNPs5. Multiple GWAS signals suggest that the genetic etiology of IA may be complex and population specific. The risk loci found in each GWA study was estimated to only account for 4.1–6.1% of the heritability in the respective cohort6.

1Montreal Neurological Institute and Hospital, McGill University, Montréal, QC, Canada. 2Department of Medicine, Faculty of Medicine, Université de Montréal, Montréal, QC, Canada. 3Department of Human Genetics, McGill University, Montréal, QC, Canada. 4Department of Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, IN, USA. 5Faculty of Medicine, Université Laval, Quebec, QC, Canada. 6Centre de recherche, Institut universitaire en santé mentale de Montréal, Montréal, QC, Canada. 7Department of Neurology and Neurosurgery, McGill University, Montréal, QC, Canada. Correspondence and requests for materials should be addressed to G.A.R. (email: guy.rouleau@mcgill.ca)
significant SNPs, representing 26 distinct loci that each reached suggestive level \( (p < 5 \times 10^{-8}) \); FC controls remained in the analysis. The genome-wide threshold for significance after Bonferroni correction was \( \lambda \). Genomic inflation factor markers, as shown in quantile-quantile (QQ) plot.

It has been reported that the French-Canadian (FC) population has higher IA/SAH incidence and that patients usually aggregate in large pedigrees \( (\text{p} = 0.002) \), with 30% of IA patients having a family history (fIA) \( (\text{p} = 0.01) \). Similarly, to the Finnish people, French-Canadians are also descended from a relatively small founder population and have population specific variants due to the population bottleneck and genetic drift. Therefore, we hypothesized that population specific and median/low frequency variants may play an important part in the disease risk in FC fIA.

Results

GWAS discovery phase. After data QC and sample pruning, 621,983 SNPs with 173 FC IA cases and 1,772 FC controls remained in the analysis. The genome-wide threshold for significance after Bonferroni correction was set to \( 5 \times 10^{-8} \). Marker-wise \( \text{p} \)-values of Cochran–Armitage trend test were performed using PLINK 1.9 \( (\text{p} = 0.0005) \) for genotyped variants. Genomic inflation factor \( \lambda = 1.02 \) indicated that there was little inflation of excessive significant markers, as shown in quantile-quantile (QQ) plot (Supplementary Figure S1).

The result of the initial trend test showed 3q13.2 as the most significant locus: rs2705520 \((\text{p} = 6.93 \times 10^{-8})\), an intronic SNP in AITG3 \((\text{autophagy} 3)\), followed by rs1877362 \((\text{p} = 1.16 \times 10^{-7})\) in CCDC80 \((\text{coiled-coil domain containing} 80)\) and rs1472107 \((\text{p} = 1.18 \times 10^{-7})\) in SLC35A5 \((\text{Solute Carrier Family 35 member A5})\) (Supplementary Figure S2).

After imputation using the Haplotype Reference Consortium (HRC) and the exclusion of low quality and low MAF variants, 7,614,484 remaining variants were included in the test of associations using SNPtest \( (\text{p} = 0.01) \). The results were shown in the Manhattan plot (Figure 1). One locus reached genome-wide significant level after imputation: 3p14.2 \((\text{p} = \text{rs1554600}, \text{OR} = 0.26, \text{p} = 4.66 \times 10^{-8})\) located in gene FHIT. TaqMan validation of rs1554600 on IA cases and controls suggested the imputation was accurate \((\text{maf.cases} = 0.0838 \text{ and maf.controls} = 0.026)\). The 28 most significant SNPs, representing 26 distinct loci that each reached suggestive level \((\text{p} < 5 \times 10^{-8})\), were prioritized for further validation (Table 1). A meta-analysis of 138 SNPs (Supplementary Table S1) located in 23 out of 26 loci comprising the current study and FIA cohort \( (\text{p} = 0.0001) \) (Table 3). However, the first two genes were less convincing with the limited number of variants in the testing \((\text{two variants for each})\). PABPC3 was unlikely to be a risk gene for IA due to its human testis-specific expression. Thresholds Test (VT) \( (\text{p} = 0.0001) \) focused on the selected genes revealed that only CCDC80 \((\text{p} = 0.01)\) in 3q13.2 reached the statistical significance after accounting for multiple testing.

In the Nunavik Inuit IA cohort, after performing Family Based Association Test (FBAT) \( (\text{p} = 0.002)\), DTNB, CCDC80, and PABPC3 were selected genes revealed that only CCDC80 \((\text{p} = 0.01)\) in 3q13.2 reached the statistical significance after accounting for multiple testing. It could be due to the limited sample size, which still suggested that FHIT variants may likely still be associated with IA.

Replication of previous GWAS risk loci. We also attempted to replicate the 12 IA risk loci identified in previous GWAS in our FC IA cohorts. 825 distinct LD blocks were established in these 12 loci in the FC.
population, the level of significance was therefore $p < 6.06 \times 10^{-5}$ after multiple correction. As a result, only one SNP rs35127791 located in 18q11.2 was replicated (maf = 0.157, $p = 5.05 \times 10^{-5}$, beta = $-0.66$, SE = 0.16) (Figure 2). LocusZoom plots covering the rest of the 23 genome-wide significant SNPs from previous GWAS in these 12 loci were shown in Supplementary Figure S3.

We further looked into the FC WES data in locus 18q11.2, which comprised 4 protein coding genes: GATA6, CTAGE1, RBBP8 and CABLES1, a variant burden test revealed exonic variants of CABLES1 seemed to be associated with IA in FC (p = 0.022, corrected) (Supplementary Table S3).

**Discussion**

In this study, we discovered a new IA associated region on 3p14.2 which encompasses the FHIT gene in the French-Canadians (Figure 3), intronic variants in FHIT also suggested an association with IA in the Nunavik Inuit population. Additionally, we found evidence suggesting exonic variants in CCDC80 within the 3q13.2 locus to be associated with the French-Canadian IA. Collectively, SNPs in FHIT and CCDC80 could explain approximately 3% of the heritability of IA in French-Canadians, higher than that was reported in the Finnish study (2.1%) and in the GWAS replication (2.5%)\(^6\). The underlying reason for this might be that French-Canadians are a more homogenous population and this study has mainly included familial cases. On the other hand, we also replicated a previous GWAS risk locus 18q11.2, the top SNP rs35127791 located approximately 200 kb upstream of FHIT, close to a H3K24Ac region in HUVEC cells. FHIT regulates the differentiative state of vascular smooth muscle cells, and an important candidate for cardiovascular development. CABLES1 proves to have an

Table 1. 26 loci (28 genes) with top SNPs reached promising level of association.

| LOCI | Marker | CHR | PDS | Allele1 | Allele2 | Zscore | P-value | Direction | PFC | SampleSize.FC | PFCIA | SampleSize.FIA |
|------|--------|-----|-----|---------|---------|--------|---------|-----------|-----|---------------|-------|---------------|
| 8p23.1 | rs76038736 | 8 | 10572792 | a | g | -4.753 | 2.01E-06 | — | 4.5E-06 | 1945 | 0.0093 | 4033 |
| 10p14 | rs7084131 | 10 | 8399121 | a | g | 4.686 | 2.78E-06 | + | 4.83E-06 | 1945 | 0.0114 | 4033 |
| 3p13.3 | rs4867353 | 5 | 31623146 | t | c | -4.608 | 4.07E-06 | — | 4.91E-06 | 1945 | 0.0148 | 4033 |

Table 2. Three loci with SNPs validated in the FIA cohort.

| LOCI | Marker | CHR | PDS | Allele1 | Allele2 | Zscore | P-value | Direction | PFC | SampleSize.FC | PFCIA | SampleSize.FIA |
|------|--------|-----|-----|---------|---------|--------|---------|-----------|-----|---------------|-------|---------------|
| 8p23.1 | rs76038736 | 8 | 10572792 | a | g | -4.753 | 2.01E-06 | — | 4.5E-06 | 1945 | 0.0093 | 4033 |
| 10p14 | rs7084131 | 10 | 8399121 | a | g | 4.686 | 2.78E-06 | + | 4.83E-06 | 1945 | 0.0114 | 4033 |
| 3p13.3 | rs4867353 | 5 | 31623146 | t | c | -4.608 | 4.07E-06 | — | 4.91E-06 | 1945 | 0.0148 | 4033 |
important role in cancer and development of neurons\textsuperscript{21}, a recent study also highlighted its function vascular cell senescence and inflammation through p21 regulation\textsuperscript{22}.

Unfortunately, we could not replicate the most significant SNP in \textit{FHIT} in another IA cohort. However, this was expected, as \textit{FHIT} variants have higher MAFs in French-Canadians compared to other European populations. The signal may only exist in French-Canadian populations and may correlate with IA cases with

Table 3. Exonic and splicing variants of 23 GWAS suggestive genes in 138 FC cases and controls from WES.

| Gene       | Loci   | Number of tested variants | P-value (SKAT) |
|------------|--------|---------------------------|----------------|
| ESRBG      | 1q44   | 2                         | 0.7073712      |
| SLC35F3    | 1q42.2 | 1                         | 0.002189217    |
| DTNB       | 2p23.3 | 2                         | 0.002747022    |
| PBR1       | 2p23.2 | 20                        | 0.1122009      |
| PAR3B      | 2p33.3 | 13                        | 0.9205563      |
| RRM3       | 3p24.1 | 1                         | 0.354907       |
| CCDC80     | 3q13.2 | 8                         | 0.0005070405   |
| CPNE4      | 3q22.1 | 1                         | 0.7073712      |
| PROM1      | 4p15.32| 12                        | 0.05065522     |
| PZD2       | 5p13.3 | 19                        | 0.7212919      |
| RSPY9      | 6p21.1 | 2                         | 0.8510591      |
| CARD11     | 7q22.1 | 4                         | 0.8172528      |
| CAMK2B     | 7p13   | 3                         | 0.09754801     |
| FTPRD      | 9p23   | 7                         | 0.2814774      |
| PA28C3     | 13q12.13| 54                        | 0.0001097577   |
| RGS6       | 14q24.2| 1                         | 0.7073712      |
| DNAJA4     | 15q25.1| 2                         | 0.5637608      |
| RBFOX1     | 16p13.2| 5                         | 1              |
| CARHSP1    | 16p13.2| 2                         | 0.3250554      |
| ANGPT4     | 20p13  | 4                         | 0.09886071     |
| SRP5G      | 20p13  | 5                         | 0.7073707      |
| HLCS       | 21q22.13| 6                        | 0.6115834      |
| SYN3       | 22q12.3| 3                         | 0.5083075      |

Figure 2. Regional association signals of the previous GWAS risk locus 18q11.2. LocusZoom plot showing the regional association of chr18.19.7-20.75 mb, including the most significant SNP rs35127791 and previous GWAS SNP rs11661542. Purple line indicates the genetic recombination rate (cM/Mb). SNPs in linkage disequilibrium with rs35127791 are shown in color gradient indicating r2 levels (hg19, 1KGP, Nov 2014, EUR).
hypertension. Among the three SNPs which were validated in the FIA cohort, rs76308736 is located in the pro-
moter region of SOX7, with its crucial function in angiogenesis and establishment of arteriovenous identity.
rs76308736 only showed suggestive significance in our discovery cohort and did not pass multiple correction in
the FIA data, the relationship of this SNP and risk for IA remains to be explored.

Although the number of cases in this study was limited, we tried improving the power by doing the follow-
ing: 1) targeting IA patients with family history, 2) focusing on individuals only with the French-Canadian eth-
nicity, and 3) validating the findings in exome sequencing results and in another founder population. Because
French-Canadians originated from a small founder population, we also included intermediate variants after
the imputation. The top SNPs that we discovered in FHIT and CCDC80 had rare or intermediate frequency
(2-4%). The MAF of rs1554600 was higher in French-Canadians (3.3%) compared to Europeans (1.9%) and
was significantly higher when compared with East Asians (0.1%), which suggest a bottleneck and/or drift may be
the reason for the accumulation of low-frequency variants that potentially associated with the risk of IA in the
French-Canadians.

FHIT (fragile histidine triad) is a tumor suppressor gene that regulates DNA replication and signals stress
responses23, and encompasses the most active of the common human chromosomal fragile regions (FRA3B).
Its expression has an important role in response to oxidative damage24. On the other hand, oxidative stress is
known to be a key contributor to IA formation and rupture25. Interestingly, a study highlighted that the SNPs
in FHIT have been associated with hypertensive traits in populations from Saguenay-Lac-St-Lean region,
mainly French-Canadians26. As over 40% of our French-Canadian IA cases were also affected with hyperten-
sion (Table 4), we considered that rs1554600 in FHIT is possibly more likely to be a risk for hypertensive IA in
French-Canadians. Further test of FHIT SNPs between IA patients with and without hypertension revealed sev-
eral SNPs in LD with rs1554600 to be significantly associated with this trait (ie. rs73098963, p = 0.002611, GW AS
p value = 2.42 × 10⁻⁸). FHIT has been reported in other GWAS studies to associate with hypertension: rs6782531,
which located at approximately 160 kb upstream and in high LD of rs1554600 have also been reported to be sig-
nificantly associated with blood pressure27.

The 3q13.2 locus is a gene-rich region with many functions in inflammatory responses. BTLA (B- and
T-lymphocyte attenuator) is involved in inflammatory responses28 and homeostasis of the immune system29. A
previous study showed BTLA expression to be up-regulated in organs after a hemorrhagic shock30. The 3q13.2
locus also comprises ATG3, which encodes a protein known to induce apoptosis31 and also to act as a regulator
of oxidant and inflammatory balance that regulates endothelial cell stress response32. The top SNP rs2705520 in
ATG3 was also reported in a GWAS to be associated with asthma33, suggesting its role in inflammatory diseases.

The most interesting gene in the 3q13.2 locus is CCDC80, also known as SSG1 (steroid-sensitive gene 1),
which is a cGMP signalling effector and was reported to be widely express in vascular smooth muscle cells34. A
previous study also showed that it has a role as a modulator of glucose and energy homeostasis35. Another study
highlighted that the product of fibroblast growth factor (FGF) regulates the expression of CCDC8036, which is
in turn also involved in cell adhesion during differentiation of fibroblast37. CCDC80 was reported as a tumor
suppressor as well38. These evidences suggest the critical function of CCDC80 in vascular formation. CCDC80
harbor a large number of rare variants in the French-Canadians (Supplementary Table S4); which also showed a
significant difference in the variation burden in cases and controls. Both ATG3 and CCDC80 are dosage sensitive genes\(^3^9,4^0\), therefore the potential different expression levels of those genes may affect the risk of IA.

In conclusion, we have provided evidence for four new loci associated with IA in French-Canadian IA cohort recruited from Montréal and Québec city, which could explain 3% of the disease heritability. Based on the findings of this study and the functions of their encoded products, two genes (FHIT and CCDC80) are potentially relevant to IA with strong aggregation of familial IA cases with high blood pressure in the French-Canadian population. FHIT is more particularly associated with hypertensive IA cases and this may be the consequence of a bottleneck and/or drift that affected the French-Canadian founder population. CCDC80 was shown to have a large number of rare variations in the French-Canadian cohort and with a significantly different variation burdens between IA cases and controls. Both the lack of association of SNPs in FHIT and CCDC80 and the replication of the only 18q11.2 locus of the previous GWAS hits suggests a genetic heterogeneity in IA, and thus additional studies targeting other high-risk populations are needed. However, the limited number of cases available in our study calls for a validation study that will have access to a larger cohort from the same founder population, thus to increase the power of detection.

Methods
Discovery cohort. The discovery cohort included 257 French-Canadian IA patients, a majority of them were with family history, which were recruited in Montréal and Québec City, Canada. The diagnoses were confirmed either by magnetic resonance angiography (MRA), or by surgical confirmation (clipped or coiled). An additional 1,992 controls, mainly comprised of unrelated FC individuals without cerebrovascular diseases were also included in the analysis. Their demographic information is listed in Table 4. Written informed consent were obtained from all participants, and this manuscript contain no identifying information for any participant. This study was approved by Comité d'éthique de la recherche du Centre hospitalier de l’Université de Montréal and McGill University ethics, all methods were performed in accordance with the relevant guidelines and regulations of McGill University (REB NEU-14-051).

Genotyping and quality control. All patients and controls were genotyped using the Illumina NeuroX SNP-chip, which contains 719,885 markers and is comprised of the backbone of Illumina HumanOmniExpress-v24 BeadChip. Raw data was processed by Illumina GenomeStudio software before the genotypes were generated. Both markers and samples were passed through a series of quality control (QC) steps. Samples were first removed if duplicated or if they had one of the following issues: 1) sex discrepancies; 2) exceeded a missing rate of 0.02; 3) ethnical admixture determined by PCA; or 4) with cryptic relateness determined by PLINK. Markers were removed if they meet one of the following criteria: 1) having a minor allele frequency (MAF) lower than 0.01; 2) deviating from Hardy Weinberg Equilibrium mined by PLINK. Markers were removed if they meet one of the following criteria: 1) exceeding a missing rate of 0.02; 3) ethnical admixture determined by PCA; or 4) with cryptic relateness determined by PLINK. Markers were removed if they meet one of the following criteria: 1) exceeding a missing rate of 0.02; 3) deviating from Hardy Weinberg Equilibrium (p < 0.0001). Principal Component Analysis (PCA) implelemented in the package EIGENSOFT 6.0\(^4^1\) was performed to assess the ethnicity of the samples. Three distinct populations CEU, CHB and YRI from 1000 Genome (1KGP) Phase III were used for clustering and CEU outliers were removed from further analysis. The remaining homogeneous population was also adjusted for the principal components in the subsequent tests for associations.

Imputation. Imputation was done by the Sanger Imputation Server (https://imputation.sanger.ac.uk/) using Haplotype Reference Consortium r1.1\(^4^2\), and were pre-phased using SHAPEIT2\(^4^3\). Imputed variants were included in the further analyses only with MAF >0.01 and with imputation quality score >0.3.

Association analysis. Frequentist additive association implemented in SNPtest\(^1^3\) was used to test for association of the imputed dataset, between FC IA cases and controls. Five major principal components were used as covariates for ancestry adjustment along with the sex of samples. Only autosomal SNPs were analyzed.

Regional association of suggestive loci were plotted using LocusZoom\(^4^4\) with LD data from 1KGP CEU population.

| Sample demographic of cases and controls | Cases | Controls |
|-----------------------------------------|-------|---------|
| Number                                  | 257   | 1192    |
| Mean age of recruit (SD)                | 53.9 (11.2) | 54.4 (16.8) |
| % Male                                  | 31.5% | 60.6%   |

| Clinical information of cases (212)     | With family history | Sporadic |
|-----------------------------------------|---------------------|---------|
| With clinical information               | 147                 | 65      |
| Multiple lesions (%)                    | 55 (37.4)           | 19 (29.2)|
| SAH (%)                                 | 51 (34.7)           | 25 (38.5)|
| Hypertension (%)                        | 60 (40.8)           | 31 (47.7)|
| Drinking (%)                            | 33 (22.4)           | 26 (40.0)|
| Smoking (%)                             | 89 (60.5)           | 50 (76.9)|
| Hypercholesterolemia (%)                | 29 (19.7)           | 24 (36.9)|

Table 4. FC sample demographics and the clinical information of IA cases.
Heritability estimation. We estimated the heritability from the original and imputed variants within the most promising loci, using methods of Estimation of Variance explained by SNPs (GREML) and GREML-LDMS programs implemented in the package of the Genome-wide Complex Trait Analysis (GCTA).

Meta-analysis. SNP identified in the current study that reached suggestive level of association ($p < 5 \times 10^{-6}$) were compared with the previous FIA GWAS summary statistics from 2,617 IA cases and 1,416 controls. METAL was used to conduct meta-analysis of the two GWAS results for these selected SNPs.

FC and Inuit IA WES data. We examined the loci from the GWAS signals that have reached suggestive level of association in the exome sequencing results of 32 selected FC IA cases and 106 FC controls. Variable Thresholds Test (VT) implemented in Variant Tools (Vtools) and Sequence Kernel Association Test (SKAT) were performed to test the exonic variation burden in the genes located in the GWAS significant regions.

Thirty-four Nunavik Inuit (Québec, Canada) families comprised of 49 IA patients and 124 family controls were also used to follow up the regions of significance. The samples were genotyped on Illumina HumanOmniExpress-v24 Beadchip which contains 730,525 SNPs. We looked into all the suggestive loci of the FC IA discovery cohort, and performed family-based association analysis (fbat) implemented in FBAT package in the Inuit SNP-chip data to test the case-control association in related individuals.

Replication of previous IA GWAS loci. 12 loci from previous GWA studies (Supplementary Figure S3) with 23 SNPs that reached genome-wide significance were selected to examine if they could be replicated in our study. 500 kb upstream/downstream of the first and last genome-wide significant SNP located in each of the 12 loci were investigated in our GWAS data. Multiple correction was performed on the number of independent tests, which was defined by the number of LD blocks within these 12 regions. LocusZoom was used for data plotting. The validated loci will further be examined in FC WES data using Genome Analysis Toolkit (GATK) and VTools.

References
1. Rinkel, G. J. Natural history, epidemiology and screening of unruptured intracranial aneurysms. J Neurol Radiol 35, 99–103, https://doi.org/10.1016/j.neurad.2007.11.004 (2008).
2. Vernooij, M. W. et al. Incidental findings on brain MRI in the general population. N Engl J Med 357, 1821–1828, https://doi.org/10.1056/NEJMoa0707972 (2007).
3. Wiebers, D. O. et al. Unruptured intracranial aneurysms: natural history, clinical outcome, and risks of surgical and endovascular treatment. Lancet 362, 103–110 (2003).
4. Bilguvar, K. et al. Susceptibility loci for intracranial aneurysm in European and Japanese populations. Nat Genet 40, 1472–1477, https://doi.org/10.1038/ng.240 (2008).
5. Foroud, T. et al. Genome-wide association study of intracranial aneurysms confirms role of Anril and SOX17 in disease risk. Stroke 43, 2846–2852, https://doi.org/10.1161/STROKEAHA.112.656397 (2012).
6. Akiyama, K. et al. Genome wide association study to identify genetic variants present in Japanese patients harboring intracranial aneurysms. J Hum Genet 55, 656–661, https://doi.org/10.1038/hjg.2010.82 (2010).
7. 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 21, 2012–2110, https://doi.org/10.1093/hmg/ddo020 (2012).
8. Foroud, T. et al. Genome-wide association study of intracranial aneurysm identifies a new association on chromosome 7. Stroke 45, 3194–3199, https://doi.org/10.1161/STROKEAHA.114.006096 (2014).
9. Kurki, M. I. et al. High risk population isolate reveals low frequency variants predisposing to intracranial aneurysms. PLoS Genet 10, e1004134, https://doi.org/10.1371/journal.pgen.1004134 (2014).
10. Halal, F., Mohr, G. & Toussi, T. & Napoleon Martinez, S. Intracranial aneurysms: a report of a large pedigree. Am J Med Genet 15, 89–95, https://doi.org/10.1002/ajmg.1320150112 (1985).
11. Mathieu, J. et al. Epidemiological study of ruptured intracranial aneurysms in the Saguenay-Lac-Saint-Jean region (Québec, Canada). Can J Neurol Sci 23, 184–188 (1996).
12. Chang, C. C. et al. Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience 4, 7, https://doi.org/10.1186/s13742-015-0047-8 (2015).
13. Marchini, J., Howie, B., Myers, S., McVean, G. & Donnelly, P. A new multtipoint method for genome-wide association studies by imputation of genotypes. Nat Genet 39, 906–913, https://doi.org/10.1038/ng2088 (2007).
14. Yang, J. et al. Common SNPs explain a large proportion of the heritability for human height. Nat Genet 42, 565–569, https://doi.org/10.1038/ng.608 (2010).
15. Yang, J., Lee, S. H., Goddard, M. E. & Visscher, P. M. GCTA: a tool for genome-wide complex trait analysis. Am J Hum Genet 88, 76–82, https://doi.org/10.1016/j.ajhg.2010.11.011 (2011).
16. Yang, J. et al. Genetic variance estimation with imputed variants finds negligible missing heritability for human height and body mass index. Nat Genet 47, 1114–1120, https://doi.org/10.1038/ng.3390 (2015).
17. Wu, M. C. et al. Rare-variant association testing for sequencing data with the sequence kernel association test. Am J Hum Genet 89, 82–93, https://doi.org/10.1016/j.ajhg.2011.05.029 (2011).
18. Price, A. L. et al. Pooled association testing for rare variants in exom-esequencing studies. Am J Hum Genet 86, 832–838, https://doi.org/10.1016/j.ajhg.2010.04.005 (2010).
19. Laird, N. M., Horvath, S. & Xu, X. Implementing a unified approach to family-based tests of association. Genet Epidemiol 19(Suppl 1), S36–42, https://doi.org/10.1002/1098-2272(2000)19:1<1::AID-GEPI6>3.0.CO;2-M (2000).
20. Tasuno, K. et al. Genome-wide association study of intracranial aneurysm identifies three new risk loci. Nat Genet 42, 420–425, https://doi.org/10.1038/ng.563 (2010).
21. Huang, J. R., Tan, G. M., Li, Y. & Shi, Z. The Emerging Role of Cables1 in Cancer and Other Diseases. Mol Pharmacol 92, 240–245, https://doi.org/10.1124/mol.116.107730 (2017).
22. Pu, Z. et al. Cables1 Inhibits Proliferation and Induces Senescence by Angiotensin II via a p21-Dependent Pathway in Human Umbilical Vein Endothelial Cells. J Vasc Res 54, 13–21, https://doi.org/10.1159/000452409 (2017).
23. Barnes, L. D. et al. Fhit, a putative tumor suppressor in humans, is a dimineoic acid3-P1,3-triphosphate hydrolase. Biochemistry 35, 11529–11535, https://doi.org/10.1021/bi961413t (1996).
24. Karras, J. R., Paisie, C. A. & Huelner, K. Replicative Stress and the FHit Gene: Roles in Tumor Suppression, Genome Stability and Prevention of Carcinogenesis. Cancers (Basel) 6, 1208–1219, https://doi.org/10.3390/cancers6021218 (2014).
25. Starke, R. M. et al. The role of oxidative stress in cerebral aneurysm formation and rupture. Curr Neurol Neurosci Rep 10, 247–253 (2013).
Acknowledgements

P.X. and Z.G.O. are the recipient of fellowship from the Canadian Institutes of Health Research. G.A.R. funded this study from the Canadian Heart and Stroke Foundation. G.A.R. holds the Canada’s Research Chair in Neurogenetics and the Wilder Penfield Chair in Neuroscience.

Author Contributions

S.Z. designed the study, performed the analyses and wrote the manuscript; Z.G.O., C.V.B. and N.D. provided samples and data; D.L. and T.M.F. provided validation data for the meta-analysis; A.A., P.X., A.D.L. and D.S. maintained the bioinformatics pipeline and helped the analyses; S.S. and J.P.R. co-performed the experiments; L.X. and P.A.D. provided suggestions to the study design and co-wrote the manuscript; G.A.R. funded the study, collected the samples and revised the manuscript.

Additional Information

Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-018-21603-7.

Competing Interests: The authors declare no competing interests.

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2018