Novel Genome-wide Interactions Mediated via *BOLL* and *EDNRA* Polymorphisms in Intracranial Aneurysm

Eun Pyo Hong

Hallym University College of Medicine

Dong Hyuk Youn

Hallym University College of Medicine

Bong Jun Kim

Hallym University College of Medicine

Jae Jun Lee

Hallym University College of Medicine

Sehyeon Nam

Hanyang University

Hyojong Yoo

Hanyang University

Jeong Jin Park

Konkuk Medical Center

Yong Jun Cho

Hallym University College of Medicine

Jin Pyeong Jeon (✉ jjjs6553@daum.net)

Hallym University College of Medicine

Research Article

Keywords: Boule, Endothelin receptor type A, Genome-wide association, Genome-wide Interaction, Intracranial aneurysm

Posted Date: December 2nd, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-1067973/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Association of *boule* (*BOLL*) and endothelin receptor type A (*EDNRA*) loci with intracranial aneurysm (IA) formation has been reported via genome-wide association studies. However, the underlying genome-wide interactions have yet to be reported. We sought to identify genome-wide interactions involving *BOLL* and *EDNRA* loci for IA. Genome-wide interaction analyses of *BOLL* and *EDNRA* involving 250 IA patients and 296 controls were performed under an additive effect model. Subsequent gene expression analyses were conducted using transcripts per million (TPM). A total of 23 and 11 SNPs suggested a genome-wide threshold ($p < 1.25 \times 10^{-8}$) interacting with rs700651 (*BOLL*) and rs6841581 (*EDNRA*), respectively. The rs1105980 (*PTCH1*) showed the most significant interaction with rs700651 ($p = 6.41 \times 10^{-11}$). The rs74585958 (*RYK*) interacted strongly with rs6841581 ($p = 1.64 \times 10^{-9}$). The *BOLL*-interacting *CXCR4* was highly overexpressed in whole blood (TPM = 419.8) and *CCDC3* was overexpressed in all artery-related tissues (TPM = 315.4 to 473.9). *EDNRA*-interacting *EIF4H* showed a comprehensively elevated expression across all tissues and cells (TPM = 85.8 to 372.0). Genome-wide interaction study shows that *BOLL* and *EDNRA* may contribute to IA formation by interacting with multiple genes in cardio-metabolic pathway. Our findings may provide insight into the functional relevant to IA susceptibility.

Introduction

Intracranial aneurysm (IA) refers to a bulge in the wall of intracranial arteries due to endothelial dysfunction and extracellular matrix remodeling of the hemodynamic response. Although the prevalence of IA is approximately 3% in the general population, the mortality rate due to subarachnoid hemorrhage (SAH) following aneurysm rupture is close to 50%^{1,2}. Clinical and radiological studies showed that IA formation and rupture were related to female gender, hypertension, smoking, larger size and posterior circulation aneurysm³.

The plethora of genome-wide association studies (GWASs) during late 2000's has increased the number of investigations into IA. These genetic studies have identified several candidate genes and loci associated with IA such as *EDNRA*, *GBA*, *CDKN2A/B*, *RBBP8*, *STARD13/KL*, and *SOX17*⁴⁻⁷. GWAS reported differences in allele frequencies of single-nucleotide polymorphisms (SNPs) based on multi-ethnic case-control studies. Accordingly, the missing heritability is explained by independent SNPs involving complex diseases in human including cerebrovascular diseases. To address this issue, robust analytical approaches such as meta-analyses, multifactorial interactions, and polygenic risk scoring systems have been performed in many populations⁸. A GWA meta-analysis reported IA-associated risk loci, including new loci (*SLC22A5*, *6q16.1*, *12q21.33*, *PSMA4*, and *NT5C2*) based on 10,754 cases and 306,882 controls⁷. Two large-scale meta-analyses also reported the successful validation of two IA-associated loci *BOLL* and *EDNRA* in an East-Asian population of nearly 20,000 individuals^{9,10}. However, few studies have reported gene-gene interactions or polygenic risk scores for IA patients. Considering to two previous findings of *BOLL* and *EDNRA* loci, here we estimated the effects of genome-wide gene-gene interactions on IA in a prospective hospital-based cohort study. Also, functional analyses in human tissues and cell lines were conducted to determine the genetic expression profile.

Materials And Methods

Study populations

The study subjects were enrolled from the multi-institutional biobanks comprising five university hospitals constituting "The First Korean Stroke Genetics Association Research", and including patients diagnosed with cerebrovascular diseases (CVDs) such as IA between March 2015 and December 2020^{5,11}. Data derived from 250 patients with IA and 296 controls, which were also used in the first Korean IA GWAS⁵, were used in the analysis. The inclusion of patients with IA was based on the following criteria: 1) adult patients more than 18 years of age, 2) patients without other types of CVD such as ischemic stroke, hemorrhagic stroke, and vascular malformation; and 3) patients without any other genetic disorders such as polycystic kidney and moyamoya disease. Control subjects were defined as adults without CVD. Medical and radiological data were collected and updated. All methods were carried out in accordance with relevant guidelines and regulations. All protocols of this study including all patients and other subjects providing written informed consents was approved by the Institutional Review Board of the Hallym University Chuncheon Sacred Heart Hospital (No. 2016-3, 2019-06-006). Initial design and protocol of this study is described in elsewhere^{5,11}.

Genotyping and quality controls

Genomic DNA derived from the peripheral blood of the study population was genotyped using the AxiomTM Asian Precision Medicine Research Array (APMRA) (Thermo Fisher Scientific, MA, USA). High quality plates were defined by a plate pass rate higher than 95% for samples and the average call rate of passing samples was greater than 99%. Out of 798,148 SNPs, 512,575 SNPs passed the quality control including genotyping call rate of 95% or higher, minor allele frequency (MAF) of at least 1%, and Hardy-Weinberg equilibrium (HWE) with P value $\geq 1 \times 10^{-6}$.

Genome-wide SNP-SNP interactions via *BOLL* and *EDNRA* genes

We investigated genome-wide SNP-SNP interactions using either *BOLL* or *EDNRA* loci associated with IA in previous GWASs and meta-analyses^{5,9,10}. The multivariate analyses of SNP interactions between rs700651 intron SNP (*BOLL*, 2q33.1) and rs6841581 upstream SNP (*EDNRA*, 4q31.22) on 512,574 SNPs were performed using the Contrived Acronym of Software for SNP Interactions (CASSI v2.5)¹². SNP-SNP

interactions were analyzed by choosing the target SNP from two given SNP windows (possibly from different pedigree files). Each pair of SNPs that passed the interaction test at a given significance level (i.e., minimized p -value = 1) was returned in the output file with possible extra information such as beta coefficient, standard error, chi-squared, and p -value. The CASSI accepted only PLINK binary files (filename.bed) in order to perform the calculations as efficiently as possible. The logistic regression epistasis test was available for the SNP-SNP interactions in this study even though this program provides the maximum number of one millions terms. Subsequent regression analyses were carried out under the additive effect model and adjusted for 10 covariates including age, sex, hypertension, diabetes mellitus, hyperlipidemia, smoking status, and four genetic ancestry factors. The effect coefficient was estimated using the natural log-transformed scale of odds ratio (i.e., $\ln(\text{OR})$). The multiple comparisons were adjusted for a genome-wide interaction threshold with a p -value less than 1.25×10^{-8} (genome-wide p -value 5×10^{-8} divided by 4 tests for interaction). Subsequent power and sample size calculations for each interaction term were estimated under the assumption of a GW interaction significance threshold, a 1:1.2 case-control ratio, and information on each SNP (i.e., allele frequency and effect size) using the Quanto v2.4 program (<https://pphs.usc.edu/download-quanto/>). Large-scale interaction analysis was completed using a multi-tasking super-computer with an Intel(R) Xeon(R) CPU E5-2667 v4 (3.20GHz), 256 GB RAM, and 16 cores. We performed Manhattan plots of *BOLL* and *EDNRA* interaction using the package of “*qqman*” in R v3.6.1 (<https://cran.r-project.org/web/packages/qqman>) and regional visualizations of the target SNP's base-pair position (BP) $\pm$ 400 kb regions using LocusZoom v1.3 written in modified Python and R scripts¹³.

Gene expression analysis

Gene expression was analyzed to evaluate the human tissue- or cell-specific expression of candidate genes based on Genotype-Tissue Expression (GTEx) Portal (<https://gtexportal.org/home/>)¹⁴. Expressions in 56,200 genes were estimated in the 13 brain tissues, 4 blood vessels (3 artery tissues and 1 cell line of EBV-transformed lymphocytes), and a whole blood cell. 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 expression (<https://gtexportal.org/home/multiGeneQueryPage>). Each gene expression in specific human cells or tissues was validated in the independent dataset, the Human Protein Atlas (HPA, <https://www.proteinatlas.org/>)¹⁵. Gene expression levels were calculated by applying the transcripts per million (TPM) method and normalized by log-transformed (TPM+1) values (page, line, highlighted).

In silico gene ontology and network analyses

Gene ontology (GO) enrichment analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway were used to evaluate the gene functional annotation of *BOLL* and *EDNRA*, and their interacting genes using the DAVID Bioinformatics Resources v6.8 (<https://david.ncicrf.gov/>)¹⁶. The gene similarity and multiple linkage thresholds were 0.5. We first filtered GO terms and/or KEGG pathways of gene clusters with p -value less than 0.05. We performed a network analysis of protein-coding interactions based on the results of functional annotation using the STRING v11.0 (<https://string-db.org/>)¹⁷.

Results

Genome-wide SNP-SNP interaction with *BOLL* (rs700651) and *EDNRA* (rs6841581) loci

The number of SNPs per gene is presented in Supplementary Table S1. Out of 512,574 SNP interaction terms, 23 and 11 SNPs reached a genome-wide interaction threshold ($p < 1.25 \times 10^{-8}$) with rs700651 intron SNP (*BOLL*, 2q33.1) and rs6841581 upstream SNP (*EDNRA*, 4q31.22), respectively (Figure 1 and Table 1). All 34 SNPs showed an MAF above 1% and an HWE p -value greater than 0.01 (Supplementary Table S2). Two target SNPs of rs700651 and rs6841581 common SNPs with MAF > 10% have been replicated in the East-Asian population including Koreans^{5,9,10}. Most of the *BOLL*- or *EDNRA*-interacting SNPs showed shared alleles in both patient and control groups (i.e., MAF > 0.23) without significant association in a single SNP analysis ($0.01 < p < 1$) (Table 1). These findings suggested that *BOLL* or *EDNRA* rarely affect IA formation at the single SNP level, but may have a synergistic effect on multifactorial pair-wise interactions. Among the interactions, the rs1105980 upstream SNP of *PTCH1* gene (9q22.32) showed the most significant interaction with rs700651 ($\ln\text{OR} = 1.53$, $p = 6.41 \times 10^{-11}$) (Figure 2A). The rs74585958 of *RYK* gene (3q22.2) interacted strongly with rs6841581 ($\ln\text{OR} = -19.91$, $p = 1.64 \times 10^{-9}$) (Figure 2B). Two strong pair-wise linkage disequilibrium were observed in the interaction of rs328025 with rs700855 (*RGP4*, 2q12.3) and between rs11925024 and rs1231 (*TNIF*, 3q26.31) ($r^2 > 0.95$, data not shown). Interestingly, two *EDNRA*-interacting gene variants of *TNIF* (rs11925024 and rs1231, $p = 1.04 \times 10^{-8}$ and 1.22×10^{-9} , respectively) and *FTO* (rs9302654, $p = 3.78 \times 10^{-9}$), and one *BOLL*-interacting *METTL4* gene variant (rs549315, $p = 4.80 \times 10^{-10}$) showed marginal interaction with *BOLL* gene ($0.001 < p < 0.05$). However, there was no direct interaction between *BOLL* and *EDNRA* variants ($\ln\text{OR} = -0.27$, $p = 0.301$). Additionally, none of the SNPs had a shared effect size mediated by *BOLL* and *EDNRA* variants. When the power calculation was estimated based on the detailed SNP information (Supplementary Table S2) and association results (Table 1), two SNPs rs74585958 (*RYK*) and rs150664966 (*EIF4H*) showed sufficient statistical power among the *EDNRA*-interacting loci (88.5% and 80.5%, respectively). The rs11672303 (*SLC7A10*) SNP in the *EDNRA*-interacting loci showed a marginal statistical power of 78.9%. In contrast, no SNP reached sufficient statistical power threshold of 80% in interaction terms by the *BOLL* locus (i.e., power < 50%). Detailed information on statistical power and sample size (at 80% power) is shown in Supplementary Table S3.

Table 1
Genome-wide interaction terms by *BOLL* or *EDNRA* polymorphisms on intracranial aneurysm

Gene	Chr	SNP	BP	M/m ^a	MAF, ca/co ^a	lnOR ^b	P for GWAS ^b	lnOR ^c	P by <i>BOLL</i> ^c	lnOR ^d	P by <i>EDNRA</i> ^d
<i>BOLL</i>	2q33.1	rs700651	198631714	A/G	0.476/0.449	1.42	0.0079	NA	NA	-0.27	0.3010
<i>EDNRA</i>	4q31.22	rs6841581	148401190	A/G	0.13/0.217	0.53	0.0006	-0.27	0.3010	NA	NA
<i>LINC01344</i>	1q25.3	rs12033118	182229747	C/T	0.022/0.027	-0.15	0.7192	-19.75	5.00E-09	0.79	0.2679
<i>RGPD4</i>	2q12.3	rs328025	108355045	G/A	0.377/0.324	0.31	0.0219	1.25	1.06E-09	-0.19	0.4897
<i>RGPD4</i>	2q12.3	rs700855	108368694	T/C	0.372/0.328	0.27	0.0454	1.31	9.91E-10	-0.18	0.5209
<i>MALL</i>	2q13	rs117802391	110862084	C/T	0.036/0.061	-0.77	0.0198	19.57	1.06E-09	-1.58	0.0762
<i>CXCR4</i>	2q22.1	rs189432614	136809235	A/G	0.016/0.025	-1.63	0.0131	-34.81	7.59E-09	1.02	0.2116
<i>RPRM</i>	2q23.3	rs5005908	154003680	G/T	0.344/0.27	0.32	0.0220	1.26	3.03E-09	-0.13	0.6388
<i>RBMS3</i>	3p24.1	rs1979271	29607405	T/A	0.406/0.39	0.07	0.5841	-1.17	6.47E-09	0.13	0.6541
<i>FOXP1</i>	3p13	rs878118	71246228	T/G	0.238/0.255	-0.15	0.3230	-1.39	3.85E-09	0.25	0.4035
<i>LINC00879</i>	3q11.2	rs4411883	94549686	T/G	0.09/0.111	-0.34	0.1404	-1.91	1.24E-08	-0.37	0.4182
<i>RYK</i>	3q22.2	rs74585958	133773362	G/A	0.054/0.041	0.57	0.0817	-0.21	0.6852	-19.91	1.64E-09
<i>TNIK</i>	3q26.31	rs11925024	171014067	A/C	0.145/0.151	-0.15	0.4431	-0.90	0.0008	-2.71	1.04E-08
<i>TNIK</i>	3q26.31	rs1231	171031233	A/T	0.144/0.154	-0.18	0.3537	-0.80	0.0021	-2.86	1.22E-08
<i>EIF2B5</i>	3q27.1	rs4350902	184352200	T/C	0.472/0.492	-0.09	0.5253	-1.18	8.83E-09	-0.01	0.9793
<i>RUFY1</i>	5q35.3	rs4075890	178997373	T/C	0.216/0.2	0.10	0.5444	1.46	8.47E-09	-0.21	0.5058
<i>C5orf60</i>	5q35.3	rs62405726	179069468	G/A	0.318/0.287	0.13	0.3646	1.46	4.21E-10	-0.02	0.9398
<i>RREB1</i>	6p24.3	rs9505086	7232186	T/C	0.286/0.307	-0.11	0.4279	1.24	1.20E-09	0.06	0.8250
<i>DST</i>	6p12.1	rs117021265	56628021	T/C	0.034/0.024	0.25	0.5494	-18.96	1.75E-09	0.16	0.8283
<i>EIF4H</i>	7q11.23	rs150664966	73594157	T/C	0.016/0.022	-0.17	0.7152	-0.05	0.9384	20.91	4.80E-09
<i>MPDZ</i>	9p23	rs1332064	12942764	T/C	0.354/0.309	0.19	0.1724	0.10	0.5916	1.75	5.10E-09

GWAS, genome-wide association study; lnOR, Natural log-transformed odds ratio; MAF, minor allele frequency.

^a M/m, major/minor allele type; MAF in case and control groups, respectively (ca/co).

^b lnORs and p-values were estimated by generalized linear model after adjusting for age, sex, hypertension, diabetes, hyperlipidemia, smoking status in the previous GWAS.

^c lnORs and p-values were estimated after *BOLL* by 500K SNPs interactions by performing CASSI program, adjusted for age, sex, hypertension, diabetes, hyperlipidemia, and smoking status.

^d lnORs and p-values were estimated after *EDNRA* by 500K SNPs interactions by performing

Gene	Chr	SNP	BP	M/m ^a	MAF, ca/co ^a	InOR ^b	P for GWAS ^b	InOR ^c	P by BOLL ^c	InOR ^d	P by EDNRA ^d
<i>PTCH1</i>	9q22.32	rs1105980	98113635	G/C	0.27/0.294	-0.08	0.5821	1.53	6.41E-11	-0.60	0.0580
<i>PFKF</i>	10p15.2	rs58183624	3107217	C/T	0.066/0.044	0.15	0.5920	-2.68	9.52E-09	0.56	0.3378
<i>CCDC3</i>	10p13	rs12412014	12911725	G/C	0.281/0.291	-0.09	0.5145	1.47	6.63E-11	-0.14	0.6273
<i>CACUL1</i>	10q26.11	rs11198727	120767097	A/G	0.382/0.429	-0.13	0.3308	0.17	0.3558	1.72	8.06E-09
<i>PLEKHA1</i>	10q26.13	rs10510110	124192430	C/T	0.399/0.372	0.17	0.2098	1.26	8.96E-09	-0.26	0.3350
<i>TRIM22</i>	11p15.4	rs7480654	5722839	T/C	0.317/0.284	0.18	0.2200	-1.21	9.59E-09	-0.26	0.3718
<i>SAP18</i>	13q12.11	rs9509543	21692404	C/T	0.346/0.356	-0.05	0.7050	-0.20	0.2731	1.85	2.87E-09
<i>LINC00457</i>	13q13.2	rs1536847	35106975	G/T	0.295/0.329	-0.18	0.2033	1.39	2.37E-10	0.25	0.3692
<i>UNC13C</i>	15q21.3	rs4774715	55140204	C/T	0.432/0.441	-0.01	0.9362	-0.09	0.6185	-1.68	7.74E-09
<i>FTO</i>	16q12.2	rs9302654	54009545	C/T	0.114/0.144	-0.34	0.0880	0.58	0.0334	-3.10	3.78E-09
<i>CDH13</i>	16q23.3	rs3848296	82550548	G/A	0.192/0.231	-0.16	0.3193	1.46	7.69E-09	-0.31	0.3982
<i>SLFN11</i>	17q12	rs77814639	33678827	A/G	0.184/0.153	0.28	0.1329	-0.43	0.1080	-18.50	4.48E-09
<i>LINC01978</i>	17q25.3	rs57851800	77896371	A/C	0.345/0.326	0.16	0.2657	-1.27	1.08E-09	-0.29	0.3135
<i>METTL4</i>	18p11.32	rs549315	2183055	G/A	0.378/0.429	-0.29	0.0389	-1.32	4.80E-10	0.59	0.0330
<i>SLC7A10</i>	19q13.11	rs11672303	33726375	T/C	0.171/0.154	0.08	0.6556	0.23	0.3238	2.16	3.55E-09
GWAS, genome-wide association study; InOR, Natural log-transformed odds ratio; MAF, minor allele frequency.											
^a M/m, major/minor allele type; MAF in case and control groups, respectively (ca/co).											
^b InORs and p-values were estimated by generalized linear model after adjusting for age, sex, hypertension, diabetes, hyperlipidemia, smoking status in the previous GWAS.											
^c InORs and p-values were estimated after BOLL by 500K SNPs interactions by performing CASSI program, adjusted for age, sex, hypertension, diabetes, hyperlipidemia, and smoking status.											
^d InORs and p-values were estimated after EDNRA by 500K SNPs interactions by performing											

Gene expression involving candidate genes interacting with BOLL and EDNRA genes

BOLL-interacting *CXCR4* (2q22.1) was highly enriched in whole blood (TPM = 419.8) and *CCDC3* (10p13) was predominant in all arterial tissues (TPM = 315.4 to 473.9) compared with little or no expression of *BOLL* gene (Figure 3A). In contrast to *BOLL* (TPM < 1), *CXCR4* expression in diverse human whole blood cells including monocytes, T cells, B cells, and NK cells was consistently validated in the independent dataset of the HPA portal (TPM for *CXCR4* = min 292.5 (non-classical monocyte) to max 2957.7 (naive B-cell)) (Supplementary Figure S1A and S1B, respectively)¹⁸. The expression of *EDNRA*-interacting *EIF4H* (7q11.23) was comprehensively elevated across brain tissues and blood-related cell lines obtained from both GTEx- and HPA-driven datasets (Figure 3B; Supplementary Figure S2). Comprehensive gene expression based on GTEx datasets is presented in Figure 3 and Supplementary Table S4.

Gene ontology and network analyses of candidate genes mediated by BOLL and EDNRA interactions

A total of 37 GO terms were statistically enriched among gene sets comprising 19 and 11 candidate genes including *BOLL* and *EDNRA* ($p < 0.05$) (Supplementary Table S5). A positive regulation of biosynthetic process (GO:0009891) was the most enriched GO in the cluster of eight genes including *EIF2B5*, *BOLL*, *PTCH1*, *CDH13*, *CCDC3*, *RREB1*, *RBMS3*, *FOXP1* (enrichment = 4.2, $p = 0.0011$). The *PTCH1* gene mostly associated with *BOLL* interaction in IA formation contributed to diverse biological processes including regulation, migration, and membrane-related functions at the molecular or cellular levels. For *EDNRA*, positive regulation of phosphate metabolism was the most enriched ($p = 0.0214$) (Supplementary Table S5). Four long intergenic non-coding (LINC) genes including *LINC01344*, *LINC00879*, *LINC00457* and *LINC01978* were filtered in the network analysis. Six genes, *PTCH1*, *CXCR4*, *CDH13*, *RREB1*, *FOXP1*, and *TRIM22*, constituted the hub network with *BOLL* gene (Figure 4A), whereas four genes including *EDNRA*, *RYK*, *TNIF*, and *CACUL1* did not (Figure 4B).

Discussion

Although *BOLL* gene, the G allele of rs700651 associated with risk showed a significant genome-wide overall effect on IA in multi-ethnic integrative meta-analyses ($p = 1.05 \times 10^{-8}$)⁹, its pathogenic mechanism remains to be identified. In our study, the protein 'boule' was rarely expressed in the brain or the blood. Most studies related to *BOLL* involved spermatogenesis due to its role in germ cell development or cancer^{19,20}. *BOLL* is a well-known gene associated with normal germ cell development²¹. The gene is predominantly expressed in secondary spermatocytes²². However, few studies investigated the role of *BOLL* in CVD including IA. Harrod et al. reported that estrogen deficiencies may lead to IA by interrupting the inflammatory response²³. In reality, earlier age at menopause increased the risk of IA, suggesting the association between estrogen deficiency and IA pathogenesis²⁴. Thus, in the case of *BOLL* gene, additional studies are needed to determine the protective effect of male hormones against IA development. The *EIF2B5* gene that interacts with *BOLL* in the current GW interaction and network analyses exhibited a homologous inhibition of cell translation. The differential expression of *EIF2B5* was moderate in human tissues and cells. Brady et al. reported that intron retention in *EIF2B5* inhibited protein translation in hypoxic cancer cells²⁵. In the case of abdominal aortic aneurysms (AAA), there was no meaningful *EIF2B5* network²⁶. Based on a review of the current literature, it is difficult to elucidate the contribution of the two genes to IA pathogenesis. Based on our findings suggesting *EIF2B5* and *BOLL* gene clustering, a further *in vivo* study is needed to investigate the function of the two genes in IA formation.

Other *BOLL*-interacting genes such as *CXCR4*, *RPRM*, *FOXP1*, *RREB1*, *PTCH1*, *TRIM22*, and *CDH13* constitute the hub network. Among them, *CXCR4* is the most connected gene. *CXCR4* is expressed in neurons, astrocytes, and microglial cells²⁷. *CXCR4* showed enhanced immune response in the peri-infarct area. Wu et al. reported the neuroprotective effect of *CXCR4* antagonist in a rodent model of stroke²⁹. *CXCR4* is one of the cysteine-X-cysteine chemokine ligands mediating transmembrane signaling of *CXCR12*²⁹. A higher expression of *CXCR4* and *CXCR12* was observed in human AAA compared with non-aneurysmal aortic tissues³⁰. In particular, *CXCR4* was mainly co-localized in the inflammatory cells of the AAA, suggesting a potential role as a biomarker of inflammation. *CXCR4/CXCR12* signaling also mediated neuroinflammation. Thus, further studies are needed to explain the neuronal response under hemodynamic stress in IA development or rupture.

The G allele of the rs6841581 located near the 5'-untranslated region of *EDNRA* (4q31.22) gene was associated with IA (10). However, detailed mechanisms of IA mediated by *EDNRA* have yet to be reported to determine the direct effect or an indirect effect. Rats with pulmonary hypertension showed higher expression of *EDNRA* genes and endothelin-1 (ET-1)³¹. Endothelial injury, followed by disruption of collagen and elastin synthesis contribute to IA³². Chronic hypertension may induce the structural changes. Thus, the inflammatory response to increased hemodynamic stress following the disruption of cerebral arteries mediated by *EDNRA* may result in IA. In our study, among the *EDNRA*-interacting genes, the *MPDZ*-centered interaction between *EIF4H*, *SAP18*, and *UNC13C* was observed. *MPDZ* is a tight junction protein, which modulates Notch signaling during angiogenesis by controlling ligand recruitment to adherent junctions³³. Feldner et al. reported that loss of *MPDZ* decreased ependymal cell integrity and caused hydrocephalus. Ependymal cells are mainly responsible for electrolyte transport between brain parenchyma and the CSF³⁴. Adult ependymal cells are highly differentiated. Ependymal cells lining the lateral ventricles are quiescent under normal physiological conditions³⁵. However, after stroke, adult ependymal cells are transformed into radial glial cells in the subventricular zone³⁶. IA formation and growth occur within the CSF space surrounding the cerebral arteries. Accordingly, a further study is required to investigate the relationship between IA and CSF mediated by ependymal cells.

The study has some limitations. First, although we identified novel loci for IA, their functional role was not investigated. The role of most of the 34 novel genes interacting with *BOLL* and *EDNRA* in our study has yet to be analyzed in IA or other CVDs. Second, the study had a potential sample size limitation of multiple interaction terms with either *BOLL* or *EDNRA* loci (Supplementary Table S3), which have been replicated in IA susceptibility involving Korean adults including 250 patients with IA and 296 controls³⁷. In addition, there is a possibility that the results of this IA genetic study are limited to the Korean population. Therefore, we may warrant these findings in the second stage GWA meta-analysis and interaction test. Nevertheless, this study evaluates the first multiple GWA interaction study by IA-targeting genes such as *BOLL* and *EDNRA*, which have been validated in previous studies^{5,9,10}. Disease is caused by various biological pathways and it is difficult to explain it based on a GWAS, which focused on differences between individual paired loci³⁸. Accordingly, a study of genetic interactions and additional insights into various compensatory functional modules is needed to elucidate complex diseases such as IA. Our genome-wide interaction study involving *BOLL* and

EDNRA loci revealed several novel genes contributing to IA pathogenesis. However, few studies investigated the role of these genes in IA compared with AAA and thoracic aortic aneurysm (TAA). Thus, the possible role of the newly discovered genes was inferred by comparing the results of this study with those of AAA or TAA assuming that all aneurysms share a common pathogenic mechanism. However, no definite evidence was available to support a polygenic overlap between these IA, AAA, and TAA³⁹. Therefore, our discussion revolves around whether the newly identified genes are really related to the IA pathogenesis. To address these limitations, it is necessary to develop a general framework for mapping complex genetic networks of IA using GWAS data combined with clinical risk factors as well as to investigate the role of the newly identified genes in actual IA pathogenesis.

In conclusion, genome-wide interaction between IA and *BOLL* or *EDNRA* revealed 34 novel loci, which were robustly associated with IA. Common susceptibility variants and their interacting factors can be used to determine the inter-individual status of IA formation. The novel gene-gene interactions reported in this study need to be corroborated via larger prospective cohort studies. Elucidation of the functional interaction between genetic variants may provide insight into the functional roles in either IA development or pathogenesis. Further studies investigating the multifactorial interactions between genetic, epigenetic, and environmental factors and the effects on IA are needed to provide personalized interventions to prevent IA and concurrently regulate the associated metabolic abnormalities.

Declarations

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

J.P.J. conceptualized, reviewed, and approved the manuscript. E.P.H. and J.P.J. drafted the manuscript, revised the article critically, and performed all of statistical analyses. D.H.K. and B.J.K. conducted DNA works, contributed toward artwork, and reviewed the manuscript. E.P.H., J.J.L., S.N., H.Y., J.J.P. Y.J.C., and J.P.J. provided critical review and revised intellectual contents. All authors contributed to the article and approved the submitted version.

COMPETING INTERESTS

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

ACKNOWLEDGMENTS

This research was supported by a 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.

References

1. Zhou, S., Dion, P.A. & Rouleau, G.A. Genetics of intracranial aneurysms. *Stroke* **49**, 780-787; doi: 10.1161/STROKEAHA.117.018152 (2018).
2. 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; doi: 10.1016/S1474-4422(11)70109-0 (2011).
3. Boehme, A.K., Esenwa, C. & Elkind, M.S. Stroke risk factors, genetics, and prevention. *Circ. Res.* **120**, 472-495; doi: 10.1161/CIRCRESAHA.116.308398 (2017).
4. Yasuno, K. *et al.* Genome-wide association study of intracranial aneurysm identifies three new risk loci. *Nat. Genet.* **42**, 420-425; doi: 10.1038/ng.563 (2010).
5. Hong, E.P. *et al.* Genomic variations in susceptibility to intracranial aneurysm in the Korean population. *J. Clin. Med.* **8**, 275; doi: 10.3390/jcm8020275 (2019).
6. Bilguvar, K. *et al.* Susceptibility loci for intracranial aneurysm in European and Japanese populations. *Nat. Genet.* **40**, 1472-1477; doi: 10.1038/ng.240 (2020).
7. 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; doi: 10.1038/s41588-020-00725-7 (2020).
8. Genin, E. Missing heritability of complex diseases: Case solved? *Hum. Genet.* **139**, 103-113; doi: 10.1007/s00439-019-02034-4 (2020).
9. Hong, E.P., Kim, B.J. & Jeon, J.P. Genome-wide association between the 2q33.1 locus and intracranial aneurysm susceptibility: An updated meta-analysis including 18,019 individuals. *J. Clin. Med.* **8**, 692; doi: 10.3390/jcm8050692 (2019).

10. Hong, E.P. *et al.* Association of endothelin receptor type a with intracranial aneurysm in 20,609 East Asians: An updated meta-analysis. *World Neurosurg.* **130**, e804-e814; doi: 10.1016/j.wneu.2019.06.228 (2019).
11. Kim, B.J. *et al.* Genome-wide blood DNA methylation analysis in patients with delayed cerebral ischemia after subarachnoid hemorrhage. *Sci. Rep.* **10**, 11419; doi: 10.1038/s41598-020-68325-3 (2020).
12. Wu, X. *et al.* A novel statistic for genome-wide interaction analysis. *PLoS Genet.* **6**, e1001131; doi: 10.1371/journal.pgen.1001131 (2010).
13. Pruim, R.J. *et al.* Locuszoom: Regional visualization of genome-wide association scan results. *Bioinformatics.* **26**, 2336-2337; doi: 10.1093/bioinformatics/btq419 (2010).
14. GTEx Consortium. Genetic effects on gene expression across human tissues. *Nature* **550**, 204-213; doi: 10.1038/nature24277 (2017).
15. Ponten, F., Schwenk, J.M., Asplund, A. & Edqvist, P.H. The human protein atlas as a proteomic resource for biomarker discovery. *J. Intern. Med.* **270**, 428-446; doi: 10.1111/j.1365-2796.2011.02427.x (2011).
16. Huang, D.W., *et al.* David bioinformatics resources: Expanded annotation database and novel algorithms to better extract biology from large gene lists. *Nucleic. Acids. Res.* **35**, W169-W175; doi: 10.1093/nar/gkm415 (2007).
17. Szklarczyk, D. *et al.* String v11: Protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic. Acids. Res.* **47**, D607-D613; doi: 10.1093/nar/gky1131 (2019).
18. Schmiedel, B.J. *et al.* Impact of genetic polymorphisms on human immune cell gene expression. *Cell* **175**, 1701-1715.e16; doi: 10.1016/j.cell.2018.10.022 (2018).
19. Luetjens, C.M. *et al.* Association of meiotic arrest with lack of boule protein expression in infertile men. *J. Clin. Endocrinol. Metab.* **89**, 1926-1933; doi: 10.1210/jc.2003-031178 (2004).
20. Kang, K.J. *et al.* Oncogenic role of boll in colorectal cancer. *Dig. Dis. Sci.* **60**, 1663-1673; doi: 10.1007/s10620-015-3533-z (2015).
21. Westerveld, G.H., Repping, S., Leschot, N.J., van der Veen, F. & Lombardi, M.P. Mutations in the human boule gene are not a major cause of impaired spermatogenesis. *Fertil. Steril.* **83**, 513-515; doi: 10.1016/j.fertnstert.2004.10.013 (2005).
22. Xu, E.Y., Moore, F.L. & Pera, R.A. A gene family required for human germ cell development evolved from an ancient meiotic gene conserved in metazoans. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 7414-7419; doi: 10.1073/pnas.131090498 (2001).
23. Harrod, C.G., Batjer, H.H. & Bendok, B.R. Deficiencies in estrogen-mediated regulation of cerebrovascular homeostasis may contribute to an increased risk of cerebral aneurysm pathogenesis and rupture in menopausal and postmenopausal women. *Med. Hypotheses.* **66**, 736-756; doi: 10.1016/j.mehy.2005.09.051 (2006).
24. Ding, C., Toll, V., Ouyang, B. & Chen, M. Younger age of menopause in women with cerebral aneurysms. *J. Neurointerv. Surg.* **5**, 327-331; doi: 10.1136/neurintsurg-2012-010364 (2013).
25. Brady, L.K. *et al.* Transcriptome analysis of hypoxic cancer cells uncovers intron retention in eif2b5 as a mechanism to inhibit translation. *PLoS Biol.* **15**, e2002623; doi: 10.1371/journal.pbio.2002623 (2017).
26. 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; doi: 10.1161/CIRCRESAHA.116.308765 (2017).
27. Lipfert, J., Odemis, V., Wagner, D.C., Boltze, J. & Engele, J. Cxcr4 and cxcr7 form a functional receptor unit for sdf-1/cxcl12 in primary rodent microglia. *Neuropathol. Appl. Neurobiol.* **39**, 667-680; doi: 10.1111/nan.12015 (2013).
28. Wu, K.J. *et al.* A novel cxcr4 antagonist cx549 induces neuroprotection in stroke brain. *Cell. Transplant.* **26**, 571-583; doi: 10.3727/096368916X693563 (2017).
29. Cheng, X. *et al.* The role of sdf-1/cxcr4/cxcr7 in neuronal regeneration after cerebral ischemia. *Front. Neurosci.* **11**, 590; doi: 10.3389/fnins.2017.00590 (2017).
30. Tanios, F. *et al.* Cxcr4: A potential marker for inflammatory activity in abdominal aortic aneurysm wall. *Eur. J. Vasc. Endovasc. Surg.* **50**, 745-753; doi: 10.1016/j.ejvs.2015.07.040 (2015).
31. Lim, K.A. *et al.* Gene expression of endothelin-1 and endothelin receptor a on monocrotaline-induced pulmonary hypertension in rats after bosentan treatment. *Korean Circ. J.* **40**, 459-464; doi: 10.4070/kcj.2010.40.9.459 (2010).
32. Inci, S. & Spetzler, R.F. Intracranial aneurysms and arterial hypertension: A review and hypothesis. *Surg. Neurol.* **53**, 530-540; discussion 540-542; doi: 10.1016/s0090-3019(00)00244-5m (2000).
33. Tetzlaff, F. *et al.* Mpdz promotes dll4-induced notch signaling during angiogenesis. *Elife* **7**, e32860; doi: 10.7554/eLife.32860 (2018).
34. Feldner, A. *et al.* Loss of mpdz impairs ependymal cell integrity leading to perinatal-onset hydrocephalus in mice. *EMBO Mol. Med.* **9**, 890-905; doi: 10.15252/emmm.201606430 (2017).
35. Carlen, M. *et al.* Forebrain ependymal cells are notch-dependent and generate neuroblasts and astrocytes after stroke. *Nat. Neurosci.* **12**, 259-267; doi: 10.1038/nn.2268 (2009).
36. Zhang, R.L. *et al.* Stroke induces ependymal cell transformation into radial glia in the subventricular zone of the adult rodent brain. *J. Cereb. Blood Flow Metab.* **27**, 1201-1212; doi: 10.1038/sj.jcbfm.9600430 (2007).

37. Hong, E.P. & Park, J.W. Sample size and statistical power calculation in genetic association studies. *Genomics Inform.* **10**, 117-122; doi: 10.5808/GI.2012.10.2.117 (2012).
38. Fang, G. *et al.* Discovering genetic interactions bridging pathways in genome-wide association studies. *Nat. Commun.* **10**, 4274; doi: 10.1038/s41467-019-12131-7 (2019).
39. van 't Hof, F.N. *et al.* Shared Genetic Risk Factors of Intracranial, Abdominal, and Thoracic Aneurysms. *J. Am. Heart Assoc.* **5**, e002603; doi: 10.1161/JAHA.115.002603 (2016).

Figures

Figure 1

Manhattan plots of genome-wide interactions with (A) rs700651 (BOLL, 2q33.1) and (B) rs6841581 (EDNRA, 4q31.22) and their effect on intracranial aneurysm based on the additive effect model. Red line indicates genome-wide significance threshold (interaction $p = 1.25 \times 10^{-8}$).

Figure 2

Regional association plots of (A) rs1105980 (PTCH1, 9q22.32) interacting with rs700651 (BOLL, 2q33.1) and (B) rs74585958 (RYK, 3q22.2) interacting with rs6841581 (EDNRA, 4q31.22) at position ± 400 kb and the effect on intracranial aneurysm (IA). X-axis indicates the chromosomal position (mega base, Mb) and Y-axis did $-\log_{10}$ transformed p-value and recombination rate, respectively. Purple triangles of rs1105980 ($p = 6.41 \times 10^{-11}$) and rs74585958 ($p = 6.41 \times 10^{-9}$) represent the most significant interactions with BOLL and EDNRA, respectively. Other up or down triangles denote other variants within the target variant ± 400 kb regions. Up and down triangles indicate positive and negative effect sizes on IA formation, respectively. Each color shows pair-wise linkage disequilibrium with either rs1105980 or rs74585958.

Figure 3

Heatmap shows multiple gene expressions involving (A) BOLL- and (B) EDNRA-interacting genes in the artery, brain, and whole blood. BOLL and EDNRA genes are included in panel A and B, respectively. Expression levels were calculated using transcripts per million (TPM). Genes and tissues are ordered by agglomerative hierarchical clustering.

Figure 4

Network of multiple protein-coding interactions using (A) 18 out of 22 BOLL-interacting genes and (B) 11 EDNRA-interacting genes associated with the risk of intracranial aneurysm (*Homo sapiens*). Individual lines indicate associations between genes sharing similar function (black line, evidence of mRNA co-expression of the associated genes; light green line, co-publication information from text mining; purple line, experimental evidence supporting interaction; light purple line, evidence suggesting homologous proteins).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [OnlineSupplementaryData.docx](#)
- [SupplementaryTableS1.xlsx](#)
- [SupplementaryTableS2.xlsx](#)
- [SupplementaryTableS3.xlsx](#)
- [SupplementaryTableS4.xlsx](#)
- [SupplementaryTableS5.xlsx](#)