



OPEN

## Fine-mapping of intracranial aneurysm susceptibility based on a genome-wide association study

Eun Pyo Hong<sup>1</sup>, Dong Hyuk Youn<sup>1</sup>, Bong Jun Kim<sup>1</sup>, Jun Hyong Ahn<sup>2</sup>, Jeong Jin Park<sup>3</sup>, Jong Kook Rhim<sup>4</sup>, Heung Cheol Kim<sup>5</sup>, Gyojun Hwang<sup>6</sup>, Hong Jun Jeon<sup>7</sup> & Jin Pyeong Jeon<sup>7</sup>✉

In addition to conventional genome-wide association studies (GWAS), a fine-mapping analysis is increasingly used to identify the genetic function of variants associated with disease susceptibilities. Here, we used a fine-mapping approach to evaluate candidate variants based on a previous GWAS involving patients with intracranial aneurysm (IA). A fine-mapping analysis was conducted based on the chromosomal data provided by a GWAS of 250 patients diagnosed with IA and 296 controls using posterior inclusion probability (PIP) and log<sub>10</sub> transformed Bayes factor (log<sub>10</sub>BF). The narrow sense of heritability ( $h^2$ ) explained by each candidate variant was estimated. Subsequent gene expression and functional network analyses of candidate genes were used to calculate transcripts per million (TPM) values. Twenty single-nucleotide polymorphisms (SNPs) surpassed a genome-wide significance threshold for credible evidence (log<sub>10</sub>BF > 6.1). Among them, four SNPs, rs75822236 (*GBA*; log<sub>10</sub>BF = 15.06), rs112859779 (*TCF24*; log<sub>10</sub>BF = 12.12), rs79134766 (*OLFML2A*; log<sub>10</sub>BF = 14.92), and rs371331393 (*ARHGAP32*; log<sub>10</sub>BF = 20.88) showed a completed PIP value in each chromosomal region, suggesting a higher probability of functional candidate variants associated with IA. On the contrary, these associations were not shown clearly under different replication sets. Our fine-mapping analysis suggested that four functional candidate variants of *GBA*, *TCF24*, *OLFML2A*, and *ARHGAP32* were linked to IA susceptibility and pathogenesis. However, this approach could not completely replace replication sets based on large-scale data. Thus, caution is required when interpreting results of fine-mapping analysis.

Intracranial aneurysm (IA) refers to an abnormal focal dilatation of a cerebral artery due to a weakening of the intima of a blood vessel wall. The prevalence of IA in the general adult population has been reported to be nearly 6%<sup>1</sup>. The incidence of unruptured intracranial aneurysms (UIAs) was 15.6 per 100,000 persons<sup>2</sup>. The asymptomatic UIA can rupture suddenly resulting in subarachnoid hemorrhage (SAH), which is associated with a higher mortality rate exceeding 50% within one month after ictus<sup>3,4</sup>. IA is a complex disease involving an interaction between clinical and genetic factors underlying its formation and growth<sup>2</sup>. Important clinical risk factors for IA include female gender, hypertension and smoking. The risk of rupture is increased when the aneurysm is located between arterial branches or in the vertebrobasilar region, in addition to larger size at diagnosis, and the presence of a bleb or daughter sac<sup>5-7</sup>. Genetic studies have been performed to identify genes associated with IA via linkage analysis and single nucleotide polymorphisms (SNPs) of known candidate genes, strongly correlated genes or genome-wide association studies (GWAS) for screening multiple candidate genes. In particular, GWAS revealed large-scale genetic associations, which were primarily correlated with traits and diseases. GWAS technically compares allele frequencies in SNPs between cases and controls. However, complex traits of

<sup>1</sup>Institute of New Frontier Research, Hallym University College of Medicine, Chuncheon, Republic of Korea. <sup>2</sup>Department of Neurosurgery, Gangwon National University College of Medicine, Chuncheon, Gangwon-do, Republic of Korea. <sup>3</sup>Department of Neurology, Konkuk University Medical Center, Seoul, Republic of Korea. <sup>4</sup>Department of Neurosurgery, Jeju National University College of Medicine, Jeju, Republic of Korea. <sup>5</sup>Department of Radiology, Hallym University College of Medicine, Chuncheon, Gangwon-do, Republic of Korea. <sup>6</sup>Department of Neurosurgery, DMC Bundang Jesaeng Hospital, Seongnam, Gyeonggi-do, Republic of Korea. <sup>7</sup>Department of Neurosurgery, Hallym University College of Medicine, 77 Sakju-ro, Chuncheon, Gangwon-do 24253, Republic of Korea. ✉email: jjs6553@daum.net

IA are not entirely attributed to a single gene, but are caused by the influence of multiple genes<sup>8,9</sup>. Gene–gene and gene–environment interactions also affect the traits and diseases. Given the inherent features of the GWAS, genetic markers included in the same linkage disequilibrium exhibit similar correlation. Accordingly, even if a candidate gene is identified via GWAS, it could merely suggest a statistically significant difference rather than represent an etiological factor. Further, the precise location of the causative gene may differ in the same linkage disequilibrium (LD) block. Thus, it is important to reduce the errors via additional data processing to identify false-positive results obtained in GWAS.

Fine-mapping is one of the post-GWAS analyses used to narrow potential candidate variants directly affecting the trait<sup>10</sup>. This approach can be used to identify the regions associated with possible susceptibility based on the population of structures with LD<sup>9</sup>. It provides a complex correlation between the candidate variants and the disease development using computational data without in vivo and in vitro molecular biology studies<sup>10,11</sup>. Sekar et al.<sup>11</sup> showed that structurally diverse alleles of the complement component 4 genes contribute to schizophrenia via excessive complement activity, resulting in reduced numbers of synapses. Consequently, it can be used to assess the functional role of the risk allele, which is a challenge to investigate based on molecular mechanisms, despite the strong genetic association. Fine-mapping of complex traits has been increasingly performed in many diseases, especially cancer and stroke, but has yet to be reported in IA. Here, for the first time, we performed a fine-mapping analysis based on previous GWAS data sets to identify the candidate variants in an effort to identify the precise genetic variants associated with IA in a Korean adult cohort. We also performed a functional gene set enrichment analysis using the optimized candidate sets to analyze the biological relationship between candidate genes and IA.

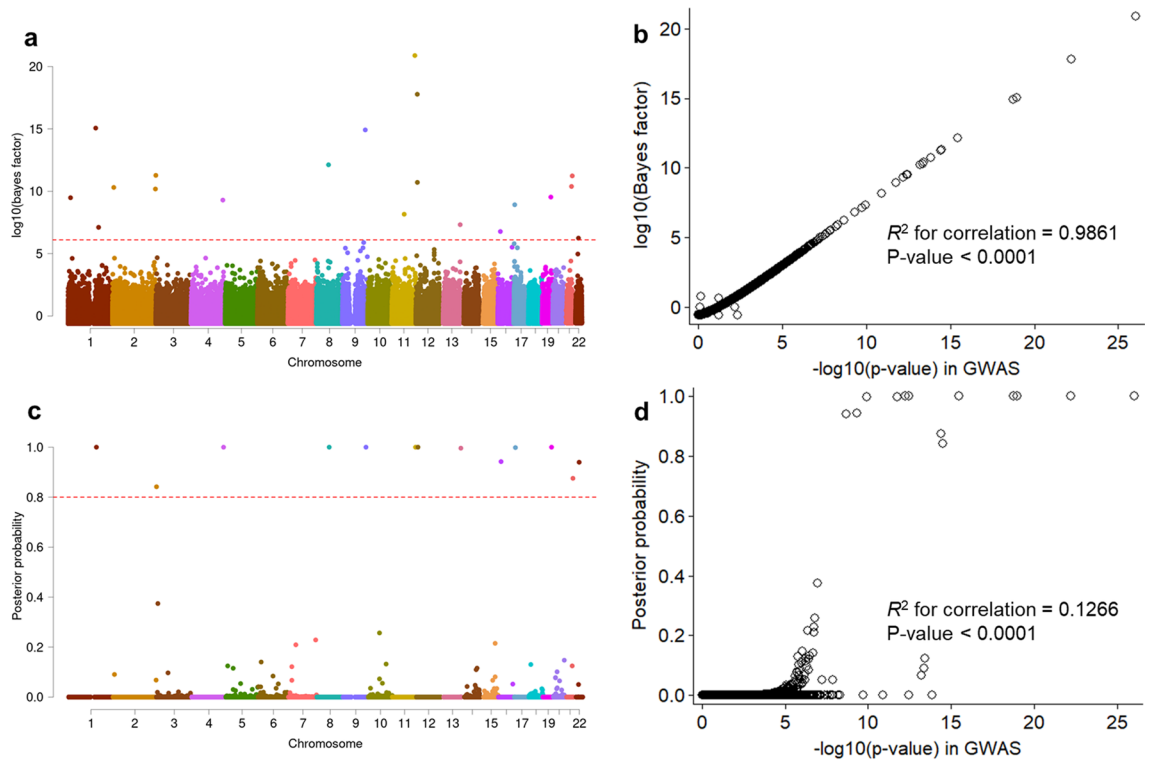
## Materials and methods

**GWAS-based summary statistics.** The analysis was based on the summary statistics provided by the previous IA GWAS. In brief, the study included 250 adult patients with saccular aneurysm and 296 controls between March 2015 and December 2020<sup>12,13</sup>. The Axiom<sup>TM</sup> Asian Precision Medicine Research Array (APMRA) (Thermo Fisher Scientific, MA, USA) were used for genotyping of the study subjects. High-quality plates were defined by a plate pass rate higher than 95% for samples. The average call rate of passing samples was greater than 99%. A total of 512,575 SNPs passed the quality control including genotyping call rate of 95% or higher, minor allele frequency (MAF) at least 1%, and Hardy–Weinberg equilibrium (HWE) with a  $p$  value  $\geq 1 \times 10^{-612}$ . GWAS-based summary statistics included allele types, MAF, and effect sizes. All methods were carried out in accordance with relevant guidelines and regulations. This study protocols including all subjects providing written informed consents has been approved by The Institutional Review Board of the Hallym University Chuncheon Sacred Heart Hospital (No. 2016-3 and 2019-06-006-012).

**Statistical analysis.** We performed a fine-mapping study to identify the role of candidate combinations in the susceptibility to IA using FINEMAP v.1.3.1<sup>12,14</sup>. Odds ratios (ORs) of individual SNPs were converted using the natural log-transformed formula (lnOR). Causality of each SNP or configuration was assessed using effect sizes, posterior inclusion probabilities (PIPs), and narrow sense of heritability ( $h^2$ ), which was explained by candidate SNPs. The log<sub>10</sub>-transformed Bayes factors (log<sub>10</sub>BF) for the individual SNPs and configurations were estimated via FINEMAP analysis. A log<sub>10</sub>BF value of above 2 was considered a creditable threshold in the FINEMAP. However, a log<sub>10</sub>BF value of greater than 6.1 suggested significant genome-wide evidence in GWAS. The fine-mapping approach requires the estimates of SNP correlations, and therefore LD matrices between SNPs were generated by PLINK v1.9 (<https://www.cog-genomics.org/plink/>)<sup>15</sup>. All the fine-mapping tests were conducted with individual chromosomes (chr1–22) due to the LD-based mapping procedure. Manhattan and regional association plots of fine-mapping results were obtained using the package of “*qqman*” in R v3.6.1 (<https://cran.r-project.org/web/packages/qqman>) and LocusZoom v1.3 written in the modified Python and R scripts<sup>16</sup>. Regional annotations and functional impact of SNPs were described using the ANNOVAR program (<http://www.openbioinformatics.org/annovar/>)<sup>17</sup>.

**Gene expression and functional network analyses.** The expressions of candidate genes was evaluated in human blood, brain-specific tissues, or cells using the Geotype-Tissue Expression (GTEx) Portal (<https://gtexportal.org/home/>)<sup>18</sup>. Transcripts per million (TPM) values of a total of 56,200 genes were calculated in 13 brain tissues, 4 blood vessels (3 arterial tissues and 1 cell line of EBV-transformed lymphocytes), and a whole blood cell. Subsequent gene functional network analysis was conducted using the GeneMANIA program (<https://genemania.org/>)<sup>19</sup>.

**Subsequent validation of 29 SNPs under different replication stages.** Targeted loci of 29 SNPs suggested by fine-mapping analysis based on initial GWAS were tested using the following two independent sets: (1) stage 2, independent subjects of 50 patients with IA and 46 controls from two Hallym University hospitals, Chuncheon and Gangdong Sacred Heart Hospitals (IA group: 24 males and 26 females aged 28–80 years; control group: 22 males and 24 females, aged 19–72 years); and (2) stage 3, the same independent set of 50 hospital-enrolled patients in stage 2 and another 575 healthy controls without underlying diseases. Controls were extracted from 8105 subjects of the Rural and Mid-size City cohort of the Korean Genome Epidemiology Study. We further accounted for a stringent significance threshold of 0.001724 according to multiple testing (0.05/29 SNPs).



**Figure 1.** (a,c) Manhattan plots show log<sub>10</sub>-transformed Bayes factors (log<sub>10</sub>BF) and posterior inclusion probability (PIP) of variant causality estimations based on the summary statistics of a genome-wide association study (GWAS) of intracranial aneurysm (IA). (b,d) Plots compare the significance of IA GWAS, log<sub>10</sub>BF, and PIP. A red dash line indicates a genome-wide significance and a strong PIP of IA formation in the panels (a) and (c), respectively (log<sub>10</sub>BF = 6.1 and PIP = 0.8). R-square ( $R^2$ ) indicates the correlation between IA GWA p-value ( $-\log_{10}$  transformed) and log<sub>10</sub>BF in panel (b) and PIP in panel (d);  $p$ -value for  $R^2$ .

## Results

**Fine-mapping analyses.** Study workflow and summarized outcomes are shown in Supplementary Fig. S1. Fine mapping analyses entailed multiple SNP combinations for each autosome chromosome. A total of 20 candidate SNPs surpassed a genome-wide significance threshold for credible evidence (log<sub>10</sub>BF > 6.1) (Fig. 1a). The log<sub>10</sub>BF values were strongly correlated with GWAS-based  $p$ -values ( $R^2$  for correlation = 0.9861,  $p < 0.0001$ ) (Fig. 1b). Among them, 13 SNPs showed a high PIP in each chromosomal region (PIP > 0.8) (Fig. 1c). The PIPs were slightly correlated with GWAS-driven  $p$ -values ( $R^2$  for correlation = 0.0974,  $p < 0.0001$ ) (Fig. 1d). Six out of 13 SNPs were located on exonic regions (Table 1). The rs371331393 (Q1932X), one SNP of *ARHGAP32* (11q24.3), exhibited a stop-gain function and was most significantly associated with increased IA risk (lnOR = 3.77, log<sub>10</sub>BF = 20.88). This variant showed the highest heritability ( $h^2 = 0.143$ ). Three SNPs including rs75822236 (R535H, *GBA*; log<sub>10</sub>BF = 15.06), rs112859779 (G141S, *TCF24*; log<sub>10</sub>BF = 12.12), and rs79134766 (A208T, *OLFML2A*; log<sub>10</sub>BF = 14.92) were associated with amino acid substitutions, which may contribute to possible DNA sequence damage and increased association with IA. Remaining 16 IA-predicting candidate loci (6 intergenic, 8 intronic, and 2 synonymous SNPs; log<sub>10</sub>BF > 6.1 and  $p < 5 \times 10^{-8}$ ) were not accompanied by any neighbor promoter or functional variations such as UTR, missense, and nonsense SNPs that showed both moderate to strong LD ( $0.5 < r^2 < 1$ ) and marginally significant associations ( $p < 0.0001$ , data not shown). This implicated that these SNPs could be commonly predictable false positive outcomes or potentially related to other mechanisms such as alternative splicing and gene fusions, but not associated with IA. Eighteen SNPs, including *PRDM2*, *FMO4*, and *RNF144A*, which exceeded genome-wide significance level in GWAS ( $p < 5 \times 10^{-8}$ ), were associated with low causality to IA formation ( $1.89 \times 10^{-11} < \text{PIP} < 0.13$ , log<sub>10</sub>BF > 5) (Supplementary Table S1). In summary, fine mapping annotated that *ARHGAP*, *GBA*, *TCF24*, and *OLFML2A* could be potentially candidate genes contributing to DNA damaging. However, functional mutations around 16 other genes including intergenic, intronic, non-coding RNA, and synonymous loci were not discovered or they showed insignificant associations in this study (log<sub>10</sub>BF < 2 or  $p > 0.05$ , data not shown). Detailed information including SNP allele frequency and HWE  $P$  value is described in Supplementary Table S1.

**Gene expression and functional network analyses.** Gene expression and functional network analyses were performed using the four candidate genes (Fig. 2 and Supplementary Table S2). *GBA* was broadly enriched in all tissues and cell lines ( $6 < \text{TPM} < 34$ ). In particular, it was highly expressed in the whole blood (TPM = 33.13). Conversely, *TCF24* was rarely expressed in all tissues and cells (TPM < 0.15). Expression in *OLFML2A* was moderate in all arteries (TPM = 2.37 to 9.54). *ARHGAP32* was rarely enriched in EBV-transformed lymphocytes and

Gene	Chr	Function	SNP	M/m <sup>a</sup>	MAF	PIP <sup>b</sup>	log10BF <sup>b</sup>	h <sup>2b</sup>	lnOR	P in GWAS	HWE p
PRDM2	1p36.21	Intronic	rs61775135	C/A	0.285	2.61E-06	9.48	0.0733	- 1.15	3.59E-13	0.626
GBA	1q22	R535H, exon11	rs75822236	C/T	0.166	1.0000	15.06	0.1088	5.08	1.09E-19	1
FMO4	1q24.3	F281F, exon8	rs3737926	C/T	0.264	1.10E-08	7.10	0.0577	- 1.01	1.83E-10	0.079
RNF144A	2p25.1	Intron	rs6741819	C/T	0.247	0.0905	10.30	0.0785	- 1.38	4.05E-14	0.022
HDAC4,LOC150935	2q37.3	Intergenic	rs59626274	C/T	0.248	0.0679	10.18	0.0777	- 1.34	5.78E-14	0.002
LINC01237	2q37.3	ncRNA, intron	rs78458145	G/A	0.285	0.8416	11.27	0.0847	- 1.41	3.14E-15	0.007
SPCS3,VEGFC	4q34.2	Intergenic	rs17688188	G/A	0.222	0.9999	9.29	0.0718	- 1.31	5.99E-13	0.030
TCF24	8q13.1	G141S, exon4	rs112859779	C/T	0.216	1.0000	12.12	0.0898	- 1.69	3.33E-16	2.02E-04
OLFML2A	9q33.3	A208T, exon4	rs79134766	G/A	0.219	1.0000	14.92	0.1072	- 1.97	1.70E-19	0.054
MYEOV,LINC01488	11q13.3	Intergenic	rs76855873	C/T	0.269	1.89E-13	8.15	0.0642	- 1.11	1.23E-11	0.023
ARHGAP32	11q24.3	Q1932X, exon22	rs371331393	G/A	0.171	1.0000	20.88	0.1435	3.77	9.32E-27	1
CD163L1	12p13.31	Splicing	rs138525217	C/T	0.161	1.0000	17.77	0.1248	4.33	6.20E-23	1
SLC2A14	12p13.31	Intron	rs118107419	C/A	0.262	8.60E-08	10.71	0.0807	- 1.40	1.44E-14	0.012
CUL4A,LAMP1	13q34	Intergenic	rs74115822	G/A	0.112	0.9960	7.32	0.0584	1.83	1.12E-10	0.001
MIR365A,PARN	16p13.12	Intergenic	rs11646803	C/T	0.376	0.9422	6.77	0.0548	- 0.79	4.76E-10	0.074
MINK1	17p13.2	Intron	rs72835045	G/A	0.220	0.9984	8.92	0.0688	- 1.38	1.69E-12	1.16E-06
NAPA-ASI	19q13.32	ncRNA, intron	rs55800589	G/C	0.364	1.0000	9.53	0.0726	- 0.96	3.35E-13	2.37E-06
DSCAM	21q22.2	Intron	rs727333	C/A	0.257	0.1247	10.38	0.0773	- 1.36	3.69E-14	4.75E-04
LRRC3	21q22.3	P63P, exon2	rs116969723	G/A	0.233	0.8753	11.23	0.0827	- 1.45	3.83E-15	0.014
RFPL2,SLC5A4	22q12.3	Intergenic	rs117398778	T/C	0.138	0.9397	6.24	0.0506	1.32	2.00E-09	1.56E-05

**Table 1.** Significant candidate loci identified by fine-mapping after genome-wide association study. GWAS genome-wide association study, *log10BF* log10 transformed Bayes factor, *lnOR* natural log-transformed odds ratio, *PIP* posterior inclusion probability, *MAF* minor allele frequency. <sup>a</sup>M/m indicates major/minor allele type, respectively. <sup>b</sup>PIP, log10BF, and heritability (*h*<sup>2</sup>) of individual variants were estimated via FINEMAP program to identify possible susceptibility to intracranial aneurysm (IA). <sup>c</sup>lnOR and *P*-value were estimated by IA GWAS.

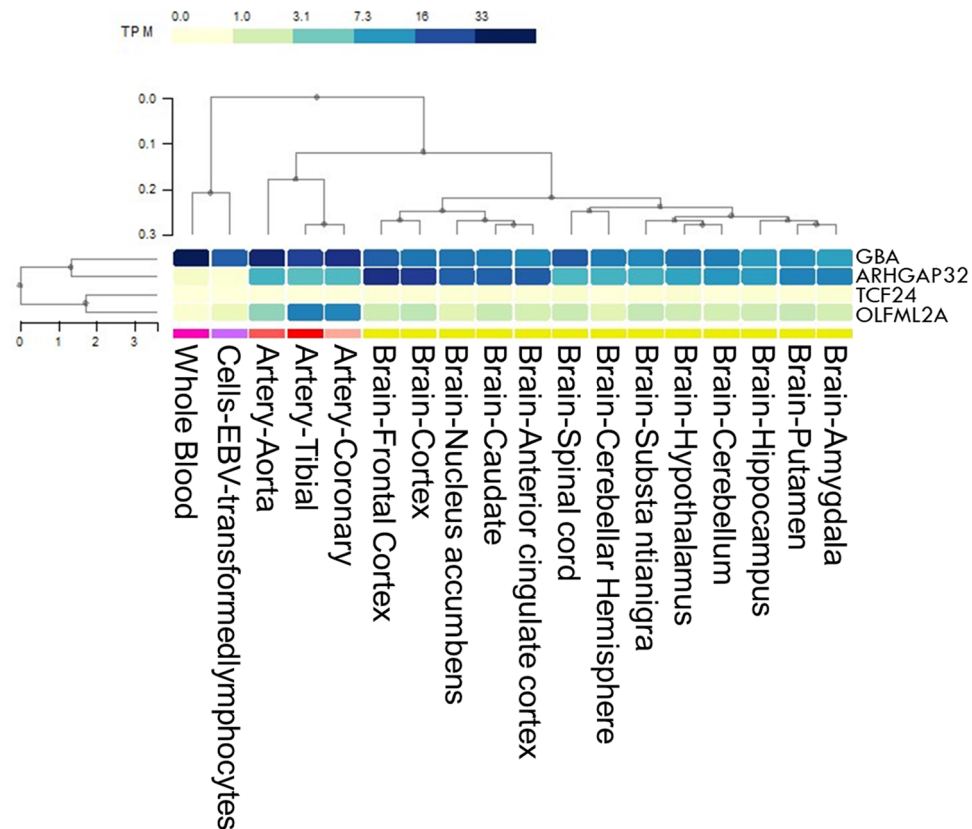
whole blood (TPM = 0.0667 and 0.2274, respectively), while it was enriched in the brain and blood vessels (TPM, between 3 and 24). No direct interaction was observed between the four candidate genes (Fig. 3). However, these genes constitute a hub network interacting with neighbor genes, especially *PSAP*, *SCARB2*, and *ASAH1* (Fig. 3).

**Subsequent validation of 29 SNPs after fine-mapping.** Out of a total 29 SNPs that revealed by fine-mapping analysis based on initial GWAS, only five SNPs such as rs79461840, rs4979583, rs7964241, rs2440154, and rs117398778 showed a statistical significance at independent stage 2 ( $0.002 < p < 0.05$ ). Five SNPs such as rs75822236, rs127561723, rs371331393, rs138525217, and rs55800589 could not be analyzed in fully at this replication stage because minor alleles of these SNPs show no frequency in the control group. Another replication stage 3 showed that only eight SNPs were matched with the GWAS-driven fine mapping analysis. Among them, two SNPs such as rs56942085 (*LINGO2*) and rs2440154 (*SLC47A1*) were closely associated with IA ( $p = 0.03301$  and  $0.01104$ , respectively). During replication studies, only one SNP rs2440154 showed a constant association with IA ( $p < 0.05$ ), but did not surpass an adjusted *p*-value less than 0.05 after multiple testing correction ( $p < 0.001724$ ) (Supplementary Table S3).

## Discussion

This study performed additional fine mapping analysis based on the original GWAS and found four candidate variations of *GBA*, *TCF24*, *OLFML2A*, and *ARHGAP32* genes potentially linked to IA in Koreans. We speculated that these genetic variants may cause dysfunctional immune response and inflammation in DNA sequences damaged by amino acid substitution or gain- or loss-of-function mutations, which affects the IA formation. The “T” allele of rs75822236 located in the exonic region of *GBA* increased the risk of IA via previous GWAS<sup>12</sup>. In addition, a fine-mapping analysis revealed a higher level of log10BF (15.06) and PIP (1.0), suggesting that this variant was a true positive for IA, although there might be concerns about its reliability due to the small sample size of the original GWAS datasets. The role of *GBA* was mainly investigated in Parkinson’s disease (PD) or Gaucher disease (GD), which is a recessive lysosomal storage disorder, and barely investigated in IA. Mata et al.<sup>20</sup> reported that *GBA* mutations and E326K carrier were related to impaired working memory and executive function in patients with PD. In GD, null or severe homozygous mutations of *GBA* showed little or no human glucocerebrosidase activity<sup>21</sup>. These findings suggested differences in phenotype due to the various *GBA* mutations. Kleinloog et al.<sup>22</sup> reported enrichment of the lysosomal pathway in ruptured IA compared with UIA based on RNA sequencing analysis of aneurysm wall. Although the lysosomal pathway does not reflect an acute reaction to IA rupture<sup>22</sup>, it is likely that it is induced by inflammation after bleeding.

*OLFML2A* and *TCF24* showed a protective effect against IA formation with log10BF levels greater than 12 and a PIP of 1.0. However, the relationship between these two genes and IA is still unclear, even though it



**Figure 2.** A heatmap of multiple gene expression involving *GBA*, *TCF24*, *OLFML2A*, and *ARHGAP32* in human cells and tissues including artery, brain, and whole blood is presented. Gene expression was estimated as transcripts per million (TPM). Genes and types of cells or tissues were ordered via agglomerative hierarchical clustering. An interactive heatmap specifically designed for rendering expression data was drawn by the GTEx Expression Map tool to report and summarize multi-gene and multi-tissue expressions (<https://gtexportal.org/home/multiGeneQueryPage>).

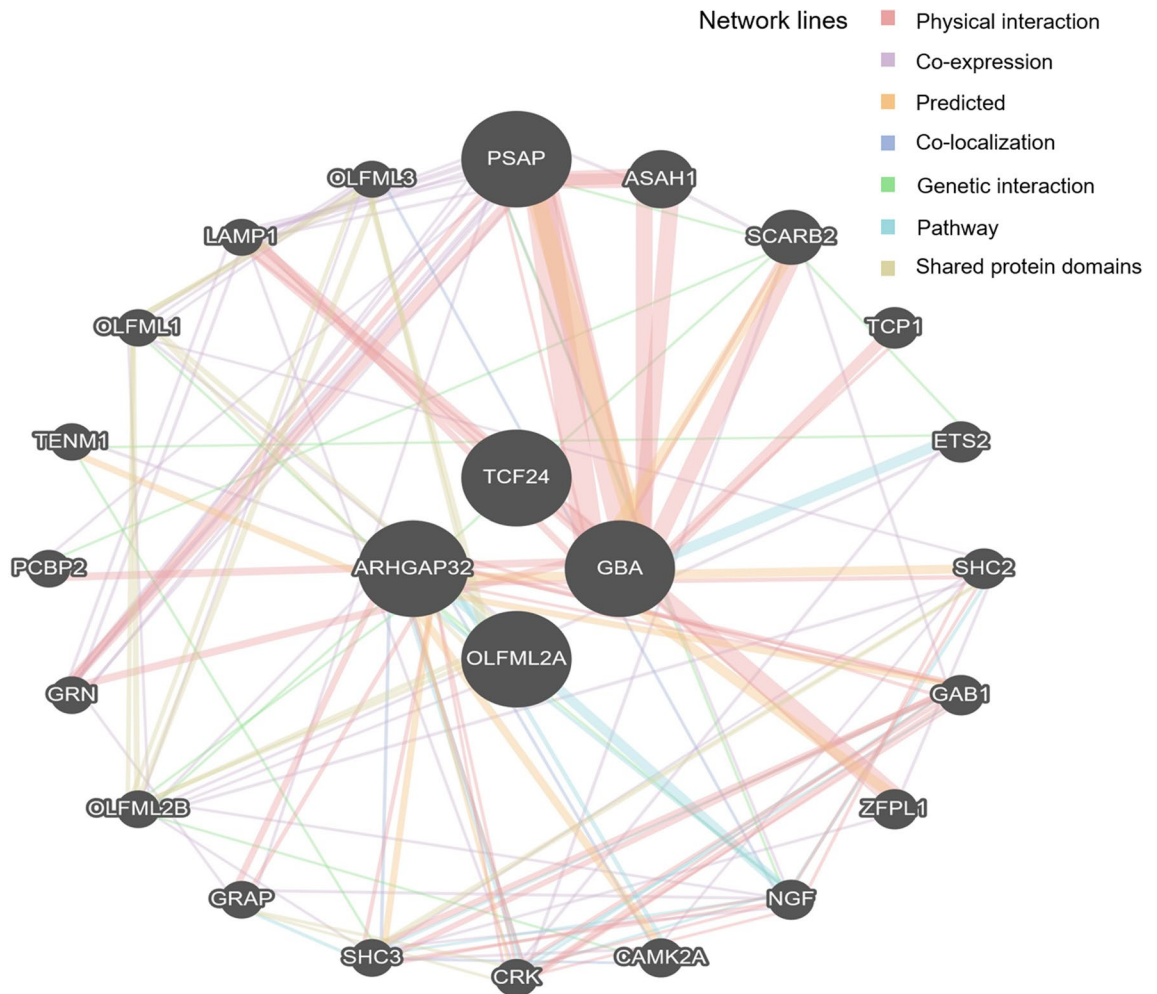
has been implicated in cardiovascular diseases. Conversely, *ARHGAP32* significantly increased the risk of IA with a  $\log_{10}BF$  of 20.88 and a PIP of 1.0. *ARHGAP32* refers to Rho GTPase-activating protein 32 and mediates *N*-methyl *D*-aspartate receptor signaling<sup>12</sup>. The role of *ARHGAP32* has been mainly investigated in the regulation of blood pressure. Rho-specific GTPase-activating protein *GRAF3* was highly expressed in smooth muscle cells (SMCs) and regulated blood pressure control by inhibiting the contractility of RhoA-mediated SMC<sup>23</sup>. *GRAF3*-deficient mice also showed increased blood pressure in response to angiotensin II and endothelin 1<sup>24</sup>. In actual clinical practice, many patients manifest both IA and hypertension. Inci et al.<sup>25</sup> reported that the rate of pre-existing hypertension was 43.5% in patients with IA, which was higher than 24.4% in the normal population. Hypertension may contribute to degeneration of the internal elastic lamina, weakening of the vessel wall, and IA formation<sup>25</sup>.

Nevertheless, it is unclear whether the role of *ARHGAP32* in IA is mediated indirectly via chronic hypertension or directly via change in vascular tone.

Functional network analyses showed that *PSAP* was an important gene in the development of IA. The role of *PSAP* gene was rarely investigated in IA and was mainly studied in PD. Oji et al.<sup>26</sup> reported that two SNPs of rs4747203 and rs885828, the intronic regions of the *PSAP* saposin D domain were linked to PD. *PSAP* mutation can also result in dopaminergic neurodegeneration and motor decline in mice. Although we did not include patients with PD, a fine-mapping analysis revealed that PD-related genes such as *GBA* and *PSAP* may contribute to IA. Lysosomal dysfunction and the resulting lysosomal storage disorder can contribute causally to PD. Putative damaging variants in at least one gene associated with lysosomal storage disorder were observed in most PD patients<sup>27</sup>. However, lysosomal dysfunction can also be observed in the arterial wall. Lysosomal changes in the vascular SMCs were attributed to the accumulation of excessive substrate levels in the lysosomes of a primate model of atherosclerosis and hypertension<sup>28</sup>. The excessive sterol accumulation in lysosomes can disrupt the lysosomal function<sup>29</sup>. Therefore, in this case, it is possible that lysosomal dysfunction may directly affect IA formation or may contribute to IA via atherosclerosis.

Hokari et al.<sup>30</sup> reported that atherosclerotic factors strongly increased the risk of middle cerebral artery aneurysm compared with paraclinoid aneurysm. After securing aneurysm, consistent statin therapy was significantly correlated with better prognosis<sup>31</sup>. Wu et al.<sup>32</sup> demonstrated that the autophagy-lysosomal pathway, which entails





**Figure 3.** Susceptibility to intracranial aneurysm (*Homo sapiens*) based on multiple protein interactions between proteins coded by four candidate hub genes including *GBA*, *TCF24*, *OLFML2A*, and *ARHGAP32*. The network included neighboring genes correlated with four hub genes. The width of individual lines indicates the intensity of the interaction between proteins. The colors in each line indicate multiple functions including physical interaction, co-expression, prediction, co-localization, genetic interaction, pathways, and shared protein domains. The multiple protein interaction map was drawn by using the GeneMANIA program (<https://genemania.org/>).

self-digestion of dysfunctional intracellular components by lysosomal enzymes, was an important pro-survival mechanism after SAH. However, the study investigated the role of the lysosome after SAH development, but not in IA development itself. Thus, additional studies are needed to investigate the role of candidate variants in lysosomal dysfunction resulting in IA formation via abnormal ECM remodeling in response to hemodynamic stress.

Large-scale GWAS and meta-analyses have been widely used to identify and validate common or novel susceptible gene variants in various medical diseases over the past decade. However, given the overall genotype–phenotype analyses, disease-modifying functional mutations and direct biological relevance to disease have yet to be elucidated completely<sup>33</sup>. In addition, the heritability of a specific trait cannot be fully explained by common SNPs of intronic or intergenic regions via GWAS because GWAS analysis was designed to identify common variants with low and modest effect size, which contribute to disease. Accordingly, even if a large number of susceptible loci were identified, a few cases showed their replication in an independent cohort. Thus, few disease-associated variants have been demonstrated in functional *in vitro* studies or used in treatment<sup>34</sup>. To overcome these limitations, an updated fine-mapping analysis was performed to identify the candidate variants associated with complex human diseases and as a cost-effective genotyping strategy<sup>9</sup>. To date, many studies underscored the need for ‘feature selection’ to identify relevant “variables” using parametric or non-parametric models. However, feature selection is not a simple challenge and requires substantial genetic investigations. It is important to identify the driver mutations linked to treatment of complex human diseases. The selection of genetic variants from GWAS is uncertain given the strongly correlated SNPs corresponding to a pairwise LD structure at the population level. A fine-mapping analysis facilitates the identification of creditable genetic variants to refine the selection bias such as false-positive variants based on the initial GWAS and to improve the findings of molecular functional studies<sup>9</sup>. Here, we performed a fine-mapping analysis based on the results of

previous IA GWAS using the statistical method developed by Benner et al.<sup>14</sup> Our findings may enable the identification of candidate variants via a pairwise LD structure and exclude potential false positives via statistically significant fine-mapping analysis of transformed GWAS results. Therefore, these analytical methods may enable the selection of functional candidate variants based on the molecular mechanisms associated with IA formation.

Our fine-mapping identified the most likely variant causality among candidate SNPs included in the analysis; however, the actual disease-associated variant might be nearby or highly located at another locus tagged by haplotype structures accompanied by ‘LD-tower’ in most GWAS. The caution is required to interpret our results through more cross validations, though this is the first fine-mapping analysis based on results of a GWAS of IA to reduce IA-associated variants with false positive. Fine-mapping analysis is inevitably affected by the original datasets. Accordingly, small sample size of the GWAS data without replication study is a concern. In particular, the effect size for the identified SNPs was extremely large. Moreover, the identified significant SNPs were not accompanied by correlated SNP via LD<sup>12</sup>. Since these issues are most likely due to the small sample size, a study based on a large number of patients is needed to address these issues. However, we cannot perform the fine-mapping analysis further based on larger patient data than the current data for two reasons. First, we performed the fine-mapping analysis by a real linkage disequilibrium (LD) based on a real genotype dataset.

Although recently large-scale GWAS meta-analysis results have been published<sup>35</sup>, we cannot obtain a real original dataset or a whole summary statistics and genotype of studies included. Second, we genotyped samples using the APMRA for the first stage GWAS. Accordingly, we identified several novel SNPs that were not found in other GWASs based on Caucasian-based chip array. However, paradoxically, this can be a limitation when comparing data results of other papers. Actually, we analyzed 17 SNPs which showed association with IA in previous genetic studies of IA, including the APMRA used in this study. Of these 17 SNPs, only two SNPs (*BOLL* and *ENDRA*) are associated with IA in Koreans<sup>12</sup>. For these two reasons, we could not perform the fine mapping analysis based on the large dataset. Currently, we are performing IA GWAS as a replication study by analyzing the currently ongoing validation dataset for 50 independent patients with IA and independent controls of hospitals and the Rural and Mid-size City cohort of the Korean Genome Epidemiology Study to validate 29 previous GWA signals showing LD < 0.8<sup>12</sup>. Unfortunately, most candidate SNPs revealed by fine-mapping analysis based on the initial GWAS did not surpass the significance threshold of IA association in replication stages. Therefore, caution is required when interpreting our results because a small number of IA patients were included in our independent replication studies. In addition, some SNPs could not be fully analyzed due to difference in the GWAS panel used.

In summary, fine-mapping analysis robustly identified four functional mutations of candidate genes (*GBA*, *TCF24*, *OLFML2A*, and *ARHGAP32*) associated with IA. Mutations in these genes may play roles in immune and inflammatory systems according to our literature review and functional annotations. At the present, our strategy will offer a good example after GWAS regardless of negative or positive outcomes. For the next step, our findings will provide a milestone of fine-mapping susceptibility to IA development.

## Data availability

The data that support the findings of this study was submitted as online supplemental material, and further detailed information is available upon request to the corresponding author. All genotype and phenotype resources are managed by “The First Korean Stroke Genetics Association Research” study constructed from the Sacred Heart Hospital Stroke Database.

Received: 3 August 2021; Accepted: 4 February 2022

Published online: 17 February 2022

## References

1. Tromp, G., Weinsheimer, S., Ronkainen, A. & Kuivaniemi, H. Molecular basis and genetic predisposition to intracranial aneurysm. *Ann. Med.* **46**, 597–606 (2014).
2. Asaithambi, G., Adil, M. M., Chaudhry, S. A. & Qureshi, A. I. Incidences of unruptured intracranial aneurysms and subarachnoid hemorrhage: Results of a statewide study. *J. Vasc. Interv. Neurol.* **7**, 14–17 (2014).
3. Zhou, S., Dion, P. A. & Rouleau, G. A. Genetics of intracranial aneurysms. *Stroke* **49**, 780–787 (2018).
4. 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).
5. Boehme, A. K., Esenwa, C. & Elkind, M. S. Stroke risk factors, genetics, and prevention. *Circ. Res.* **120**, 472–495 (2017).
6. Jeon, J. S. et al. A retrospective analysis on the natural history of incidental small paraclinoid unruptured aneurysm. *J. Neurol. Neurosurg. Psychiatry.* **85**, 289–294 (2014).
7. Hackenberg, K. A. M., Hanggi, D. & Etminan, N. Unruptured intracranial aneurysms. *Stroke* **49**, 2268–2275 (2018).
8. Freebern, E. et al. Gwas and fine-mapping of livability and six disease traits in holstein cattle. *BMC. Genom.* **21**, 41. <https://doi.org/10.1186/s12864-020-6461-z> (2020).
9. Schaid, D. J., Chen, W. & Larson, N. B. From genome-wide associations to candidate causal variants by statistical fine-mapping. *Nat. Rev. Genet.* **19**, 491–504 (2018).
10. Benner, C. et al. Prospects of fine-mapping trait-associated genomic regions by using summary statistics from genome-wide association studies. *Am. J. Hum. Genet.* **101**, 539–551 (2017).
11. Sekar, A. et al. Schizophrenia risk from complex variation of complement component 4. *Nature* **530**, 177–183 (2016).
12. Hong, E. P. et al. Genomic variations in susceptibility to intracranial aneurysm in the Korean population. *J. Clin. Med.* **8**, 275. <https://doi.org/10.3390/jcm8020275> (2019).
13. Kim, B. J. et al. Genome-wide blood DNA methylation analysis in patients with delayed cerebral ischemia after subarachnoid hemorrhage. *Sci. Rep.* **10**, 11419. <https://doi.org/10.1038/s41598-020-68325-3> (2020).
14. Benner, C. et al. Finemap: Efficient variable selection using summary data from genome-wide association studies. *Bioinformatics* **32**, 1493–1501 (2016).
15. Chang, C. C. et al. Second-generation plink: Rising to the challenge of larger and richer datasets. *Gigascience.* **4**, 7. <https://doi.org/10.1186/s13742-015-0047-8> (2015).
16. Pruim, R. J. et al. Locuszoom: Regional visualization of genome-wide association scan results. *Bioinformatics* **26**, 2336–2337 (2010).

17. Wang, K., Li, M. & Hakonarson, H. ANNOVAR: Functional annotation of genetic variants from high-throughput sequencing data. *Nucleic. Acids. Res.* **38**, e164. <https://doi.org/10.1093/nar/gkq603> (2010).
18. GTEx Consortium. Genetic effects on gene expression across human tissues. *Nature* **550**, 204–213 (2017).
19. Warde-Farley, D. *et al.* The genemania prediction server: Biological network integration for gene prioritization and predicting gene function. *Nucleic. Acids. Res.* **38**, W214–220 (2010).
20. Mata, I. F. *et al.* Gba variants are associated with a distinct pattern of cognitive deficits in Parkinson's disease. *Mov. Disord.* **31**, 95–102 (2016).
21. Beutler, E., Gelbart, T. & Scott, C. R. Hematologically important mutations: Gaucher disease. *Blood Cells Mol. Dis.* **35**, 355–364 (2005).
22. Kleinloog, R. *et al.* Rna sequencing analysis of intracranial aneurysm walls reveals involvement of lysosomes and immunoglobulins in rupture. *Stroke* **47**, 1286–1293 (2016).
23. Dee, R. A., Bai, X., Mack, C. P. & Taylor, J. M. Molecular regulation of the rhogap graf3 and its capacity to limit blood pressure in vivo. *Cells* **9**, 1042. <https://doi.org/10.3390/cells9041042> (2020).
24. Bai, X. *et al.* The smooth muscle-selective rhogap graf3 is a critical regulator of vascular tone and hypertension. *Nat. Commun.* **4**, 2910. <https://doi.org/10.1038/ncomms3910> (2013).
25. Inci, S. & Spetzler, R. F. Intracranial aneurysms and arterial hypertension: A review and hypothesis. *Surg. Neurol.* **53**, 530–540 (2000) (**discussion 540–532**).
26. Oji, Y. *et al.* Variants in saposin d domain of prosaposin gene linked to Parkinson's disease. *Brain* **143**, 1190–1205 (2020).
27. Robak, L. A. *et al.* Excessive burden of lysosomal storage disorder gene variants in Parkinson's disease. *Brain* **140**, 3191–3203 (2017).
28. Wolinsky, H., Goldfischer, S., Daly, M. M., Kasak, L. E. & Coltoff-Schiller, B. Arterial lysosomes and connective tissue in primate atherosclerosis and hypertension. *Circ. Res.* **36**, 553–561 (1975).
29. Jerome, W. G. Lysosomes, cholesterol and atherosclerosis. *Clin. Lipidol.* **5**, 853–865 (2010).
30. Hokari, M. *et al.* The impact of atherosclerotic factors on cerebral aneurysm is location dependent: Aneurysms in stroke patients and healthy controls. *J. Stroke Cerebrovasc. Dis.* **23**, 2301–2307 (2014).
31. Song, T. J., Oh, S. H. & Kim, J. The impact of statin therapy after surgical or endovascular treatment of cerebral aneurysms. *J. Neurosurg.* <https://doi.org/10.3171/2019.3.JNS183500> (2019).
32. Wu, H. *et al.* The autophagy-lysosomal system in subarachnoid haemorrhage. *J. Cell Mol. Med.* **20**, 1770–1778 (2016).
33. Tam, V. *et al.* Benefits and limitations of genome-wide association studies. *Nat. Rev. Genet.* **20**, 467–484 (2019).
34. Lohmueller, K. E., Pearce, C. L., Pike, M., Lander, E. S. & Hirschhorn, J. N. Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. *Nat. Genet.* **33**, 177–182 (2003).
35. Bakker, M. K. *et al.* Genome-wide association study of intracranial aneurysms identifies 17 risk loci and genetic overlap with clinical risk factors. *Nat. Genet.* **52**, 1303–1313 (2020).

## Acknowledgements

This research is supported by the grant of the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (Grant number: HR21C0198) and Hallym Research Fund. For collecting controls for independent stage 4, this study was conducted using bioresources from National Biobank of Korea, the Korea Disease Control and Prevention Agency, Republic of Korea (KBN-2021-057).

## Author contributions

E.P.H. and J.P.J. designed and managed this study. Analysis was done by E.P.H. and B.J.K. Sample preparation and data collection were done by D.H.Y., B.J.K., E.P.H. and J.P.J. Drafting and reviewing of manuscript were done by E.P.H., D.H.Y., B.J.K., J.H.A., J.J.P., J.K.R. H.C.K., G.H., H.J.J., and J.P.J.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-06755-x>.

**Correspondence** and requests for materials should be addressed to J.P.J.

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

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



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

© The Author(s) 2022