



# Genome-wide association study of intracranial aneurysms identifies 17 risk loci and genetic overlap with clinical risk factors

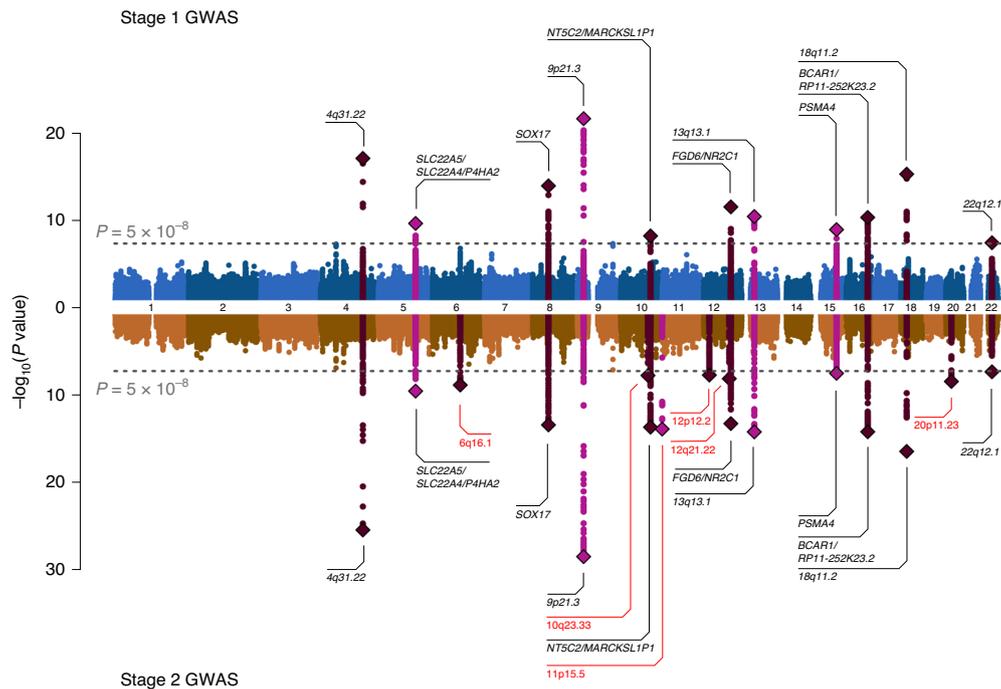
Mark K. Bakker<sup>1</sup>✉, Rick A. A. van der Spek<sup>1</sup>, Wouter van Rheenen<sup>1</sup>, Sandrine Morel<sup>2,3</sup>, Romain Bourcier<sup>4,5</sup>, Isabel C. Hostettler<sup>6,7</sup>, Varinder S. Alg<sup>6</sup>, Kristel R. van Eijk<sup>1</sup>, Masaru Koido<sup>8,9</sup>, Masato Akiyama<sup>8,10,11</sup>, Chikashi Terao<sup>8</sup>, Koichi Matsuda<sup>12,13</sup>, Robin G. Walters<sup>14,15</sup>, Kuang Lin<sup>14</sup>, Liming Li<sup>16</sup>, Iona Y. Millwood<sup>14,15</sup>, Zhengming Chen<sup>14,15</sup>, Guy A. Rouleau<sup>17</sup>, Sirui Zhou<sup>18</sup>, Kristiina Rannikmäe<sup>19</sup>, Cathie L. M. Sudlow<sup>19,20</sup>, Henry Houlden<sup>21</sup>, Leonard H. van den Berg<sup>1</sup>, Christian Dina<sup>4</sup>, Olivier Naggara<sup>22,23</sup>, Jean-Christophe Gentric<sup>24</sup>, Eimad Shotar<sup>25</sup>, François Eugène<sup>26</sup>, Hubert Desal<sup>4,5</sup>, Bendik S. Winsvold<sup>27,28</sup>, Sigrid Børte<sup>28,29,30</sup>, Marianne Bakke Johnsen<sup>28,29,30</sup>, Ben M. Brumpton<sup>28</sup>, Marie Søfteland Sandvei<sup>31,32</sup>, Cristen J. Willer<sup>33</sup>, Kristian Hveem<sup>28,34</sup>, John-Anker Zwart<sup>27,28,30</sup>, W. M. Monique Verschuren<sup>35,36</sup>, Christoph M. Friedrich<sup>37,38</sup>, Sven Hirsch<sup>39</sup>, Sabine Schilling<sup>39</sup>, Jérôme Dauvillier<sup>40</sup>, Olivier Martin<sup>40</sup>, HUNT All-In Stroke\*, China Kadoorie Biobank Collaborative Group\*, BioBank Japan Project Consortium\*, The ICAN Study Group\*, CADISP Group\*, Genetics and Observational Subarachnoid Haemorrhage (GOSH) Study investigators\*, International Stroke Genetics Consortium (ISGC)\*, Gregory T. Jones<sup>41</sup>, Matthew J. Bown<sup>42,43</sup>, Nerissa U. Ko<sup>44</sup>, Helen Kim<sup>45,46,47</sup>, Jonathan R. I. Coleman<sup>48,49</sup>, Gerome Breen<sup>48,49</sup>, Jonathan G. Zaroff<sup>50</sup>, Catharina J. M. Klijn<sup>51</sup>, Rainer Malik<sup>52</sup>, Martin Dichgans<sup>53,54</sup>, Muralidharan Sargurupremraj<sup>55,56</sup>, Turgut Tatlisumak<sup>57</sup>, Philippe Amouyel<sup>58</sup>, Stéphanie Debette<sup>55,56</sup>, Gabriel J. E. Rinkel<sup>1</sup>, Bradford B. Worrall<sup>59</sup>, Joanna Pera<sup>60</sup>, Agnieszka Slowik<sup>60</sup>, Emília I. Gaál-Paavola<sup>61,62</sup>, Mika Niemelä<sup>61</sup>, Juha E. Jääskeläinen<sup>63,64</sup>, Mikael von Und Zu Fraunberg<sup>63,64</sup>, Antti Lindgren<sup>63,64</sup>, Joseph P. Broderick<sup>65</sup>, David J. Werring<sup>6</sup>, Daniel Woo<sup>65</sup>, Richard Redon<sup>4</sup>, Philippe Bijlenga<sup>3</sup>, Yoichiro Kamatani<sup>8,12</sup>, Jan H. Veldink<sup>1,72</sup> and Ynte M. Ruigrok<sup>1,72</sup>✉

**Rupture of an intracranial aneurysm leads to subarachnoid hemorrhage, a severe type of stroke. To discover new risk loci and the genetic architecture of intracranial aneurysms, we performed a cross-ancestry, genome-wide association study in 10,754 cases and 306,882 controls of European and East Asian ancestry. We discovered 17 risk loci, 11 of which are new. We reveal a polygenic architecture and explain over half of the disease heritability. We show a high genetic correlation between ruptured and unruptured intracranial aneurysms. We also find a suggestive role for endothelial cells by using gene mapping and heritability enrichment. Drug-target enrichment shows pleiotropy between intracranial aneurysms and antiepileptic and sex hormone drugs, providing insights into intracranial aneurysm pathophysiology. Finally, genetic risks for smoking and high blood pressure, the two main clinical risk factors, play important roles in intracranial aneurysm risk, and drive most of the genetic correlation between intracranial aneurysms and other cerebrovascular traits.**

An intracranial aneurysm is a balloon-shaped dilatation, usually located at a branch of an intracranial artery. It is present in 3% of the population<sup>1</sup>. Rupture of an intracranial aneurysm causes an aneurysmal subarachnoid hemorrhage (aSAH), a severe type of stroke. Approximately one-third of patients die, and another

third remain dependent for daily life activities<sup>2</sup>. Intracranial aneurysms occur in relatively young people with a mean age of 50 years and is twice as common in women aged >50 years compared with men of that age. Genetic predisposition plays an important role in the disease with an aSAH heritability of 41%, as estimated in a twin study<sup>3</sup>.

A full list of author affiliations appears at the end of the paper.



**Fig. 1 | GWAS meta-analysis association results.** SAIGE logistic mixed-model association  $P$  values of the stage 1 (upward direction) and stage 2 (downward direction) GWAS meta-analyses. The horizontal axis indicates chromosomal position. The vertical axis indicates  $-\log_{10}(P \text{ value})$  of the association. The dotted lines indicate the genome-wide-significance threshold of  $P = 5 \times 10^{-8}$ . Lead SNPs of each locus are highlighted with a diamond, and SNPs in close proximity ( $\pm 500$  kb) are colored in pink or purple, depending on the chromosome index parity. Labels are gene or locus names annotated using SMR, eCAVIAR and TWAS, or prior information of intracranial aneurysm-associated genes. Labels or loci identified only in the stage 2 GWAS are shown in red.

Much is still unknown about the genetic architecture of intracranial aneurysms<sup>4,5</sup>. Family-based studies identified a number of variants with Mendelian inheritance<sup>6–10</sup>, but genome-wide association studies (GWAS) have identified multiple common variants, suggesting a polygenic model of inheritance<sup>5,11–13</sup>. The largest GWAS published to date, involving 2,780 cases and 12,515 controls, identified 6 risk loci<sup>11,13</sup>. Based on that GWAS, the explained SNP-based heritability of intracranial aneurysms was estimated as being only 4.1–6.1%, depending on the population<sup>5</sup>.

We aimed to further characterize the genetic architecture of intracranial aneurysms by performing a cross-ancestry GWAS meta-analysis on a total of 10,754 cases and 306,882 controls from a wide range of European and East Asian ancestries. We included cases with both unruptured intracranial aneurysms and aSAHs (that is, with ruptured intracranial aneurysms), enabling us to identify potential risk factors specific for intracranial aneurysm rupture. We also looked for genetic similarities between intracranial aneurysms and related traits, including other types of stroke, vascular malformations and other aneurysms, and analyzed whether known risk factors for intracranial aneurysms play a causal genetic role. Furthermore, we investigated enrichment of genetic associations in functional genetic regions, tissue subtypes and drug classes to provide insight into intracranial aneurysm pathophysiology.

## Results

**GWAS of intracranial aneurysms.** Our GWAS meta-analysis on intracranial aneurysms consisted of two stages. The stage 1 meta-analysis included all individuals of European ancestry and consisted of individual-level genotypes from 23 different cohorts, which were merged into 9 European-ancestry strata based on genotyping platform and country. These strata were each analyzed in a logistic mixed model<sup>14</sup> and then meta-analyzed, while also including summary statistics from a population-based cohort

study: the Nord-Trøndelag Health Study (the HUNT Study). This resulted in 7,495 cases, 71,934 controls and 4,471,083 SNPs passing quality control (QC) thresholds (see Methods and Supplementary Table 1). Stage 2 was a cross-ancestry meta-analysis including all stage 1 strata and summary statistics of East Asian individuals from two population-based cohort studies: the BioBank Japan (BBJ) and the China Kadoorie Biobank (CKB). This totaled 10,754 cases, 306,882 controls and 3,527,309 SNPs in stage 2 (Supplementary Table 1).

The stage 1 association study resulted in 11 genome-wide-significant loci ( $P \leq 5 \times 10^{-8}$ ; Fig. 1 and Supplementary Table 2). Transancestry genetic correlation analysis showed a strong correlation between the stage 1 meta-analysis of European ancestry and an analysis including only East Asian ancestry samples (genetic correlation,  $\rho_g = 0.938 \pm 0.165$  (s.e.) for genetic impact and  $0.908 \pm 0.146$  for genetic effect; Supplementary Table 3). Stage 2 increased the number of genome-wide-significant loci to 17 (Table 1 and Fig. 1). All but two loci (8q11.23, rs6997005 and 15q25.1, rs10519203) were also associated with intracranial aneurysms in the samples of East Asian ancestry added in stage 2 ( $P < 0.05/11$ ), and two loci were monomorphic in East Asians (Table 1). The stage 2 loci included 11 new risk loci and 6 previously reported risk loci<sup>11</sup>. We used conditional and joint (COJO) analysis (GCTA v.1.91.1beta)<sup>15</sup> to condition the stage 1 GWAS summary statistics on the lead SNP in each locus. We found that none of the loci consisted of multiple independent SNPs and that each locus tagged a single causal variant (data not shown). Genomic inflation factors ( $\lambda_{GC}$ ) were 1.050 for the stage 1 meta-analysis and 1.065 for stage 2 (Supplementary Fig. 1 and Supplementary Table 4). The linkage disequilibrium score regression (LDSR) intercept was  $0.957 \pm 0.008$  (s.e.) for the stage 1 meta-analysis and  $0.982 \pm 0.008$  for the East Asian subset. This indicated that, in all GWAS analyses, observed inflation was due to polygenic architecture.

**Table 1 | Lead associations of genome-wide-significant risk loci**

SNP	Locus	Chromosome	Position	EA	OA	Stage	EA F	$\beta$	s.e.	P value	Annotated genes	Associated traits
rs6841581	4q31.22 <sup>a</sup>	4	148401190	A	G	Stage 1	0.131	-0.262	0.031	$1.08 \times 10^{-17b}$	-	CAD
						East Asian	0.297	-0.181	0.028	$6.55 \times 10^{-11}$		
						Stage 2	0.222	-0.218	0.021	$3.22 \times 10^{-26}$		
rs4705938	5q31.1	5	131694077	T	C	Stage 1	0.549	0.120	0.019	$2.55 \times 10^{-10}$	SLC22A5/ SLC22A4/ P4HA2	Lung function
						East Asian	N/A	N/A	N/A	N/A <sup>c</sup>		
						Stage 2	0.549	0.120	0.019	$2.55 \times 10^{-10}$		
rs11153071	6q16.1	6	97039741	A	G	Stage 1	0.185	0.158	0.032	$5.86 \times 10^{-7b}$	-	SBP, migraine, sleep quality
						East Asian	0.113	0.143	0.041	$5.29 \times 10^{-4}$		
						Stage 2	0.158	0.153	0.025	$1.25 \times 10^{-9}$		
rs62516550	8q11.23 <sup>a</sup>	8	55467028	T	C	Stage 1	0.389	0.169	0.023	$1.44 \times 10^{-13b}$	SOX17	-
						East Asian	0.087	0.102	0.049	$3.70 \times 10^{-2}$		
						Stage 2	0.335	0.157	0.021	$3.44 \times 10^{-14}$		
rs1537373	9p21.3 <sup>a</sup>	9	22103341	T	G	Stage 1	0.514	-0.186	0.019	$2.60 \times 10^{-22}$	-	IS, AAA, CAD
						East Asian	0.342	-0.165	0.029	$1.43 \times 10^{-8}$		
						Stage 2	0.462	-0.180	0.016	$2.86 \times 10^{-29}$		
rs11187838	10q23.33	10	96038686	A	G	Stage 1	0.415	-0.075	0.019	$1.24 \times 10^{-4}$	-	SBP, migraine, fat free mass
						East Asian	0.473	-0.108	0.025	$1.81 \times 10^{-5}$		
						Stage 2	0.436	-0.087	0.015	$1.55 \times 10^{-8}$		
rs79780963	10q24.32 <sup>a</sup>	10	104952499	T	C	Stage 1	0.078	-0.225	0.039	$6.82 \times 10^{-9}$	NT5C2/ MARCKSL1P1	-
						East Asian	0.371	-0.163	0.032	$3.11 \times 10^{-7}$		
						Stage 2	0.254	-0.188	0.025	$2.34 \times 10^{-14}$		
rs2280543	11p15.5	11	203788	T	C	Stage 1	0.041	0.162	0.053	$2.19 \times 10^{-3}$	-	-
						East Asian	0.131	0.277	0.038	$2.87 \times 10^{-13}$		
						Stage 2	0.101	0.238	0.031	$1.16 \times 10^{-14}$		
rs11044991	12p12.2	12	20174364	A	G	Stage 1	0.038	-0.142	0.053	$7.47 \times 10^{-3}$	-	Mean arterial pressure
						East Asian	0.476	-0.125	0.025	$6.74 \times 10^{-7}$		
						Stage 2	0.395	-0.128	0.023	$1.74 \times 10^{-8}$		
rs2681472	12q21.33	12	90008959	A	G	Stage 1	0.844	0.086	0.029	$2.86 \times 10^{-3}$	-	SBP, DBP, pulse pressure, CVD, CAD
						East Asian	0.629	0.131	0.026	$5.29 \times 10^{-7}$		
						Stage 2	0.719	0.116	0.020	$6.71 \times 10^{-9}$		
rs7137731	12q22	12	95490999	T	C	Stage 1	0.647	-0.138	0.020	$3.31 \times 10^{-12b}$	FGD6/NR2C1	-
						East Asian	0.640	-0.086	0.026	$1.01 \times 10^{-3}$		
						Stage 2	0.644	-0.119	0.016	$4.88 \times 10^{-14}$		
rs3742321	13q13.1 <sup>a</sup>	13	33704065	T	C	Stage 1	0.764	-0.148	0.022	$4.10 \times 10^{-11}$	-	-
						East Asian	0.756	-0.135	0.032	$2.71 \times 10^{-5}$		
						Stage 2	0.762	-0.144	0.018	$5.47 \times 10^{-15}$		
rs8034191	15q25.1	15	78806023	T	C	Stage 1	0.659	-0.115	0.022	$1.22 \times 10^{-7b}$	PSMA4	Smoking behavior, lung function, COPD
						East Asian	0.976	-0.161	0.091	$7.69 \times 10^{-2}$		
						Stage 2	0.676	-0.117	0.021	$2.75 \times 10^{-8}$		
rs7184525	16q23.1	16	75437186	A	G	Stage 1	0.450	0.148	0.023	$8.80 \times 10^{-11b}$	BCAR1/ RPT1-252K23.2	-
						East Asian	0.459	0.123	0.028	$1.04 \times 10^{-5}$		
						Stage 2	0.453	0.138	0.018	$5.60 \times 10^{-15}$		
rs11661542	18q11.2 <sup>a</sup>	18	20223695	A	C	Stage 1	0.516	-0.166	0.021	$5.74 \times 10^{-16}$	-	-
						East Asian	0.401	-0.087	0.026	$6.82 \times 10^{-4}$		
						Stage 2	0.471	-0.135	0.016	$3.17 \times 10^{-17}$		

Continued

**Table 1 | Lead associations of genome-wide-significant risk loci (Continued)**

SNP	Locus	Chromosome	Position	EA	OA	Stage	EAF	$\beta$	s.e.	P value	Annotated genes	Associated traits
rs4814863	20p11.23	20	19469685	A	G	Stage 1	0.248	0.096	0.024	$6.71 \times 10^{-5}$	-	-
						East Asian	0.513	0.110	0.025	$1.10 \times 10^{-5}$	-	-
						Stage 2	0.375	0.103	0.017	$3.22 \times 10^{-9}$	-	-
rs39713	22q12.1	22	30343186	T	C	Stage 1	0.088	0.182	0.033	$4.10 \times 10^{-8}$	-	-
						East Asian	N/A	N/A	N/A	N/A <sup>c</sup>	-	-
						Stage 2	0.088	0.182	0.033	$4.10 \times 10^{-8}$	-	-

Association statistics were derived using the SAIGE logistic mixed model. P values are unadjusted from a two-sided test. Risk loci reaching a genome-wide-significance threshold ( $P < 5 \times 10^{-8}$ ) in the stage 2 GWAS of European and East Asian ancestry individuals are shown. Position, base-pair position on GRCh37; EA, effect allele; OA, other allele; Stage 1, European ancestry only GWAS meta-analysis; East Asian, subset of samples from Japan and China; Stage 2, meta-analysis of European ancestry and East Asian data; EAF, effect allele frequency; s.e., s.e. of  $\beta$ ; N/A, not available. Annotated genes are potentially causative genes identified using summary SMR, eCAVIAR and TWAS. Associated traits are cardiovascular traits and stroke risk factors with which the lead SNP is associated. CAD, coronary artery disease; COPD, chronic obstructive pulmonary disease; CVD, cardiovascular disease. \*Known locus, described in Hussain et al.<sup>31</sup>; <sup>b</sup>Another SNP in this locus ( $r^2 > 0.8$  with the stage 2 lead SNP) has a lower P value due to differences in LD patterns between European and East Asian populations. For locus 15q25.1, another SNP in that locus reaches genome-wide significance in stage 1. <sup>c</sup>For two SNPs, no East Asian association statistics could be obtained because these SNPs are monomorphic in Japanese and Chinese populations (LDlink, <https://ldlink.nci.nih.gov>).

Conditioning the stage 1 GWAS summary statistics on GWAS summary statistics for systolic and diastolic blood pressure (BP, Neale lab summary statistics, <http://www.nealelab.is/blog/2017/7/19/rapid-gwas-of-thousands-of-phenotypes-for-337000-samples-in-the-uk-biobank>) using multi-trait conditional and joint (mtCOJO)<sup>16</sup> analysis resulted in one additional genome-wide-significant locus (rs2616406,  $P = 6.22 \times 10^{-8}$  in the stage 1 GWAS,  $P = 4.50 \times 10^{-9}$  after mtCOJO with BP). The mtCOJO analysis with smoking pack-years summary statistics or including genetic risk scores (GRSs) for smoking (cigarettes per day)<sup>17</sup> or BP-related traits<sup>18</sup> as covariates did not result in additional loci (data not shown).

**Characterization of GWAS loci.** An overview of the genic position, alleles, effect size and P value of the strongest association per locus is shown in Table 1. We used summary statistics-based Mendelian randomization (SMR), co-localization analysis using eCAVIAR and transcriptome-wide association study (TWAS, <http://gusevlab.org/projects/fusion>) to annotate potential causative genes in these loci (Supplementary Tables 5–9 and Supplementary Fig. 2). A description of this annotation process is described in the Supplementary Note. As SMR, eCAVIAR and TWAS all require linkage disequilibrium (LD) reference panels, we limited the annotation to the loci identified in the European ancestry stage 1 GWAS meta-analysis. This resulted in ten potential causative genes at five unique loci: *SLC22A5/SLC22A4/P4HA2* (chromosome (chr) 5), *NT5C2/MARCKSLP1* (chr10), *FGD6/NR2C1* (chr12), *PSMA4* (chr15) and *BCAR1/RP11-252K23.2* (chr16) (Table 1 and Supplementary Table 5). Although we did not find evidence for involvement of *SOX17* in the chr8 locus, previous studies did find functional evidence for *SOX17* (refs. <sup>19,20</sup>). Therefore, we annotated the chr8 locus as *SOX17*, totaling 11 genes in 6 loci.

In the stage 2 GWAS, six additional loci were identified: 6q16.1, 10q23.33, 11p15.5, 12p12.2, 12q21.22 and 20p11.23. Due to the combined European and East Asian LD structures, these loci cannot reliably be mapped to genes using the above-mentioned techniques. Of the six additional loci, four have previously been linked to BP, namely 6q16.1 (rs11153071)<sup>31</sup>, 10q23.33 (rs11187838)<sup>22</sup>, rs11044991 (12p12.2)<sup>23</sup> and rs2681492 (12q21.22)<sup>23,24</sup>. A detailed description of the genes and loci is found in the Supplementary Note.

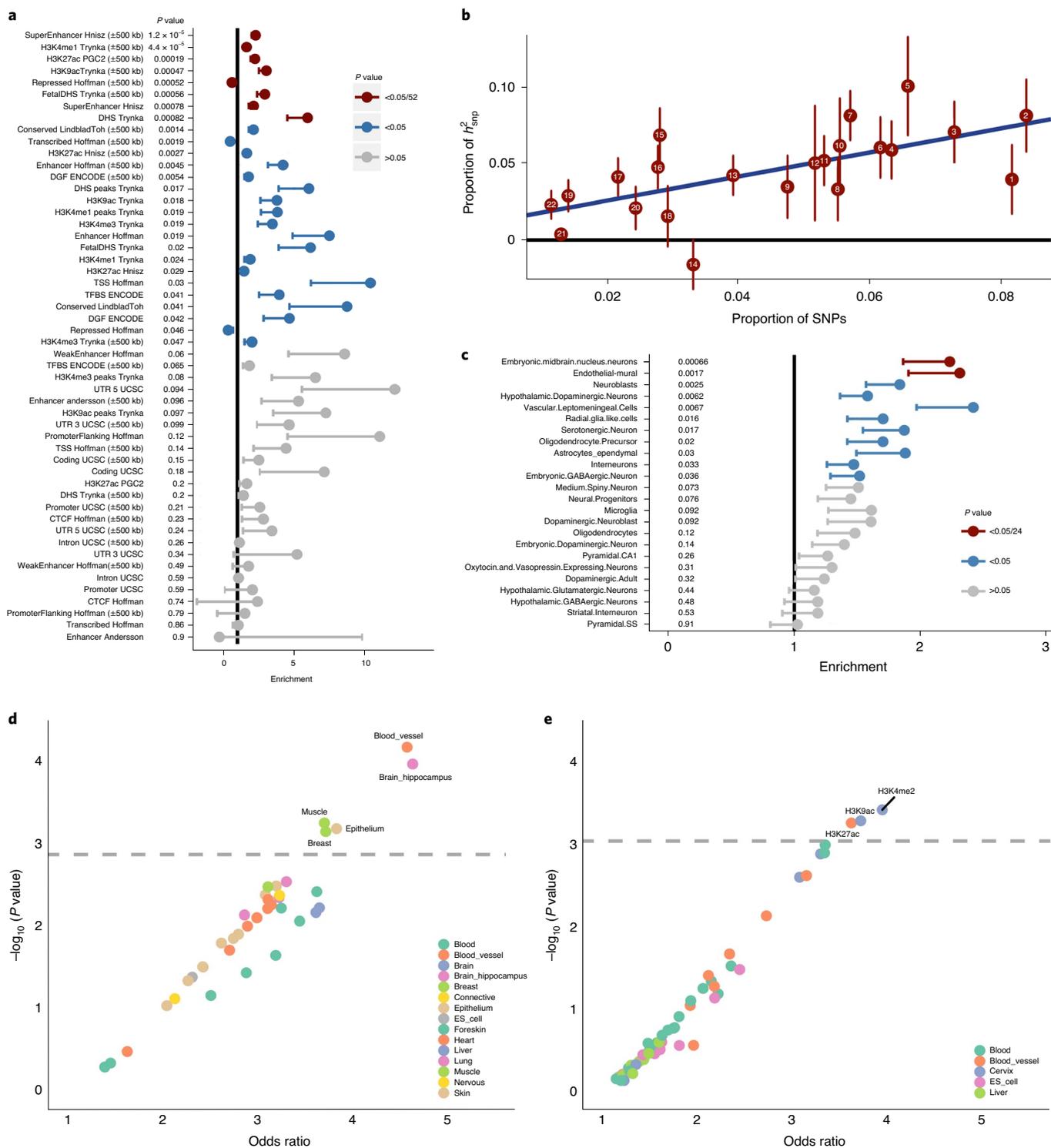
The product of the potentially causative gene *FGD6* (ref. <sup>25</sup>) plays a role in angiogenesis, and defects may lead to a compromised formation of blood vessels. *FGD6* is a vascular endothelial cell (vEC)-signaling gene involved in stress signaling in vECs<sup>26</sup>. Loss-of-function mutations in *THSD1* and *SOX17* lead to subarachnoid hemorrhage in animal models. Products of both these genes have key roles in vECs<sup>7,19,27</sup>. *BCAR1* is a ubiquitously expressed gene with a protein product that is a sensor for mechanical stress<sup>28</sup>.

The *PSMA4* locus is known for associations with a number of smoking and respiratory system traits<sup>29–32</sup>.

**Predictors of intracranial aneurysm rupture.** We assessed whether genetic risk factors differed between ruptured and unruptured intracranial aneurysms using stratified GWAS analysis. The number of cases with an unruptured intracranial aneurysm was small ( $n = 2,070$ ). Therefore, in addition to performing a stratified GWAS on patients with a ruptured aneurysm versus patients with an unruptured intracranial aneurysm (aSAH-versus-uIA), we also performed a stratified GWAS on just the patients with ruptured intracranial aneurysm versus controls (aSAH-only) and a stratified GWAS on just the patients with an unruptured intracranial aneurysm versus controls (uIA-only) (Supplementary Table 4 and Supplementary Fig. 1e–j). Overall, 69% of intracranial aneurysm cases had a ruptured intracranial aneurysm and 28% an unruptured intracranial aneurysm, whereas 3.8% had an unknown rupture status. The aSAH-only and uIA-only GWAS identified a number of genome-wide-significant loci, all of which reached genome-wide significance in the stages 1 and 2 GWAS meta-analyses of intracranial aneurysms. In the aSAH-versus-uIA GWAS, we found no genome-wide-significant loci. Furthermore, genetic correlation analysis showed a high correlation of  $0.970 \pm 0.133$  (s.e.) between ruptured and unruptured intracranial aneurysms (Supplementary Table 3). Together these findings indicate a strong similarity in genetic architecture between ruptured and unruptured intracranial aneurysms.

**SNP-based heritability.** We estimated the SNP-based heritability of intracranial aneurysms to be  $21.6 \pm 2.8\%$  (s.e.) on the liability scale with LDSR (tool named LDSC<sup>33</sup>, <https://github.com/bulik/ldsc>) and  $29.9 \pm 5.4\%$  using SumHer<sup>34</sup> (<http://dougsspeed.com/sumher>) (Table 2). This corresponds to an explained fraction of the twin-based heritability ( $h^2 = 41\%$ )<sup>3</sup> of 53–73% depending on the method used (LDSC or SumHer). We used a prevalence for unruptured intracranial aneurysms of 3%<sup>1</sup> for the conversion to the liability scale. As this GWAS was an admixture of patients with ruptured and unruptured intracranial aneurysms, this prevalence may not be representative of the whole study population. Therefore, we calculated a liability scale heritability using a range of prevalence values (Supplementary Fig. 3a). This showed that, when using lower prevalence estimates ( $K$ ), the explained SNP-based heritability was substantial ( $K = 0.02$ :  $h^2 = 19.3 \pm 2.5\%$  (LDSC),  $26.8 \pm 4.8\%$  (SumHer);  $K = 0.01$ :  $16.3 \pm 2.1\%$  (LDSC),  $22.6 \pm 4.1\%$  (SumHer)).

A substantial SNP-based heritability is also found for ruptured intracranial aneurysms (SAH-only,  $h^2 = 0.140 \pm 0.020$ ) and unruptured intracranial aneurysms (uIA-only,  $h^2 = 0.223 \pm 0.044$ ).



**Fig. 2 | Heritability and functional enrichment analyses.** **a**, Partitioned LDSCR enrichment of regulatory elements. Labels indicate type of regulatory element or histone mark. On the horizontal axis, the enrichment is shown. Enrichment = 1 indicates no enrichment. Statistical significance was defined as  $P < 0.05$  divided by the number of annotations (52). Effective  $n$  varies per SNP (see Methods). Points are estimates and error bars denote 1 s.e. in the direction of no effect. Statistics were derived from two-sided, weighted linear regression. There was no  $P$ -value adjustment. **b**, Partitioned LDSCR heritability analysis per chromosome. On the horizontal axis, the proportion of SNPs on each chromosome is shown, and on the vertical axis, the proportion of SNP-based heritability. The linear regression line is shown in blue. Data are presented as point estimates  $\pm$  s.e. Statistics are the same as used for **a**. **c**, Partitioned LDSCR enrichment analysis of scRNA-seq of brain cell types. Format and statistics are the same as used for **a**. **d**, GARFIELD analysis of tissues. On the horizontal axis, the enrichment of annotations is shown; on the vertical axis, the corresponding  $-\log_{10}(P\text{ value})$  is shown. The dashed line indicates the significance threshold of  $P = 0.05$  divided by the number of annotations. Odds ratios are derived by logistic regression.  $P$  values are unadjusted, derived from two-sided tests. **e**, GARFIELD analysis of regulatory regions defined by histone modifications. Format and statistics are the same as used for **d**.

**Table 2 | SNP heritability estimates**

Trait	Method	$h^2_{\text{obs}}$	s.e. ( $h^2_{\text{obs}}$ )	Prevalence	$h^2_{\text{liab}}$	s.e. ( $h^2_{\text{liab}}$ )	Cases	Controls	$n_{\text{eff}}$
Intracranial aneurysms (stage 1)	LDSC	0.295	0.038	0.03	0.216	0.028	7,495	71,934	24,253
Intracranial aneurysm (stage 1)	SumHer	0.409	0.074	0.03	0.299	0.054	7,495	71,934	24,253
Intracranial aneurysm (stage 1)	SumHer (LDSC)	0.276	0.037	0.03	0.202	0.027	7,495	71,934	24,253
aSAH-only	LDSC	0.296	0.043	0.005	0.140	0.020	5,140	71,952	17,019
uIA-only	LDSC	0.393	0.075	0.03	0.223	0.044	2,070	71,952	7,721

Values are given on the observed scale ( $h^2_{\text{obs}}$ ) and liability scale ( $h^2_{\text{liab}}$ ). Prevalence used for conversion to the liability scale is shown. Effective number samples were used for the conversion, as described in the Supplementary Note. For SumHer, two analyses were done: one with settings suggested by the SumHer authors, using LD reference data from the HRS, and one to mimic LDSC, with the same settings and reference panel (HapMap3, hm3).  $n_{\text{eff}}$ , effective sample size.

The difference between the heritability estimates could suggest differences in genetic architecture, but estimates depend on the prevalence estimate (Supplementary Fig. 3b,c), meaning that these differences should be interpreted with caution.

**Enrichment of genomic regions.** To understand the disease mechanisms of intracranial aneurysms, we applied several heritability enrichment analyses using LDSR. Partitioning on functional genomic elements showed a clear enrichment of heritability in regulatory elements, including enhancer and promoter histone marks H3K4me1, H3K27Ac and H3K9Ac, super-enhancers and DNase I hypersensitivity sites (Fig. 2a). Such enrichment of regulatory elements in the genome is also seen in other polygenic traits and indicates that the architecture of intracranial aneurysms is polygenic<sup>35</sup>. Partitioning heritability per chromosome further supported a polygenic architecture because heritability was associated with the number of SNPs on a chromosome (Fig. 2b).

Tissue-specific LDSR did not show enrichment for any tissue (Supplementary Tables 10 and 11). We then performed cell-type enrichment analysis using single-cell RNA-sequencing (scRNA-seq) reference data derived from mouse brains<sup>36</sup>. No enrichment was found using a scRNA-seq dataset of mouse brain blood vessels<sup>37</sup> (Supplementary Table 12). Using a larger dataset defining cell types in the mouse brain<sup>36</sup>, we found enrichment in ‘endothelial mural cells’, which is a combined set of vECs and mural cells (enrichment =  $2.31 \pm 0.41$  (s.e.),  $P = 1.65 \times 10^{-3}$ , Fig. 2c), and in midbrain neurons (enrichment =  $2.23 \pm 0.37$ ,  $P = 6.56 \times 10^{-4}$ ).

LD-pruned enrichment analysis using GARFIELD showed that genes specific for blood vessels were enriched (Fig. 2d and Supplementary Table 13) and further supports the role of promoters and enhancers (Fig. 2e).

**Causal genetic roles of BP and smoking.** To assess which phenotypes causally influence the risk of intracranial aneurysms, we performed generalized SMR (GSMR) using summary statistics for all phenotypes available in the UK Biobank (Supplementary Table 14). We used the stage 1 summary statistics excluding the UK Biobank data as an outcome. In this analysis, we chose a stringent value for the multiple testing threshold of 376, which was the number of traits passing the GSMR quality control parameters. After correction for multiple testing 16 traits were statistically significant (Fig. 3a). All statistically significant traits were related to either smoking or BP, which are the two main clinical risk factors for unruptured intracranial aneurysms and aSAHs<sup>1,38,39</sup>. To determine whether genetic predisposition for smoking and BP was a causal genetic risk factor for both, independent of each other, we conditioned the stage 1 GWAS summary statistics on GWAS summary statistics for smoking and BP using mtCOJO analysis. We used summary statistics for both systolic BP (SBP) and diastolic BP (DBP) combined to condition on BP and summary statistics for pack-years to condition on smoking (Fig. 3a and Supplementary Table 14). All GSMR effects diminished

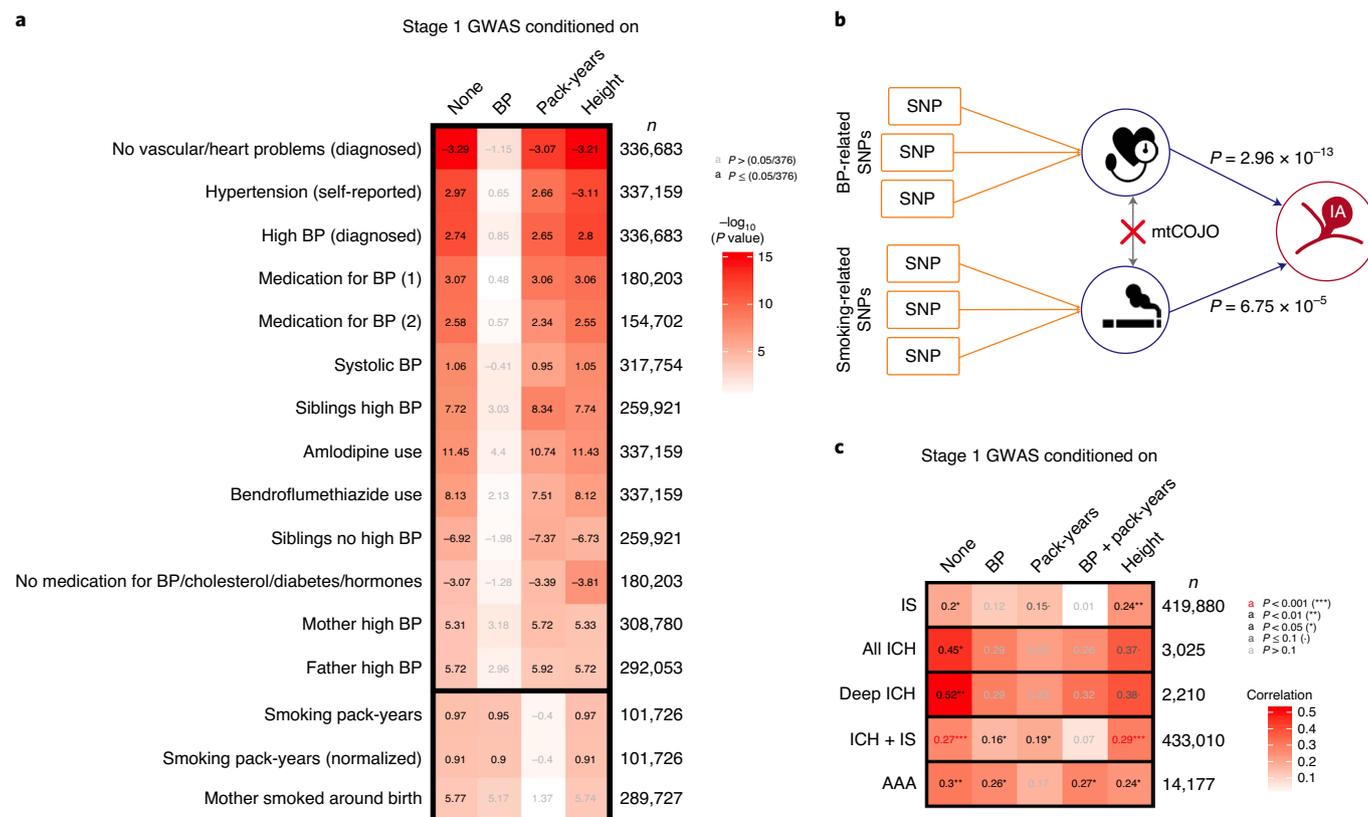
after conditioning on either BP or pack-years and remained when conditioning on the other risk factor. The mtCOJO analysis method itself did not affect the effect size estimates because conditioning on standing height did not affect the estimates. These findings provide strong evidence that the genetic predisposition for BP and smoking are independent genetic causes of intracranial aneurysms (Fig. 3b).

As the phenotype values of the exposure traits were inverse rank-normalized, the GSMR effect size of SBP (effect of exposure  $x$  on outcome  $y$ ,  $\beta_{xy} = 1.058 \pm 0.187$ ) and pack-years ( $\beta_{xy} = 0.973 \pm 0.236$ ) cannot easily be interpreted. Therefore, we performed an additional GSMR analysis for BP with an updated version of the UK Biobank GWAS (<http://www.nealelab.is/uk-biobank>), including raw phenotype values for quantitative traits (Supplementary Table 15). For BP traits, the GSMR analysis resulted in an effect size estimate of  $0.095 \pm 0.019$  for DBP and  $0.047 \pm 0.011$  for SBP, meaning an 8–12% increase in intracranial aneurysm risk per mmHg increase of DBP and a 3.7–6.0% increase in intracranial aneurysm risk per mmHg increase of SBP, assuming a linear effect of BP on intracranial aneurysm liability. In addition, age at high BP diagnosis had a significant GSMR effect ( $P = 1.79 \times 10^{-4}$ ,  $\beta_{xy} = 0.163 \pm 0.044$ ), indicating an increase in intracranial aneurysm risk of 13–23% for each year of additional high BP exposure. We did not include smoking quantitative traits because these were not normally distributed (data not shown) and could, therefore, lead to a biased effect estimate.

We then tested whether the effects of smoking and BP were different between ruptured (SAH-only) and unruptured intracranial aneurysms (uIA-only; Supplementary Table 16). The GSMR effect sizes followed the same trend for all phenotypes, but ‘hypertension (self-reported)’ had a stronger effect on ruptured intracranial aneurysms (SAH-only:  $\beta_{xy} = 6.74 \pm 0.61$  (s.e.), all intracranial aneurysms:  $2.97 \pm 0.42$  and uIA-only:  $2.38 \pm 0.70$ ), whereas amlodipine use had a weaker effect on unruptured intracranial aneurysms and became statistically nonsignificant (uIA-only:  $\beta_{xy} = 4.77 \pm 3.90$ ,  $P = 0.22$ , all intracranial aneurysms:  $\beta_{xy} = 11.4 \pm 2.10$ ,  $P = 5.25 \times 10^{-8}$ , and SAH-only:  $\beta_{xy} = 13.1 \pm 2.60$ ,  $P = 5.25 \times 10^{-7}$ ). Although the effect of self-reported hypertension on SAH-only was stronger, conditioning on BP using mtCOJO analysis mitigated the effect ( $\beta_{xy} = 1.02 \pm 0.45$ ,  $P = 0.024$ , data not shown). As the power to detect GSMR effects in the uIA-only sample is much lower compared with all intracranial aneurysms and SAH-only due to limited sample size, further investigation is required to make inferences about genetic risk factors for rupture.

Traits influencing female hormones are suggested to play a role in aSAH risk<sup>40</sup>. Only two female hormone-related traits had enough genome-wide-significant risk loci to pass GSMR QC. These were ‘age when periods started (menarche)’ and ‘had menopause’. Neither of these showed a causal relationship with intracranial aneurysms in the GSMR analysis (Supplementary Table 14).

**Drivers of genetic correlation with vascular traits.** To identify traits correlated with intracranial aneurysms, we analyzed stage



**Fig. 3 | Cross-trait analyses.** **a**, GSNR analysis of UK Biobank predictors on the stage 1 intracranial aneurysm GWAS, conditioned on traits depicted by column labels with mtCOJO analysis. Numeric values are the GSNR effect sizes. The top 13 traits are BP-related traits. The bottom three traits are smoking related. Statistical significance was defined as  $P < 0.05$  divided by the number of traits that passed QC (376). Square fill colors indicate  $-\log_{10}(P \text{ value})$  of the GSNR effect. All 16 traits that pass the multiple testing threshold for significance in the unconditioned analysis are shown. Presented *n* is the sample size in the UK Biobank GWAS. For intracranial aneurysms, effective *n* per SNP was used. *P* values are from two-sided linear regressions, unadjusted. **b**, Causality diagram further explaining the analyses of **a**: GSNR analysis showed that genetic risks for smoking and BP cause intracranial aneurysms. Using mtCOJO analysis, it was found that the genetic factors associated with BP and smoking cause intracranial aneurysms through independent mechanisms. Statistics are the same as used for **a**. BP,  $n = 317,754$  samples; smoking,  $n = 101,726$  samples. **c**, Genetic correlation analysis with LDSR. Genetic correlation estimates are indicated by color and numeric value. Axis labels on the left denote the trait correlated with intracranial aneurysms. Labels on the top denote the trait for which the stage 1 intracranial aneurysm GWAS was conditioned using mtCOJO analysis. More details are provided in Supplementary Table 3. Presented *n* is an effective sample size for trait on the left, except for IS and ICH + IS, where an *n* per SNP was used and average *n* is shown.

1 summary statistics using LDHub<sup>41</sup>. LDHub includes a subset of the summary statistics used for GSNR and a number of summary statistics from publicly available sources. Traits that showed correlations that reached Bonferroni's threshold for multiple testing ( $P = 0.05/464$ ) included several BP-related traits, including DBP ( $\rho_g = 0.223$ ,  $P = 5.40 \times 10^{-9}$ ) and SBP ( $\rho_g = 0.256$ ,  $P = 1.34 \times 10^{-8}$ ) and smoking traits, such as pack-years ( $\rho_g = 0.330$ ,  $P = 7.87 \times 10^{-8}$ ) (Supplementary Table 17).

We used LDSR to calculate the genetic correlation of intracranial aneurysms with other stroke subtypes (ischemic stroke (IS)<sup>42</sup> and intracerebral hemorrhage (ICH)), with other vascular malformation types (intracranial arteriovenous malformation (AVM)<sup>43</sup> and cervical artery dissection<sup>44</sup>) and with abdominal aortic aneurysm (AAA)<sup>45</sup>. For IS, a correlation of  $0.195 \pm 0.079$  ( $P = 0.014$ ) was found with intracranial aneurysms (Fig. 3c and Supplementary Table 3). After conditioning the intracranial aneurysm GWAS on either BP or pack-years, which are clinical risk factors for both IS and intracranial aneurysms<sup>1,38,39,46</sup>, the correlation was no longer statistically significant and reduced to  $0.121 \pm 0.081$  for BP and  $0.147 \pm 0.084$  for pack-years. The correlation disappeared after conditioning on both risk factors ( $\rho_g = 0.009 \pm 0.083$ ,  $P = 0.916$ ). When conditioning on an unrelated but heritable trait (standing height), the correlation

remained ( $\rho_g = 0.238 \pm 0.081$ ,  $P = 0.003$ ). No genetic correlation was found for any of the IS subtypes.

We found a statistically significant genetic correlation between intracranial aneurysms and ICH ( $\rho_g = 0.447 \pm 0.184$ ,  $P = 0.015$ ), which was mainly driven by deep ICH ( $\rho_g = 0.516 \pm 0.198$ ,  $P = 0.009$ ), and not by lobar ICH ( $P = 0.534$ ). After conditioning the intracranial aneurysm GWAS on either BP or pack-years, which are also important risk factors for ICH<sup>47</sup>, the correlation with deep ICH decreased ( $\rho_g = 0.288 \pm 0.189$  for BP and  $0.234 \pm 0.192$  for pack-years) and was no longer statistically significant. Conditioning on height had a much smaller effect ( $\rho_g = 0.380 \pm 0.196$ ).

A genetic correlation was found between intracranial aneurysms and AAAs ( $\rho_g = 0.302 \pm 0.105$ ,  $P = 0.004$ ). Conditioning on pack-years strongly reduced the correlation between intracranial aneurysms and AAAs ( $\rho_g = 0.173 \pm 0.117$ ,  $P = 0.138$ ), whereas BP did not ( $\rho_g = 0.264 \pm 0.117$ ,  $P = 0.024$ ).

There was no genetic correlation between intracranial aneurysms and carotid artery dissection ( $\rho_g = 0.151 \pm 0.180$ ,  $P = 0.401$ ), whereas, for vertebral artery dissection and the combined set of vertebral and carotid artery dissection, a larger, albeit nonstatistically significant, estimate was observed ( $\rho_g = 0.281 \pm 0.159$ ,  $P = 0.077$  and  $\rho_g = 0.174 \pm 0.149$ ,  $P = 0.066$ , respectively) (Supplementary Table 3).

For AVMs, a negative SNP-based heritability was estimated, which could be due to the small sample size of this GWAS (1,123 cases and 1,935 controls). Therefore, we performed a lookup of all SNPs identified in the stage 1 and 2 intracranial aneurysm GWAS in the summary statistics of the AVM GWAS<sup>43</sup>, but were unable to replicate any of these SNP associations ( $P < 0.05/17$ ) (Supplementary Table 18).

**Drug-target enrichment.** To identify pleiotropic pathways between intracranial aneurysms and other diseases that contain known drug targets, we assessed enrichment in genes targeted by drugs and drug classes<sup>48</sup>. Gene-based  $P$  values were calculated using MAGMA, resulting in 29 genes that passed Bonferroni's threshold for multiple testing ( $P < 0.05/18,106$ , Supplementary Table 19). The antihypertensive drugs ambrisentan and macitentan showed a statistically significant enrichment ( $P = 1.35 \times 10^{-5}$ , Supplementary Table 20), which was driven by a single gene (*EDNRA*). Drug-class enrichment analysis showed that drugs in the class 'antiepileptics' were enriched (area under the curve (AUC) = 0.675,  $P = 8 \times 10^{-5}$ ; Supplementary Table 21). The most statistically significant enriched drugs within this class are blockers of  $\text{Na}^+$  and  $\text{Ca}^{2+}$  channels, namely phenytoin, zonisamide and topiramate<sup>49</sup> (Supplementary Table 20). These channels are important in BP regulation, as well as in several other biological mechanisms. The other enriched drug class is 'sex hormones and modulators of the genital system' (AUC = 0.652,  $P = 2.02 \times 10^{-4}$ ). We also used MAGMA to study enrichment in gene pathways but found no statistically significant results (Supplementary Table 22).

## Discussion

We identified 11 new risk loci for intracranial aneurysms and confirmed 6 previously identified risk loci, yielding a total of 17 risk loci for intracranial aneurysms. A SNP-based heritability of 21.6% was found, explaining over half of the total heritability. We showed strong evidence that the majority of intracranial aneurysm heritability is polygenic. Our results further highlight several major features of the genetic architecture of intracranial aneurysms. First, we identified ECs as a key cell type in intracranial aneurysm risk. Second, we showed that, of 375 tested traits, smoking and BP predisposition were the main genetic risk factors for intracranial aneurysms. Third, we showed that the main drivers of the genetic correlation between intracranial aneurysms and other stroke types, and between intracranial aneurysms and AAAs, are genetic predisposition for smoking and BP. Last, we found pleiotropic characteristics of antiepileptic drugs and sex hormones with intracranial aneurysms.

Through gene mapping incorporating gene expression datasets and distinct bioinformatics analyses, we were able to identify 11 potential causative genes within 6 of the stage 1 risk loci. Many of these genes have known or putative roles in blood vessel function and BP regulation. We found heritability enrichment in genes that are specifically expressed in a combined set of ECs and mural cells, and not in other vascular cell types. Together, the identified potential causative genes and heritability enrichment analyses suggest an important role of the vEC in intracranial aneurysm development and rupture.

Through genetic correlation and formal causal inference methods, we established that genetic predisposition for smoking and BP are the most important independent genetic risk factors for intracranial aneurysms<sup>1</sup>. First, using causal inference with GSMR, we showed that genetic predisposition for these traits drives a causal increase in intracranial aneurysm risk. Then, using mtCOJO analysis, we showed that smoking and high BP are causative of intracranial aneurysms, independent of each other. By using nontransformed continuous SBP and DBP measures in the UK Biobank, we estimated the increase in intracranial aneurysm risk per 1 mmHg increase of SBP to be 3.7–6.0%, and that of DBP to be 8–12%. These strong effects provide genetic evidence for clinical

prevention by lowering BP. As smoking dose is not normally distributed, we were not able to estimate a quantitative effect of smoking on intracranial aneurysms, but this has been done before using nongenetic methods<sup>50–52</sup>. Future studies that model risk prediction using polygenic risk scores should determine whether the polygenic risks of genetic risk factors for intracranial aneurysms are clinically relevant risk factors for the disease.

We found that genetic correlations of intracranial aneurysms with IS and ICH are mainly driven by genetic predisposition for smoking and BP. For ICH, conditioning on smoking and BP did not completely mitigate the genetic correlation with intracranial aneurysms, suggesting additional shared genetic causes. For vertebral artery dissection, a substantial but not statistically significant correlation with intracranial aneurysms was found, whereas this was absent in carotid artery dissection. We showed that the genetic correlation between intracranial aneurysms and AAAs was driven by smoking, but not by BP. This implies that intracranial aneurysms are more dependent on BP compared with AAAs. This observation could be a result of different ratios of unruptured and ruptured aneurysms included in the two GWAS. The AAA GWAS consists of mainly unruptured AAAs<sup>45</sup> and, although the role of BP on AAA rupture is clear, the effect on developing AAAs is a matter of debate<sup>53</sup>.

One of the main aims of intracranial aneurysm research is to prevent rupture of intracranial aneurysms and thus avoid the devastating consequences of an aSAH. We performed various analyses in an attempt to identify genetic predictors specific for intracranial aneurysm rupture. Instead, we found a very strong genetic correlation between ruptured and unruptured intracranial aneurysms. These analyses together indicate that the common variant genetic architecture of ruptured and unruptured aneurysms is strikingly similar.

The heritability of unruptured intracranial aneurysms has never been studied in twins and may therefore be a suboptimal estimate for intracranial aneurysm heritability. One twin study estimated the heritability of aSAH at 41%<sup>3</sup>. Our finding that the genetic architecture of uIA and aSAH is similar suggests that this heritability estimate may also be accurate for unruptured intracranial aneurysms. This means that, in European ancestry populations, 53–73% of the heritability of intracranial aneurysms can be explained by variants tagged in this GWAS.

Using transancestry genetic correlation, we found a remarkable similarity of genetic architecture between the European ancestry and East Asian ancestry GWAS of  $>90.8 \pm 14.6\%$  (s.e.). This indicates that most common-variant genetic causes are the same, regardless of ancestry. However, as the LD structures remain distinct, current methods for summary statistics-based enrichment analysis cannot effectively account for population-specific variation in a cross-ancestry GWAS.

Drug-class enrichment showed pleiotropic characteristics of antiepileptic drugs and sex hormones with the genetic associations of intracranial aneurysms. It has been suggested that sex hormones might play a role in intracranial aneurysms<sup>40</sup>, potentially explaining why women have a higher intracranial aneurysm risk than men<sup>1</sup>. However, as causal inference analysis with GSMR did not show evidence for the involvement of female hormones, further investigation is required. Enrichment of the antiepileptic drug class may indicate shared disease mechanisms between intracranial aneurysms and epilepsy. The main mechanism of antiepileptic drugs is through blocking  $\text{Na}^+$  and  $\text{Ca}^{2+}$  ion channels<sup>49</sup>. Together with other ion channels, these play essential roles in contraction and relaxation of the blood vessels<sup>54</sup>. Mutations in the ion-channel gene *PKD2* (*TRRP2*) are known to cause intracranial aneurysms. This gene product, along with other members of the *TRP* gene family, regulates systemic BP through vasoconstriction and vasodilatation<sup>55,56</sup>. More research on the effect of antiepileptics on vascular tension and BP will enhance our understanding of the disease-causing mechanisms.

Furthermore, this could help to identify methods of intracranial aneurysm prevention using antiepileptics or related drugs.

In conclusion, we performed a GWAS meta-analysis of intracranial aneurysms, identifying 11 new risk loci, confirming 6 previously identified risk loci and explaining over half of the heritability of intracranial aneurysms. We found strong evidence for a polygenic architecture. Through gene mapping and heritability enrichment methods, we discovered a possible role for ECs in intracranial aneurysm development. We showed that the genetic architecture of unruptured and ruptured aneurysms is very similar. The well-known clinical risk factors, smoking and hypertension, were identified as the main genetic drivers of intracranial aneurysms. These risk factors also explained most of the similarity to other stroke types, IS and deep ICH, which could open a window for clinical prevention. We also found pleiotropic effects between intracranial aneurysms and antiepileptic drugs, which require further investigation to understand the shared mechanisms of intracranial aneurysms and epilepsy. Our findings represent a major advance in understanding the pathogenesis of intracranial aneurysms and an important step toward the development of effective genetic risk prediction and prevention of intracranial aneurysm development and subsequent aSAH in the future.

### Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41588-020-00725-7>.

Received: 24 February 2020; Accepted: 24 September 2020;

Published online: 16 November 2020

### References

1. Vlak, M. H., Algra, A., Brandenburg, R. & Rinkel, G. J. Prevalence of unruptured intracranial aneurysms, with emphasis on sex, age, comorbidity, country, and time period: a systematic review and meta-analysis. *Lancet Neurol.* **10**, 626–636 (2011).
2. Nieuwkamp, D. J. et al. Changes in case fatality of aneurysmal subarachnoid haemorrhage over time, according to age, sex, and region: a meta-analysis. *Lancet Neurol.* **8**, 635–642 (2009).
3. Korja, M. et al. Genetic epidemiology of spontaneous subarachnoid hemorrhage: Nordic Twin Study. *Stroke* **41**, 2458–2462 (2010).
4. Kurki, M. I. et al. High risk population isolate reveals low frequency variants predisposing to intracranial aneurysms. *PLoS Genet.* **10**, e1004134 (2014).
5. Yasuno, K. et al. Common variant near the endothelin receptor type A (*EDNRA*) gene is associated with intracranial aneurysm risk. *Proc. Natl Acad. Sci. USA* **108**, 19707–19712 (2011).
6. Yan, J. et al. Genetic study of intracranial aneurysms. *Stroke* **46**, 620–626 (2015).
7. Santiago-Sim, T. et al. THSD1 (Thrombospondin Type 1 Domain Containing Protein 1) mutation in the pathogenesis of intracranial aneurysm and subarachnoid hemorrhage. *Stroke* **47**, 3005–3013 (2016).
8. Bourcier, R. et al. Rare coding variants in *ANGPTL6* are associated with familial forms of intracranial aneurysm. *Am. J. Hum. Genet.* **102**, 133–141 (2018).
9. Lorenzo-Betancor, O. et al. PCNT point mutations and familial intracranial aneurysms. *Neurology* **91**, e2170–e2181 (2018).
10. Zhou, S. et al. RNF213 is associated with intracranial aneurysms in the French-Canadian population. *Am. J. Hum. Genet.* **99**, 1072–1085 (2016).
11. Hussain, I., Duffis, E. J., Gandhi, C. D. & Prestigiacomo, C. J. Genome-wide association studies of intracranial aneurysms: an update. *Stroke* **44**, 2670–2675 (2013).
12. Foroud, T. et al. Genome-wide association study of intracranial aneurysms confirms role of *Anril* and *SOX17* in disease risk. *Stroke* **43**, 2846–2852 (2012).
13. Yasuno, K. et al. Genome-wide association study of intracranial aneurysm identifies three new risk loci. *Nat. Genet.* **42**, 420–425 (2010).
14. Zhou, W. et al. Efficiently controlling for case-control imbalance and sample relatedness in large-scale genetic association studies. *Nat. Genet.* **50**, 1335–1341 (2018).
15. Yang, J. et al. Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nat. Genet.* **44**, 369–375 (2012).
16. Zhu, Z. H. et al. Causal associations between risk factors and common diseases inferred from GWAS summary data. *Nat. Commun.* **9**, 224 (2018).
17. Tobacco Genetics Consortium. Genome-wide meta-analyses identify multiple loci associated with smoking behavior. *Nat. Genet.* **42**, 441–447 (2010).
18. Evangelou, E. et al. Genetic analysis of over 1 million people identifies 535 new loci associated with blood pressure traits. *Nat. Genet.* **50**, 1412–1425 (2018).
19. Lee, S. et al. Deficiency of endothelium-specific transcription factor *Sox17* induces intracranial aneurysm. *Circulation* **131**, 995–1005 (2015).
20. Laarman, M. D. et al. Chromatin conformation links putative enhancers in intracranial aneurysm-associated regions to potential candidate genes. *J. Am. Heart Assoc.* **8**, e011201 (2019).
21. Giri, A. et al. Trans-ethnic association study of blood pressure determinants in over 750,000 individuals. *Nat. Genet.* **51**, 51–62 (2019).
22. Kichaev, G. et al. Leveraging polygenic functional enrichment to improve GWAS power. *Am. J. Hum. Genet.* **104**, 65–75 (2019).
23. Takeuchi, F. et al. Interethnic analyses of blood pressure loci in populations of East Asian and European descent. *Nat. Commun.* **9**, 5052 (2018).
24. Hoffmann, T. J. et al. Genome-wide association analyses using electronic health records identify new loci influencing blood pressure variation. *Nat. Genet.* **49**, 54–64 (2017).
25. Huang, L. et al. A missense variant in *FGD6* confers increased risk of polypoidal choroidal vasculopathy. *Nat. Genet.* **48**, 640–647 (2016).
26. Romanoski, C. E. et al. Systems genetics analysis of gene-by-environment interactions in human cells. *Am. J. Hum. Genet.* **86**, 399–410 (2010).
27. Haasdijk, R. A. et al. THSD1 preserves vascular integrity and protects against intraplaque haemorrhaging in *ApoE*<sup>-/-</sup> mice. *Cardiovasc. Res.* **110**, 129–139 (2016).
28. Camacho Leal Mdel, P. et al. p130Cas/BCAR1 scaffold protein in tissue homeostasis and pathogenesis. *Gene* **562**, 1–7 (2015).
29. Nedeljkovic, I. et al. Understanding the role of the chromosome 15q25.1 in COPD through epigenetics and transcriptomics. *Eur. J. Hum. Genet.* **26**, 709–722 (2018).
30. David, S. P. et al. Genome-wide meta-analyses of smoking behaviors in African Americans. *Transl. Psychiatry* **2**, e119 (2012).
31. Liu, M. et al. Association studies of up to 1.2 million individuals yield new insights into the genetic etiology of tobacco and alcohol use. *Nat. Genet.* **51**, 237–244 (2019).
32. Lutz, S. M. et al. A genome-wide association study identifies risk loci for spirometric measures among smokers of European and African ancestry. *BMC Genet.* **16**, 138 (2015).
33. Bulik-Sullivan, B. K. et al. LD score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* **47**, 291–295 (2015).
34. Speed, D. & Balding, D. J. SumHer better estimates the SNP heritability of complex traits from summary statistics. *Nat. Genet.* **51**, 277–284 (2019).
35. Watanabe, K. et al. A global overview of pleiotropy and genetic architecture in complex traits. *Nat. Genet.* **51**, 1339–1348 (2019).
36. Skene, N. G. et al. Genetic identification of brain cell types underlying schizophrenia. *Nat. Genet.* **50**, 825–833 (2018).
37. He, L. et al. Single-cell RNA sequencing of mouse brain and lung vascular and vessel-associated cell types. *Sci. Data* **5**, 180160 (2018).
38. Backes, D., Rinkel, G. J., Laban, K. G., Algra, A. & Vergouwen, M. D. Patient- and aneurysm-specific risk factors for intracranial aneurysm growth: a systematic review and meta-analysis. *Stroke* **47**, 951–957 (2016).
39. Muller, T. B., Vik, A., Romundstad, P. R. & Sandvei, M. S. Risk factors for unruptured intracranial aneurysms and subarachnoid hemorrhage in a prospective population-based study. *Stroke* **50**, 2952–2955 (2019).
40. Algra, A. M., Klijn, C. J., Helmerhorst, F. M., Algra, A. & Rinkel, G. J. Female risk factors for subarachnoid hemorrhage: a systematic review. *Neurology* **79**, 1230–1236 (2012).
41. Zheng, J. et al. LD Hub: a centralized database and web interface to perform LD score regression that maximizes the potential of summary level GWAS data for SNP heritability and genetic correlation analysis. *Bioinformatics* **33**, 272–279 (2017).
42. Malik, R. et al. Multiancestry genome-wide association study of 520,000 subjects identifies 32 loci associated with stroke and stroke subtypes. *Nat. Genet.* **50**, 524–537 (2018).
43. Weinsheimer, S. et al. Genome-wide association study of sporadic brain arteriovenous malformations. *J. Neurol. Neurosurg. Psychiatry* **87**, 916–923 (2016).
44. Debette, S. et al. Common variation in *PHACTR1* is associated with susceptibility to cervical artery dissection. *Nat. Genet.* **47**, 78–83 (2015).
45. Jones, G. T. et al. Meta-analysis of genome-wide association studies for abdominal aortic aneurysm identifies four new disease-specific risk loci. *Circ. Res.* **120**, 341–353 (2017).

46. Hankey, G. J. Stroke. *Lancet* **389**, 641–654 (2017).
47. An, S. J., Kim, T. J. & Yoon, B. W. Epidemiology, risk factors, and clinical features of intracerebral hemorrhage: an update. *J. Stroke* **19**, 3–10 (2017).
48. Gaspar, H. A. & Breen, G. Drug enrichment and discovery from schizophrenia genome-wide association results: an analysis and visualisation approach. *Sci. Rep.* **7**, 12460 (2017).
49. Rogawski, M. A. & Loscher, W. The neurobiology of antiepileptic drugs. *Nat. Rev. Neurosci.* **5**, 553–564 (2004).
50. Lindbohm, J. V., Kaprio, J., Jousilahti, P., Salomaa, V. & Korja, M. Sex, smoking, and risk for subarachnoid hemorrhage. *Stroke* **47**, 1975–1981 (2016).
51. Vlak, M. H., Rinkel, G. J., Greebe, P. & Algra, A. Risk of rupture of an intracranial aneurysm based on patient characteristics: a case-control study. *Stroke* **44**, 1256–1259 (2013).
52. Juvela, S., Poussa, K. & Porras, M. Factors affecting formation and growth of intracranial aneurysms: a long-term follow-up study. *Stroke* **32**, 485–491 (2001).
53. Kobeissi, E., Hibino, M., Pan, H. & Aune, D. Blood pressure, hypertension and the risk of abdominal aortic aneurysms: a systematic review and meta-analysis of cohort studies. *Eur. J. Epidemiol.* **34**, 547–555 (2019).
54. Cheng, J. et al. Ion channels and vascular diseases. *Arterioscler. Thromb. Vasc. Biol.* **39**, e146–e156 (2019).
55. Bulley, S. et al. Arterial smooth muscle cell PKD2 (TRPP1) channels regulate systemic blood pressure. *eLife* **7**, e42628 (2018).
56. Perrone, R. D., Malek, A. M. & Watnick, T. Vascular complications in autosomal dominant polycystic kidney disease. *Nat. Rev. Nephrol.* **11**, 589–598 (2015).

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

© The Author(s), under exclusive licence to Springer Nature America, Inc. 2020

<sup>1</sup>Department of Neurology and Neurosurgery, University Medical Center Utrecht Brain Center, Utrecht University, Utrecht, the Netherlands. <sup>2</sup>Department of Pathology and Immunology, Faculty of Medicine, University of Geneva, Geneva, Switzerland. <sup>3</sup>Neurosurgery Division, Department of Clinical Neurosciences, Faculty of Medicine, Geneva University Hospitals, Geneva, Switzerland. <sup>4</sup>l'institut du thorax Université de Nantes, CHU Nantes, INSERM, CNRS, Nantes, France. <sup>5</sup>CHU Nantes, Department of Neuroradiology, Nantes, France. <sup>6</sup>Stroke Research Centre, University College London Queen Square Institute of Neurology, London, UK. <sup>7</sup>Department of Neurosurgery, Klinikum rechts der Isar, Technical University Munich, Munich, Germany. <sup>8</sup>Laboratory for Statistical and Translational Genetics, RIKEN Center for Integrative Medical Sciences, Yokohama, Japan. <sup>9</sup>Department of Cancer Biology, Institute of Medical Science, The University of Tokyo, Tokyo, Japan. <sup>10</sup>Department of Ophthalmology, Graduate School of Medical Sciences, Kyushu University, Fukuoka, Japan. <sup>11</sup>Department of Ocular Pathology and Imaging Science, Graduate School of Medical Sciences, Kyushu University, Fukuoka, Japan. <sup>12</sup>Graduate School of Frontier Sciences, The University of Tokyo, Tokyo, Japan. <sup>13</sup>Laboratory of Clinical Genome Sequencing, Graduate School of Frontier Sciences, The University of Tokyo, Tokyo, Japan. <sup>14</sup>Clinical Trial Service Unit and Epidemiological Studies Unit, Nuffield Department of Population Health, University of Oxford, Oxford, UK. <sup>15</sup>Medical Research Council Population Health Research Unit, University of Oxford, Oxford, UK. <sup>16</sup>Department of Epidemiology, School of Public Health, Peking University Health Science Center, Beijing, China. <sup>17</sup>Montréal Neurological Institute and Hospital, McGill University, Montréal, QC, Canada. <sup>18</sup>Lady Davis Institute, Jewish General Hospital, McGill University, Montréal, QC, Canada. <sup>19</sup>Centre for Medical Informatics, Usher Institute, University of Edinburgh, Edinburgh, UK. <sup>20</sup>UK Biobank, Cheadle, Stockport, UK. <sup>21</sup>Neurogenetics Laboratory, The National Hospital of Neurology and Neurosurgery, London, UK. <sup>22</sup>Pediatric Radiology, Necker Hospital for Sick Children, Université Paris Descartes, Paris, France. <sup>23</sup>Department of Neuroradiology, Sainte-Anne Hospital and Université Paris Descartes, INSERM UMR, S894 Paris, France. <sup>24</sup>Department of Neuroradiology, University Hospital of Brest, Brest, France. <sup>25</sup>Department of Neuroradiology, Pitié-Salpêtrière Hospital, Paris, France. <sup>26</sup>Department of Neuroradiology, University Hospital of Rennes, Rennes, France. <sup>27</sup>Department of Research, Innovation and Education, Division of Clinical Neuroscience, Oslo University Hospital, Oslo, Norway. <sup>28</sup>K. G. Jebsen Center for Genetic Epidemiology, Department of Public Health and Nursing, Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology, Trondheim, Norway. <sup>29</sup>Research and Communication Unit for Musculoskeletal Health (FORMI), Department of Research, Innovation and Education, Division of Clinical Neuroscience, Oslo University Hospital, Oslo, Norway. <sup>30</sup>Institute of Clinical Medicine, Faculty of Medicine, University of Oslo, Oslo, Norway. <sup>31</sup>Department of Public Health and Nursing, Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology, Trondheim, Norway. <sup>32</sup>The Cancer Clinic, St Olavs Hospital, Trondheim University Hospital, Trondheim, Norway. <sup>33</sup>Department of Internal Medicine, Division of Cardiovascular Medicine, University of Michigan, Ann Arbor, MI, USA. <sup>34</sup>HUNT Research Center, Department of Public Health and Nursing, Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology, Trondheim, Norway. <sup>35</sup>Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, the Netherlands. <sup>36</sup>National Institute for Public Health and the Environment, Bilthoven, the Netherlands. <sup>37</sup>Dortmund University of Applied Science and Arts, Dortmund, Germany. <sup>38</sup>Institute for Medical Informatics, Biometry and Epidemiology (IMIBE), University Hospital Essen, Essen, Germany. <sup>39</sup>Zurich University of Applied Sciences, School of Life Sciences and Facility Management, Zurich, Switzerland. <sup>40</sup>SIB Swiss Institute of Bioinformatics, Lausanne, Switzerland. <sup>41</sup>Department of Surgery, University of Otago, Dunedin, New Zealand. <sup>42</sup>Department of Cardiovascular Sciences and National Institute for Health Research, University of Leicester, Leicester, UK. <sup>43</sup>Leicester Biomedical Research Centre, University of Leicester, Glenfield Hospital, Leicester, UK. <sup>44</sup>Department of Neurology, University of California at San Francisco, San Francisco, CA, USA. <sup>45</sup>Department of Anesthesia and Perioperative Care, Center for Cerebrovascular Research, University of California, San Francisco, CA, USA. <sup>46</sup>Department of Epidemiology and Biostatistics, University of California, San Francisco, CA, USA. <sup>47</sup>Institute for Human Genetics, University of California, San Francisco, CA, USA. <sup>48</sup>Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK. <sup>49</sup>UK National Institute for Health Research (NIHR) Biomedical Research Centre (BRC), South London and Maudsley NHS Foundation Trust, London, UK. <sup>50</sup>Division of Research, Kaiser Permanente of Northern California, Oakland, CA, USA. <sup>51</sup>Department of Neurology, Donders Institute for Brain, Cognition and Behaviour, Radboud University Medical Center, Nijmegen, the Netherlands. <sup>52</sup>Institute for Stroke and Dementia Research, University Hospital, Ludwig-Maximilians-University, Munich, Germany. <sup>53</sup>Munich Cluster for Systems Neurology (SyNergy), Munich, Germany. <sup>54</sup>Deutsches Zentrum für Neurodegenerative Erkrankungen (DZNE), Munich, Germany. <sup>55</sup>INSERM U1219 Bordeaux Population Health Research Center, University of Bordeaux, Bordeaux, France. <sup>56</sup>Department of Neurology, Institute for Neurodegenerative Disease, Bordeaux University Hospital, Bordeaux, France. <sup>57</sup>Department of Clinical Neuroscience at Institute of Neuroscience and Physiology, University of Gothenburg, Gothenburg, Sweden. <sup>58</sup>Institut Pasteur de Lille, UMR1167 LabEx DISTALZ – RID-AGE Université de Lille, INSERM, Centre Hospitalier Université de Lille Lille, Lille, France. <sup>59</sup>Departments of Neurology and Public Health Sciences, University of Virginia School of Medicine, Charlottesville, VA, USA. <sup>60</sup>Department of Neurology, Faculty of Medicine, Jagiellonian University Medical College, Krakow, Poland. <sup>61</sup>Department of Neurosurgery, Helsinki University Hospital, University of Helsinki, Helsinki, Finland. <sup>62</sup>Clinical Neurosciences, University of Helsinki, Helsinki, Finland. <sup>63</sup>Neurosurgery NeuroCenter, Kuopio University Hospital, Kuopio, Finland. <sup>64</sup>Institute of Clinical Medicine, Faculty of Health Sciences, University of Eastern Finland, Kuopio, Finland. <sup>65</sup>University of Cincinnati College of Medicine, Cincinnati, OH, USA. <sup>66</sup>These authors jointly supervised this work: Jan H. Veldink, Ynte M. Ruigrok. \*Lists of authors and their affiliations appear at the end of the paper. <sup>✉</sup>e-mail: [m.k.bakker-25@umcutrecht.nl](mailto:m.k.bakker-25@umcutrecht.nl); [ij.m.ruigrok@umcutrecht.nl](mailto:ij.m.ruigrok@umcutrecht.nl)

## HUNT All-In Stroke

**Bendik S. Winsvold<sup>27,28</sup>, Sigrid Børte<sup>28,29,30</sup>, Marianne Bakke Johnsen<sup>28,29,30</sup>, Ben M. Brumpton<sup>28</sup>, Marie Søfteland Sandvei<sup>31,32</sup>, Cristen J. Willer<sup>33</sup>, Kristian Hveem<sup>28,34</sup> and John-Anker Zwart<sup>27,28,30</sup>**

A full list of members and their affiliations appears in the Supplementary Information.

## China Kadoorie Biobank Collaborative Group

**Zheng Bian<sup>66</sup>, Junshi Chen<sup>67</sup>, Yiping Chen<sup>14,15</sup>, Zhengming Chen<sup>14,15</sup>, Robert Clarke<sup>14</sup>, Rory Collins<sup>14</sup>, Yu Guo<sup>66</sup>, Xiao Han<sup>66</sup>, Michael Hill<sup>14,15</sup>, Liming Li<sup>16</sup>, Kuang Lin<sup>14</sup>, Depei Liu<sup>66</sup>, Jun Lv<sup>16</sup>, Iona Millwood<sup>14,15</sup>, Richard Peto<sup>14</sup>, Sam Sansome<sup>14</sup>, Robin Walters<sup>14,15</sup>, Xiaoming Yang<sup>14</sup> and Canqing Yu<sup>66</sup>**

<sup>66</sup>Chinese Academy of Medical Sciences, Beijing, China. <sup>67</sup>China National Center for Food Safety Risk Assessment, Beijing, China. A full list of members and their affiliations appears in the Supplementary Information.

## BioBank Japan Project Consortium

**Masaru Koido<sup>8,9</sup>, Masato Akiyama<sup>8,10,11</sup>, Chikashi Terao<sup>8</sup>, Koichi Matsuda<sup>12,13</sup> and Yoichiro Kamatani<sup>8,12</sup>**

A full list of members and their affiliations appears in the Supplementary Information.

## The ICAN Study Group

**Hubert Desal<sup>4,5</sup>, Romain Bourcier<sup>4,5</sup>, Richard Redon<sup>4</sup>, Christian Dina<sup>4</sup>, Olivier Naggara<sup>22,23</sup>, François Eugène<sup>26</sup>, Jean-Christophe Gentric<sup>24</sup> and Eimad Shotar<sup>25</sup>**

A full list of members and their affiliations appears in the Supplementary Information.

## CADISP Group

**Muralidharan Sargurupremraj<sup>55,56</sup>, Turgut Tatlisumak<sup>57</sup> and Stéphanie Debette<sup>55,56</sup>**

A full list of members and their affiliations appears in the Supplementary Information.

## Genetics and Observational Subarachnoid Haemorrhage (GOSH) Study investigators

**David J. Werring<sup>6</sup>, Henry Houlden<sup>21</sup>, Varinder S. Alg<sup>6</sup>, Isabel C. Hostettler<sup>6,7</sup>, Stephen Bonner<sup>68</sup>, Daniel Walsh<sup>69</sup>, Diederik Bulters<sup>70</sup>, Neil Kitchen<sup>71</sup>, Martin Brown<sup>6</sup> and Joan Grieve<sup>71</sup>**

<sup>68</sup>Department of Anaesthesia, The James Cook University Hospital, Middlesbrough, UK. <sup>69</sup>Department of Neurosurgery, King's College Hospital NHS Foundation Trust, London, UK. <sup>70</sup>Department of Neurosurgery, University Hospital Southampton NHS Foundation Trust, Southampton, UK. <sup>71</sup>Department of Neurosurgery, The National Hospital of Neurology and Neurosurgery, London, UK. A full list of members and their affiliations appears in the Supplementary Information.

## International Stroke Genetics Consortium (ISGC)

**Mark K. Bakker<sup>1</sup>, Romain Bourcier<sup>4,5</sup>, Robin G. Walters<sup>14,15</sup>, Rainer Malik<sup>52</sup>, Martin Dichgans<sup>52,53,54</sup>, Muralidharan Sargurupremraj<sup>55,56</sup>, Turgut Tatlisumak<sup>57</sup>, Stéphanie Debette<sup>55,56</sup>, Gabriel J. E. Rinkel<sup>1</sup>, Bradford B. Worrall<sup>59</sup>, Joanna Pera<sup>60</sup>, Agnieszka Slowik<sup>60</sup>, Joseph P. Broderick<sup>65</sup>, David J. Werring<sup>6</sup>, Daniel Woo<sup>65</sup>, Philippe Bijlenga<sup>3</sup>, Yoichiro Kamatani<sup>8,12</sup> and Ynte M. Ruigrok<sup>1</sup>**

A full list of members and their affiliations appears in the Supplementary Information.



**Mendelian randomization.** To infer causal genetic effects of exposure traits on intracranial aneurysms (the outcome), we used GSMR<sup>16</sup>. We used a meta-analysis of all European ancestry strata, except the UK Biobank (stratum sUK2), as the outcome. As exposures we used summary statistics of 2,419 traits analyzed using UK Biobank data, prepared by the Neale lab, release 2017 (<http://www.nealelab.is/blog/2017/7/19/rapid-gwas-of-thousands-of-phenotypes-for-337000-samples-in-the-uk-biobank>). For a second GSMR run with raw quantitative phenotypes, we used the 2019 GWAS release from the same group. GSMR was run using the GCTA wrapper (v.1.92.2). More details on the method and settings are described in the Supplementary Note.

To determine which of the top significant GSMR traits were independent genetic causes of intracranial aneurysms, the stage 1 GWAS summary statistics were conditioned on the top traits, that is, smoking and BP. Conditioning was done using mtCOJO analysis (GCTA v.1.92.2 beta) as described in Conditional genetic correlation.

**Drug-target enrichment.** Drug-target enrichment analysis was performed according to a previously described method<sup>48</sup>. Gene-wise, *P* values were calculated with MAGMA v.1.06 using a combined approach of average and top *P* values per gene region. Gene-set analysis was performed using MAGMA, with pathways curated from MSigDB<sup>65,66</sup> and TargetValidation (<https://www.targetvalidation.org>), and with drug-target sets described previously<sup>49</sup>. Drug-class enrichment analysis was performed using a Wilcoxon–Mann–Whitney *U*-test. Drug-gene-set *P* values were tested for enrichment in drug classes. Enrichment was expressed as the AUC. AUCs were compared between drug-gene sets within a drug class and all other drug-gene sets.

**Statistics.** The different statistical tests used in the different analysis methods are as follows: (1) SAIGE: logistic mixed model with saddle-point approximation for *P* values; the resulting  $\beta$  values are on the logit scale; (2) METAL: inverse-variance weighted meta-analysis; the resulting  $\beta$  values are on the same scale as the input (here, logit scale); (3) eCAVIAR: directly calculates a colocalization posterior probability from expression and trait GWAS effect sizes using Bayes' rule; (4) TWAS: used to calculate a *z*-score, which is tested against a null distribution of mean zero and unit variance to calculate a *P* value; (5) SMR: the Mendelian randomization effect of exposure (gene expression) *x* on outcome *y* is the ratio of the estimate of the effect of SNP *z* on outcome *y* and SNP *z* on exposure *x*; the SNP effect *z*-scores are used to calculate a  $\chi^2$  statistic with one degree of freedom; (6) LDSC: weighted linear regression, where weights are the inverse of the LD score of a SNP; the slope is divided by sample size and multiplied by the number of SNPs; s.e. values are obtained by the jackknife method; (7) GARFIELD: calculates enrichment odds ratios using logistic regression, accounting for LD, distance to transcription start site and binary annotations; (8) POPCORN: maximum likelihood test; s.e. is calculated using a block jackknife method; (9) GSMR: two-sided linear regression after excluding pleiotropic SNPs using a 'heterogeneity in dependent instrument' test; (10) MAGMA (gene test): uses a multiple linear regression to calculate gene effects; subsequent *P* value is derived from two-sided *F*-test; (11) MAGMA (gene-set test): drug *P* values are calculated by comparing gene *z*-scores (derived from *P* values reported in Supplementary Table 19) in the gene set to those outside the gene set; *P* values are derived from a one-sided Student's *t*-test; and (12) SumHer: conceptually similar to LDSC, but with different weight based on LD and MAF.

**Ethical statement.** All participants provided written informed consent. The Biobank Research Ethics Committee of the University Medical Center Utrecht reviewed and approved the study protocol (TCBio 17-087). The following local data access and ethics committees approved collection and use of genetic data for this study. @neurIST: Medisch Ethische Toetsings Commissie Erasmus MC (METC), Research Committee of the Hospital Clinic de Barcelona, Central Office for Research Ethics Committees (COREC) NHS, and Commission centrale d'éthique de la recherche sur l'être humain de la république et canton de Genève. ARIC: NHLBI Data Access Committee (through dbGaP). Busseleton: GABRIEL Consortium Data Access Committee (through EGA). Utrecht 1: University Medical Center Utrecht Ethics Committee. The Netherlands (EGA): Wellcome Trust Case-Control Consortium Data Access Committee (through EGA). Utrecht 2: University Medical Center Utrecht Ethics Committee. Doetinchem Cohort Study: Scientific Advisory Group of the Netherlands National Institute for Public Health and the Environment. Project MinE: Project MinE GWAS Consortium. French Canadian: Comité d'éthique de la recherche du Centre hospitalier de l'Université de Montréal and McGill University ethics. Finland (EGA): Wellcome Trust Case-Control Consortium Data Access Committee (through EGA). Finland: The ethics committee of Kuopio University Hospital and Helsinki University Hospital. NFBC1966: Ethics Committee of Northern Ostrobothnia Hospital District, Finland. ICAN: Institutional Review Boards (Comité consultatif sur le traitement de l'information en matière de recherche dans le domaine de la santé, Commission Nationale de l'Informatique et des Libertés) and Groupe Nantais d'Éthique dans le Domaine de la Santé (GNEDES). PREGO: Research Ethics Committee (CPP of Nantes). GAIN: NHLBI Data Access Committee (through dbGaP). FIA: University of Cincinnati ethics committee. NonGAIN: NHLBI Data Access Committee

(through dbGaP). Poland: Institutional review board of the Jagiellonian University. NBS: Wellcome Trust Case-Control Consortium Data Access Committee (through EGA). UK Biobank: UK Biobank Data Access Committee. GOSH controls: Central London REC 3 committee. GOSH cases: central London REC 3 committee. NBS+1958BBC: Wellcome Trust Case-Control Consortium Data Access Committee (through EGA). HUNT Study: the Norwegian Data Inspectorate, the Norwegian Board of Health and the Regional Committee for Ethics in Medical Research. China Kadoorie Biobank: Oxford University ethical committee and the China National CDC. The BBJ: research ethics committees at the Institute of Medical Science, University of Tokyo. More details can be found in the Nature Research Reporting Summary.

**Reporting Summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

## Data availability

Summary statistics for stages 1 and 2 GWAS meta-analyses, the SAH-only and uIA-only GWAS, and a meta-analysis consisting of just East Asian samples, including effective sample size per SNP, can be accessed through Figshare (<https://doi.org/10.6084/m9.figshare.11303372>) and through the Cerebrovascular Disease Knowledge Portal (<http://www.cerebrovascularportal.org>). Detailed information on access to publicly available data is given in the Nature Research Reporting Summary.

## References

- Willer, C. J., Li, Y. & Abecasis, G. R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190–2191 (2010).
- Hormozdiari, F. et al. Colocalization of GWAS and eQTL signals detects target genes. *Am. J. Hum. Genet.* **99**, 1245–1260 (2016).
- Gusev, A. et al. Integrative approaches for large-scale transcriptome-wide association studies. *Nat. Genet.* **48**, 245–252 (2016).
- Zhu, Z. et al. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nat. Genet.* **48**, 481–487 (2016).
- Finucane, H. K. et al. Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nat. Genet.* **47**, 1228–1235 (2015).
- Iotchkova, V. et al. GARFIELD classifies disease-relevant genomic features through integration of functional annotations with association signals. *Nat. Genet.* **51**, 343–353 (2019).
- Woo, D. et al. Meta-analysis of genome-wide association studies identifies 1q22 as a susceptibility locus for intracerebral hemorrhage. *Am. J. Hum. Genet.* **94**, 511–521 (2014).
- Brown, B. C., Asian Genetic Epidemiology Network-Type 2 Diabetes Consortium, Ye, C. J., Price, A. L. & Zaitlen, N. Transethnic genetic-correlation estimates from summary statistics. *Am. J. Hum. Genet.* **99**, 76–88 (2016).
- Mootha, V. K. et al. PGC-1 $\alpha$ -responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat. Genet.* **34**, 267–273 (2003).
- Subramanian, A. et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102**, 15545–15550 (2005).

## Acknowledgements

This research has been conducted using the UK Biobank Resource under application no. 2532. We thank R. McLaughlin for the advice on population-based heritability analysis. We thank M. Gunel and K. Yasuno for their help with genotyping DNA samples of the Utrecht 1, Finland and @neurIST cohorts. We thank the staff and participants of all CADISP centers for their important contributions. We acknowledge the contribution of participants, project staff, and the China National Center for Disease Control and Prevention (CDC) and its regional offices to the CKB. China's National Health Insurance provided electronic linkage to all hospital treatments. We thank K. Jepsen for genotyping quality control and imputation of the HUNT Study. For providing clinical information and biological samples collected during the @neurIST project, we thank J. Macho, T. Dóczy, J. Byrne, P. Summers, R. Risselada, M. Sturkenboom, U. Patel, S. Coley, A. Waterworth, D. Rüfenacht, C. Proust and F. Cambien. We acknowledge the support from the Netherlands Cardiovascular Research Initiative: an initiative with support of the Dutch Heart Foundation, CVON2015-08 ERASE. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant no. 852173). This project has received funding from the ERC under the European Union's Horizon 2020 research and innovation program (grant no. 772376—ESCORIAL). The BBJ project was supported by the Ministry of Education, Culture, Sports, Sciences and Technology of the Japanese government, and the Japan Agency for Medical Research and Development (19km0605001). The CADISP study has been supported by INSERM, Lille 2 University, Institut Pasteur de Lille and Lille University Hospital, and received funding from the European Regional Development Fund (FEDER funds) and Région Nord-Pas-de-Calais in the framework of Contrat de Projets Etat-Region 2007–2013 Région Nord-Pas-de-Calais

(grant no. 09120030), Centre National de Génotypage, the Emil Aaltonen Foundation, the Paavo Ilmari Ahvenainen Foundation, the Helsinki University Central Hospital Research Fund, the Helsinki University Medical Foundation, the Päivikki and Sakari Sohlberg Foundation, the Aarne Koskelo Foundation, the Maire Taponen Foundation, the Aarne and Aili Turunen Foundation, the Lilly Foundation, the Alfred Kordelin Foundation, the Finnish Medical Foundation, the Orion Farnos Research Foundation, the Maud Kuistila Foundation, the Finnish Brain Foundation, the Biomedicum Helsinki Foundation, Projet Hospitalier de Recherche Clinique Régional, Fondation de France, Génomôle de Lille, Adrinord, the Basel Stroke Funds, and the Käthe-Zingg-Schwichtenberg-Fonds of the Swiss Academy of Medical Sciences and the Swiss Heart Foundation. S.D. received funding from the French National Funding Agency (ANR), and the ERC under the European Union's Horizon 2020 research and innovation program (grant no. 640643). J.P. was supported by a Jagiellonian University Medical College (grant no. K/ZDS/001456). CKB was supported as follows: baseline survey and first re-survey: Hong Kong Kadoorie Charitable Foundation; long-term follow-up: UK Wellcome Trust (grant nos. 202922/Z/16/Z, 104085/Z/14/Z, 088158/Z/09/Z), National Natural Science Foundation of China (grant nos. 81390540, 81390541, 81390544) and National Key Research and Development Program of China (grant nos. 2016YFC 0900500, 0900501, 0900504, 1303904). DNA extraction and genotyping: GlaxoSmithKline, UK Medical Research Council (grant nos. MC\_PC\_13049, MC\_PC\_14135). Core funding to the Clinical Trial Service Unit and Epidemiological Studies Unit at Oxford University was provided by the British Heart Foundation, UK Medical Research Council and Cancer Research UK. S.Z. and G.A.R. received funding from the Canadian Institutes of Health Research (CIHR). This project has received funding from the European Union's Horizon 2020 research and innovation program (no. 666881), SVDs@target (to M.D.) and CoSTREAM (Common Mechanisms and Pathways in Stroke and Alzheimer's Disease; grant no. 667375, to M.D.); the DFG (Deutsche Forschungsgemeinschaft) as part of the Munich Cluster for Systems Neurology (EXC 2145 SyNergy—ID 390857198) and the CRC 1123 (B3, to M.D.); the Corona Foundation (to M.D.); the Fondation Leducq (Transatlantic Network of Excellence on the Pathogenesis of Small Vessel Disease of the Brain, to M.D.); the e:Med program (e:AtheroSysMed, to M.D.); and the FP7/2007-2103 European Union project CVgenes@target (grant no. Health-F2-2013-601456, to M.D.). K.R. is funded by the Health Data Research UK (HDRUK) fellowship MR/S004130/1. C.L.M.S. was funded by the UK Biobank, HDRUK and Scottish Funding Council. I.C.H. received funding from the Alzheimer Research UK and Dunhill Medical Trust Foundation. J.P.B. and D.W. were supported by National Institutes of Health (NIH) funding. D.J.W. and V.S.A. received funding support from the Stroke Association. D.J.W. and H.H. received funding for genotyping from the NIHR University College London Hospitals Biomedical Research

Center. The Nord-Trøndelag Health Study (HUNT Study) is a collaboration by the HUNT Research Center, Faculty of Medicine at the Norwegian University of Science and Technology (NTNU), the Norwegian Institute of Public Health and the Nord-Trøndelag County Council. The genotyping was financed by the NIH, University of Michigan, the Norwegian Research Council and Central Norway Regional Health Authority and the Faculty of Medicine and Health Sciences, NTNU. P.B. and C.M.F. were supported by EU commission FP6—IST – 027703 @neurIST-Integrated biomedical informatics for the management of cerebral aneurysms. P.B., S.M., S.H., S.S., J.D. and O.M. were supported by the grant (no. MRD 2014/261) from the Swiss SystemsX.ch initiative and evaluated by the Swiss National Science Foundation (AneuX project).

### Author contributions

J.H.V. and Y.M.R. contributed equally to this study. M.K.B., Y.M.R. and J.H.V. wrote the manuscript. Y.M.R. and J.H.V. supervised the project. I.C.H., S.D., B.B.W., J.P., A.S., E.I.G.-P., M.N., J.E.J., M.v.U.Z.F., A.L., J.P.B., D.J.W., D.W., R.R., P.B., Y.K., J.H.V. and Y.M.R. designed the study. M.K.B., R.A.A.v.d.S. and W.v.R. performed the association analyses and scripts. M.K.B., R.A.A.v.d.S. and W.v.R. performed the functional analyses and scripts. S.M., R.B., C.M.F., S.H., S.S., J.D., O.M. and P.B. prepared the phenotypes. R.A.A.v.d.S. and K.R.v.E. provided technical assistance. M.K.B., Y.M.R., G.J.E.R., J.H.V., L.H.v.d.B., P.B., S.M., E.I.G.-P., M.N., J.P., A.S., J.E.E., M.v.U.Z.F., A.L., G.A.R., S.Z., N.U.K., R.M., K.R., C.L.M.S., D.J.W., I.C.H., H.H., V.S.A., J.P.B., D.W., R.R., R.B., C.D., O.N., J.-C.G., E.S., F.E., H.D. and W.M.M.V. contributed the phenotypes and genotypes. Y.K., M.K., M.A., C.T., K.M. (BBJ), R.G.W., K.L., L.L., I.Y.M., Z.C. (CKB), B.S.W., S.B., M.B.J., B.M.B., M.S.S., C.J.W., K.H., J.-A.Z. (HUNT), M.J.B., G.T.J. (AAA), H.K., J.G.Z., C.J.M.K., N.U.K. (AVM), D.W. (ICH), R.M., M.D. (IS), S.D., T.T., M.S. and P.A. (cervical artery dissection) summarized the statistical contributions. J.R.I.C. and G.B. performed drug-target and MAGMA pathway enrichment analyses. D.J.W. critiqued the output for important intellectual content.

### Competing interests

When this study was conducted, C.L.M.S. was chief scientist for the UK Biobank study.

### Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41588-020-00725-7>.

**Correspondence and requests for materials** should be addressed to M.K.B. or Y.M.R.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

## Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

### Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided  
*Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g.  $F$ ,  $t$ ,  $r$ ) with confidence intervals, effect sizes, degrees of freedom and  $P$  value noted  
*Give  $P$  values as exact values whenever suitable.*
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's  $d$ , Pearson's  $r$ ), indicating how they were calculated

*Our web collection on [statistics for biologists](#) contains articles on many of the points above.*

### Software and code

Policy information about [availability of computer code](#)

Data collection

No software was used for data collection.

Data analysis

R 3.2.2 and 3.6.1. For creating plots, and running R-based programs SAIGE, FUSION, and GARFIELD. (<https://www.r-project.org>)  
 plink 1.90 for genotype QC and genetic risk score (GRS) analysis. (<https://www.cog-genomics.org/plink2/>)  
 SAIGE 0.29.3 for genome-wide logistic mixed model association analysis. (<https://github.com/weizhouUMICH/SAIGE>)  
 GCTA v1.91.1beta as wrapper for COJO (<http://cnsgenomics.com/software/gcta/>)  
 GCTA v1.92.2beta as wrapper for mtCOJO and GSMR (<http://cnsgenomics.com/software/gcta/>)  
 SMR 0.712 (linux, <http://cnsgenomics.com/software/smr/>)  
 FUSION (no version) for TWAS analysis. ([https://github.com/gusevlab/fusion\\_twass](https://github.com/gusevlab/fusion_twass))  
 LDSC v1.0.0 for population stratification estimate, genetic correlation analysis, SNP heritability estimate, and heritability enrichment analyses. (<https://github.com/bulik/ldsc>)  
 GARFIELD v2 for functional enrichment analysis. (<https://www.ebi.ac.uk/birney-srv/GARFIELD/>)  
 METAL release 2011-03-25 for meta-analysis of genome-wide association analyses per stratum. (<http://csg.sph.umich.edu/abecasis/Metal/>)  
 POPCORN 0.9.9 for trans-ancestry genetic correlation analysis. (<https://github.com/brielin/Popcorn>)  
 eCAVIAR v2.2 (<http://genetics.cs.ucla.edu/caviar/>)  
 MAGMA v1.06 (<https://ctg.cncr.nl/software/magma>)  
 SumHer is integrated in Ldak5 (linux, <http://dougsspeed.com/sumher/>)

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

## Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Summary statistics for the Stage 1 and Stage 2 GWAS meta analyses, as well as summary statistics for the ruptured IA versus controls, unruptured IA versus controls, and the analysis including only East Asian participants can be accessed using doi: 10.6084/m9.figshare.11303372. And through the cerebrovascular disease knowledge portal <http://www.cerebrovascularportal.org>.

The following datasets have been accessed and used.

Database of genotypes and phenotypes (dbGaP): <https://dbgap.ncbi.nlm.nih.gov>  
Health and Retirement Study (HRS) genotype data (dbGaP accession code phs000428.v2.p2)  
Atherosclerosis Risk in Communities (ARIC) genotype data (dbGaP accession code phs000280.v5.p1)  
Genetic Association Information Network (GAIN) genotype data (dbGaP accession code phs000021.v3.p2)  
Schizophrenia controls, not included in the GAIN study (nonGAIN, dbGaP accession code phs000167.v1.p1)

European Genome-Phenome Archive (EGA): <https://ega-archive.org>  
Genotypes from the Busselton cohort (EGA accession code EGAD00000000076)  
Genotypes from Netherlands, Finland and Italy (EGAD00010000288)  
Genotypes from National blood donors cohort (EGAD00000000023 and EGAD00010000290)  
Genotypes from 1958 British birth cohort (EGAD00000000021)

UK Biobank summary statistics from the Neale lab (<http://www.nealelab.is/uk-biobank>)  
SMR reference data (<https://cnsgenomics.com/software/smr/#eQTLsummarydata>)  
LDSC reference data (<https://github.com/bulik/ldsc>)  
Mouse brain single-cell RNAseq data for LDSC (Supplementary Table 4 of Skene, N.G. et al. Genetic identification of brain cell types underlying schizophrenia. Nat Genet 50, 825-833 (2018).)  
Mouse brain vasculature single-cell RNaseq data for LDSC. Published as dataset in: He, L. et al. Single-cell RNA sequencing of mouse brain and lung vascular and vessel-associated cell types. Sci Data 5, 180160 (2018).  
TWAS reference data (<http://gusevlab.org/projects/fusion/#reference-functional-data>)  
eCAVIAR reference data (Downloaded from GTEx resource: [https://storage.googleapis.com/gtex\\_analysis\\_v7/single\\_tissue\\_eqtl\\_data/GTEx\\_Analysis\\_v7\\_eQTL\\_all\\_associations.tar.gz](https://storage.googleapis.com/gtex_analysis_v7/single_tissue_eqtl_data/GTEx_Analysis_v7_eQTL_all_associations.tar.gz))

Access to summary statistics of the Aortic Abdominal Aneurysms (AAA), arteriovenous malformations (AVM), intracerebral hemorrhage (ICH), ischemic stroke (IS) and cervical artery dissection studies are specified in their respective publications:

AAA: Jones, G.T. et al. Meta-Analysis of Genome-Wide Association Studies for Abdominal Aortic Aneurysm Identifies Four New Disease-Specific Risk Loci. Circ Res 120, 341-353 (2017).

AVM: Weinsheimer, S. et al. Genome-wide association study of sporadic brain arteriovenous malformations. J Neurol Neurosurg Psychiatry 87, 916-23 (2016).

ICH: Woo, D. et al. Meta-analysis of genome-wide association studies identifies 1q22 as a susceptibility locus for intracerebral hemorrhage. Am J Hum Genet 94, 511-21 (2014).

IS: Malik, R. et al. Multiancestry genome-wide association study of 520,000 subjects identifies 32 loci associated with stroke and stroke subtypes. Nat Genet 50, 524-537 (2018).

Cervical artery dissection: Debette, S. et al. Common variation in PHACTR1 is associated with susceptibility to cervical artery dissection. Nat Genet 47, 78-83 (2015).

## Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences       Behavioural & social sciences       Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](http://nature.com/documents/nr-reporting-summary-flat.pdf)

## Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size

Sample sizes per cohort, stratum, and used in each meta analysis, are shown in Table 2 and Supplementary Table 1. In some instances, in particular when converting observed scale heritability to the liability scale, effective sample size is a more appropriate measure for sample size. The calculation for this is described in the Supplementary Note, section Analysis rationale and parameters. In instances where effective sample size is used, ascertainment (the fraction of cases in the sample) is set at 0.5.

Data exclusions

A description of quality control (QC) metrics, is described in the Supplementary Note, section Quality control parameters. The recruitment methods, including inclusion and exclusion criteria are described briefly in the Online Methods, section Recruitment and diagnosis and in detail in the Supplementary Note, section Cohort description.

Replication	No replication was performed
Randomization	Cases and controls were defined by disease status (intracranial aneurysm and/or subarachnoid hemorrhage), and could therefore not be randomized. Control datasets were matched based on genotyping platform and population, creating strata that were as homogeneous as possible. Extensive (QC) was performed to exclude population and batch outliers, as described in the Supplementary Note, section Quality control parameters. We used a mixed model, including genetic relationship matrix to account for population stratification and batch effect, and used principal components we included in the association analyses to further account for this. Sex was used as covariate to account for sex-specific confounding due to differences in prevalence between sexes.
Blinding	Blinding was not applicable, since groups were defined by disease status.

## Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

### Materials & experimental systems

### Methods

n/a	Included in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input type="checkbox"/>	<input checked="" type="checkbox"/> Human research participants
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data

n/a	Included in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

## Human research participants

Policy information about [studies involving human research participants](#)

Population characteristics	An overview of number of cases and controls, sex distribution, and genotyping platform is given in Supplementary Table 1. The age of the participants was not registered for most of the cohorts included in this study.
Recruitment	<p>The recruitment methods, including inclusion and exclusion criteria are described briefly in the Online Methods, section Recruitment and diagnosis and in detail in the Supplementary Note, section Cohort description.</p> <p>Given this recruitment method, we see two potential forms of bias.</p> <p>The first and largest source of potential bias is caused by the fact that there is no routine screening for intracranial aneurysms. This leads to an imbalance toward the inclusion of more subarachnoid hemorrhage cases, compared to unruptured intracranial aneurysm cases. It is therefore important to realize that the presented results for a large part represent the genetics of aneurysmal subarachnoid hemorrhage.</p> <p>Second, the most severe cases of subarachnoid hemorrhage are more likely to die before reaching a hospital, and therefore less likely to be included in the study. This could lead to an underrepresentation of severe cases. However, this is the not case for the included cohort studies (UK Biobank, Biobank Japan, HUNT study, China Kadoorie Biobank), where the cases were recruited before the clinical events. Since these cohort studies represent a large portion of the study population, we expect the effect of this potential bias to be small.</p>
Ethics oversight	<p>Local Data Access Committees reviewed and accepted the protocols for data collection and use for each cohort. Guidelines for the study procedure were provided by the Biobank Research Ethics Committee of the University Medical Center Utrecht (TCBio 17-087).</p> <p>The local data access and ethical committee who approved collection and use of genetic data are shown below</p> <p>-@neurIST Rotterdam: Medisch Ethische Toetsings Commissie Erasmus MC (METC) Barcelona (both centers): Research Committee of the Hospital Clinic de Barcelona Oxford and Sheffield: Central Office for Research Ethics Committees (COREC) NHS Geneva: Commission centrale d'éthique de la recherche sur l'être humain de la république et canton de Genève -ARIC NHLBI Data Access Committee (through dbGaP) -Busselton GABRIEL Consortium Data Access Committee. (through EGA) -Utrecht 1 University Medical Center Utrecht Ethics Committee -Netherlands (EGA) Wellcome Trust Case-Control Consortium Data Access Committee. (through EGA)</p>

-Utrecht 2  
 University Medical Center Utrecht Ethics Committee  
 -Doetinchem Cohort Study  
 Scientific Advisory Group of the Netherlands National Institute for Public Health and the Environment  
 -Project MinE  
 Project MinE GWAS Consortium  
 -French Canadian  
 Comité d'éthique de la recherche du Centre hospitalier de l'Université de Montréal and McGill University ethics  
 -Finland (EGA)  
 Wellcome Trust Case-Control Consortium Data Access Committee. (through EGA)  
 -Finland  
 The ethics committee of Kuopio University Hospital and Helsinki University Hospital  
 -NFBC1966  
 Ethics Committee of Northern Ostrobothnia Hospital District, Finland  
 -ICAN  
 Institutional Review Boards (Comité consultatif sur le traitement de l'information en matière de recherche dans le domaine de la santé, Commission Nationale de l'Informatique et des Libertés) and Groupe Nantais d'Ethique dans le Domaine de la Santé (GNEDS)  
 -PREGO  
 Research Ethics Committee (CPP of Nantes)  
 -GAIN  
 NHLBI Data Access Committee (through dbGaP)  
 -FIA  
 University of Cincinnati ethics committee  
 -nonGAIN  
 NHLBI Data Access Committee (through dbGaP)  
 -Poland  
 Institutional review board of the Jagiellonian University.  
 -NBS  
 Wellcome Trust Case-Control Consortium Data Access Committee. (through EGA)  
 -UK Biobank  
 UK Biobank Data Access Committee  
 -GOSH controls  
 Central London REC 3 committee  
 -GOSH cases  
 Central London REC 3 committee  
 -NBS+1958BBC  
 Wellcome Trust Case-Control Consortium Data Access Committee. (through EGA)  
 -HUNT study  
 The Norwegian Data Inspectorate, the Norwegian Board of Health, and the Regional Committee for Ethics in Medical Research  
 -China Kadoorie Biobank  
 Oxford University ethical committee and the China National CDC  
 -Biobank Japan  
 Research ethics committees at the Institute of Medical Science, the University of Tokyo

Note that full information on the approval of the study protocol must also be provided in the manuscript.