screening of ≥ 40 years old, (c) no medical history of any stroke including IA or SAH, and (d) no family history of IA or SAH in first-degree relatives.[10]

Individual and family history and life style was obtained by interviews. Past history and co-morbidity were also examined by clinical charts at the hospitals or interview charts at the brain check-ups.

The study was approved by the Ethics Committee of Kyoto University Institutional Review Board and appropriate informed consent was obtained from all subjects.

Direct sequencing and prediction of functional analysis for detected variants

All exons, intron-exon boundaries, putative promoter sequences and the 3'UTR were analyzed by direct sequencing of 9 genes for 58 cases (the first cohort). For sequencing, we referred to *TNFRSF13B*, *M-RIP*, *COPS3*, *RAI1*, *SREBF1*, *GRAP*, *MAPK7*, *MFAP4* and *AKAP10*, NCBI Map Viewer, http://www.ncbi.nih.gov/mapview/maps.cgi?. Primers for coding exons were designed from an intronic sequence >50 base pairs (bp) away from the intron-exon boundary and commercially synthesized by PROLIGO (PROLIGO Primers & Probes, Kyoto, Japan; http://www.proligo.com). For regulatory regions, about 500 bp upstream to the first exon was sequenced. However, if the database suggested the existence of a regulatory region further upstream, sequencing was done to cover the entire regulatory region. After PCR amplification and purification, sequencing was done on an ABI Prism 3100 Avant DNA sequencer (Applied

Biosystems, USA). We checked the SNP database (dbSNP) as a reference (http://www.ncbi.nlm.nih.gov/SNP/index.html). Primers and PCR conditions for each gene are available from the Supplement Table 2.

Among all the sequence changes identified by direct sequencing, we selected nonsense mutations and non-synonymous variants as primary candidate variants. Then, we conducted functional analysis for each non-synonymous variant by PolyPhen.

The number of subjects who had rare non-synonymous or deleterious changes was compared between the second cohort and the third cohort by Fisher's exact test using SAS software (Version 8.2. SAS Institute Inc).

Testing segregation in pedigrees

Three variants (K154X, c.585-586insA, G76C) of *TNFRSF13B* found in three probands were investigated for concordance of segregation in these families (Pedigree10, Pedigree26 and Pedigree15).[9]

Association study

SNPs of *TNFRSF13B* with allele frequency ≥1% in 58 cases (the first cohort: 29 probands of the pedigrees and 29 unrelated cases) were all genotyped by direct sequencing (P251L and S277S) or by polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP) using AlwI for c-247G>T and BfuAI for IVS3+25C>A in 304 unrelated cases (the second

cohort) and 332 controls (the third cohort).

Haplotypes were constructed using sequence variants with allele frequency ≥1% in the third cohort by THESIAS[14] (http://genecanvas.ecgene.net/). We used the following criteria to choose a set of haplotypes for association study: a set of the minimum number of haplotypes of which cumulative haplotype frequency was ≥80%[15] or a set of all halotypes of which frequencies were ≥5%.[16] Associations were analyzed by adjusting for covariates including sex, hypertension, smoking, and drinking habit. Bonferroni correction was done for comparison of multiple haplotypes, not for experiment-wide multiple testing.

Linkage Disequilibrium (LD) was analyzed and visualized with the Genotype2LDBlock (http://cgi.uc.edu/cgi-bin/kzhang/genotype2LDBlock.cgi).

Population attributable risk

The population attributable risk for a given haplotype was calculated as follows:

Population attributable risk= (Odds ratio-1)*IE*P/IT

Where IE is incidence of IA in the control cohort, IT is the incidence of IA in the general population and P is the reference haplotype frequency in the general population. We assumed that IE was equal to IT, and P in the control cohort was equal to that in the general population. Thus the population attributable risk will be obtained as follows:

Population attributable risk= (Odds ratio-1)* P

The authors had full access to the data and take full responsibility for its integrity. All authors have read and agree to the manuscript as written.

Results

Demographic features of the three cohorts

As shown in Table 2, among unrelated subjects the proportion of females or that of hypertension was higher in the second cohort than in the third cohort. Age at diagnosis was lower in the second cohort. No significant difference was found for either smoking or drinking habits.

Candidate genes

The primary candidate gene set after excluding genes based on a defined criteria was found to be related with immunity (*TNFRSF13B)*, regulatory component (*M-RIP*), protein kinase (*COPS3*, *MAPK7*), transcriptional factor (*RAI1*, *SREBF1*), signaling protein (*GRAP*), cell adhesion (*MFAP4*) and signal transduction (*AKAP10*).

Detected sequence changes in 58 cases (the first cohort) and their predicted effects on function are shown in Supplement Table 3. We identified seven sequence changes in *TNFRSF13B*, 20 sequence changes in *M-RIP*, six sequence changes in *COPS3*, 23 sequence changes in *RAI1*, nine sequence changes in *SREBF1*, nine sequence changes in *GRAP*, five sequence changes in *MAPK7*, two sequence changes in *MFAP4* and ten sequence changes in

AKAP10.

TNFRSF13B had two nonsense mutations and two non-synonymous variants predicted as "probably damaging" by PolyPhen: K154X and frame shift (c.585-586insA) in exon4, G76C in exon3 and P251L in exon5. Apparent deleterious variants, including nonsense mutations or non-synonymous variants, which were predicted to be "probably damaging", were identified only in *TNFRSF13B* (Table 3).

Segregation of the *TNFRSF13B* variants with the IA phenotype in pedigrees

Two nonsense mutations and one non-synonymous variant ("probably damaging" by PolyPhen) were found in probands in three pedigrees (Figure 1) out of 29 families. In one family (Pedigree10);[9] K154X was found in two affected siblings and one daughter while it was not detected in an unaffected younger brother. Insertion A (c.585-586insA) was found in one family (Pedigree26)[9]; an affected mother and her son had this mutation. This mutation was also found in an unaffected sibling of the mother, that later developed stroke but was not investigated for pathogenesis. G76C was found in another family (Pedigree15)[9]; two affected sisters had this variant but an unaffected sister did not.

Direct sequencing exons 3 to 5 in 304 unrelated cases (the second cohort) and 332 controls (the third cohort) in *TNFRSF13B*

An extensive search was done in exons 3 to 5 because there are two deleterious variants and

two probably damaging variants in these regions in *TNFRSF13B*. We further found additional sequence variants in the second and third cohort (Table 3). The number of subjects having rare non-synonymous changes, a splicing acceptor site change and a frame shift in *TNFRSF13B* were significantly larger in the 304 unrelated cases than in the 332 controls (Fisher's exact test, p=0.035, Table 4).

Locations of these variants were summarized in Figure 2. These non-synonymous variants were located on the region critical for function.[17-19] The 70th peptide S was conserved in *Xenopus laevis* but not in mice, dogs or rats. On the other hand, 74th E, 76th G, and 177th C were conserved in mice, dogs and rats. It should also be pointed out that G76S and C177R were found in more than one unrelated subjects. But, no one had more than one variant.

Association study

Allele frequencies of four SNPs were found to be ≥1% in the third cohort (Table 3). We thus used four SNPs (c-247G>C, IVS3+25C>A, P251L, S277S) of *TNFRSF13B* to construct haplotypes. LD structure was shown in Figure 3. Application of the selection criteria chose four hypotypes, covering 87% of all haplotypes (Table 5). Haplotype H1 (GACC) was found to be protective (odds=0.69, 95%CI: 0.52-0.92, p=0.012) as compared with the major haplotype H4 (TCTC). After Bonferroni correction for multiple comparison, the p value of H1 was still statistically significant (p_{corr}=0.048).

Population attributable risks

The population attributable risk was calculated to be about –8% for H1 haplotype against H4 compared with 24% for smoking against nonsmoking. Therefore the attributable risk for the *TNFRSF13B* variants was approximately one-third that of smoking.

Discussion

Extensive efforts have been made to search for susceptibility gene(s) for IA. So far, three genome-wide linkage analyses have been done for the general population. Except *ELN*,[6] *LOX*[20] and *COL1A2*,[21] no gene has been claimed as a candidate gene. There have been contradictions, however, in terms of involvement of *ELN* in IA.[22,23]

In the present study, we have conducted a systematic approach targeting a linked region on chromosome 17. We selected nine candidates from 108 genes and sequenced entire coding exons and regulatory regions in 58 cases (the first cohort). Since we found several variants including obvious deleterious mutations in *TNFRSF13B*, we searched variants in 304 unrelated cases (the second cohort) and 332 control subjects (the third cohort), although searches were limited to those in exons 3 to 5 which covered the critical areas CRD2 (cysteine-rich domain 2), TM (trans-membrane) and ICD (intracellular regions). The rare variants were significantly more frequent in IA unrelated cases than in control subjects. In addition, deleterious variants [K154X,

frame shift (c.585-586insA) and G76C] were clearly segregated in the families except in a family sibling who had c.585-586insA but did not have IA, implying that the penetrance of IA is not complete as was expected.[24,25] Finally, case-control studies using sequence variants revealed a protective haplotype (GACC) against the most common haplotype (TCTC). With these lines of evidence, *TNFRSF13B* emerges as a candidate for susceptibility for IA.

TACI (Transmembrane activator and calcium-modulator and cyclophilin ligand interactor) encoded by *TNFRSF13B* mediates isotype switching in B cells. The mutations in *TNFRSF13B* have recently been reported to be associated with common variable immunodeficiency (CVID) and IgA deficiency in humans.[26,27] In one of these studies, 11 mutations (4.1%) were found in 270 chromosomes from 135 sporadic CVID cases.[26] It is of particular interest that most sporadic cases with CVID had only one mutant allele, suggesting a mechanism of gain of function or haploinsufficiency.

Given that mutations of *TNFRSF13B* are associated with CVID or IgA deficiency, an unanswered question is why variants in *TNFRSF13B* are associated with IA. It is interesting that in our study 12 out of 17 rare variants in IA cases, and 3 out of 5 rare variants in control subjects were found in the CRD2 domain, while the majority of mutations in cases with CVID or IgA deficiency were found at the C terminal side to the CRD2 region, which transfers signals from cell surface to intracellular domains. We postulate that variants at the ligand binding site may

cause quantitative changes while mutations in signal transduction results in qualitative changes. Different modes of functional impairments might be associated with different phenotypes. Studies are needed to investigate this further.

In the present study, we found three nonsense mutations (one stop codon, one splicing acceptor site change, and one frame shift) and five rare non-synonymous changes in 17 cases. Each case had a single variant. It is interesting that these variants are novel and none was found in Caucasians.[26,27] The most common mutation among Japanese with IA is G76S (8/17) while that in Caucasians with CVID is A181E, suggesting founder mutations specific to ethnic groups. If so, genetic preposition to IA or CVID or IgA deficiency may be predicted by these founder mutations in the future.

The present study has several limitations. First, population attributable risks of IA are calculated to be 7% to 10% while that of smoking observed is about 24%, suggesting that the risk attributable to *TNFRSF13B* is about one-third of smoking in the present cohort. However, further studies are needed because only a small fraction of the risk is explained by *THNFRS13B*. Second, we have selected only nine genes as the primary gene set from 108 genes. We excluded genes of which functions are not well characterized or those of which well characterized functions are not considered to be involved in IA. Although this is primary screening, this study cannot be free from selection bias. In the next study, we are expanding the gene set so that it

includes some genes of which functions are unknown. Third, we tested with PolyPhen whether non-synonymous variants were functional or not. Bioinformatics approaches may sometime be misleading.[28] We should explore other genes which had "Possibly damaging" or "unknown" variants in future. Effects of variants on function of TACI should also be confirmed experimentally in future. Fourth, there may be an argument for the rare variants contribute to a common diseases hypothesis. It should be addressed, however, that the hypothesis can provide criteria for positive selection of a susceptibility gene, which would have been overlooked by a haplotype-based association study. Fifth, in the present study, we did not determine CVID related parameters such as B-cell expression of TACI and serum levels of immunoglobulin. Finally, we did not explore genes in LD with *TNFRSF13B*. The International HapMap Project (http://www.hapmap.org) suggests that there is LD between *LOC96597* and *TNFRSF13B*. Further exploration will be needed in future.

With positive findings, above rationale and reasonable background, we proposed that *TNFRSF13B* is one of the candidate genes for susceptibility for IA notwithstanding several limitations. This in turn proposes that immunological mechanisms may play a role in IA development to a discernible extent. Our hypothesis is in accordance with clinical experiences where IA is often found in subjects with autoimmune diseases.[29,30] Further studies are needed to strengthen our hypothesis. In addition, the present results might pave the way for investigating a

link between immunological events and IA development.

Acknowledgments

This work was supported by a grant from the Ministry of Education, Science, Sports, and Culture of Japan to AK (15012231, 16012232 and 17019034) and a grant from the Japan Society for the Promotion of Science to AK (Kiban Kenkyu A: 14207016 and S: 17109007). We thank Norio Matsuura for technical assistance, and the following doctors for patient recruitment and help in ascertaining MRA examinations: Shinsuke Tominaga, Hiroshi Hasegawa, and Toshihiko Inui (Tominaga Hospital), Shyunichi Yoneda and Yoshito Naruo (Nihonbashi Hospital), Yoo Kang and Shoichi Tani (Osaka Saiseikai Izuo Hospital), Kouji Murakawa (Takayama Red Cross Hospital), Hiroyasu Yamakawa (Gero-spring Hospital), Atsushi Kawarazaki (Kawarazaki Hospital), Masayuki Matsuda (Shiga University of Medical Science), Michiyasu Suzuki and Sadahiro Nomura (Yamaguchi University School of Medicine), Takaaki Kaneko and Nozomu Murai (Hikone Municipal Hospital), Tatsuhito Yamagami (Kyoto Kizugawa Hospital), Hikaru Ohishi, Kiminari Ohtaka, Junko Sasaki, Koki Iwaya and Masaya Iwakawa (Senboku Kumiai Sougou Hospital), Kenji Kikuchi and Yutaka Yamazaki (Yuri Kumiai Sougou Hospital), Shiro Nagasawa and Nobuhisa Mabuchi (Soseikai General Hospital), Yasuhiko Tokuriki (Fukui Red Cross Hospital), Tomoo Tokime (Tenri Hospital), Sen Yamagata (Kurashiki Central Hospital), Kenji Hashimoto (Hyogo Prefectural Tsukaguchi Hospital), Atsushi Okumura (Takeda Hospital), Yoshihiko Uemura (Kyoto City Hospital), Tomohiko Iwai

(Gifu Municipal Hospital), Kiyohiro Houkin and Osamu Honmou (Sapporo Medical University School of Medicine), Izumi Nagata (Nagasaki University School of Medicine), Ichiro Nakahara and Toshio Higashi (Kokura Memorial Hospital), and Takashi Yoshizawa and Kenjiro Ito (Yokohama Shintoshi Neurosurgical Hospital), Jun Takahashi, Nobuhiro Mikuni, Ken-ichiro Kikuta, and Yasushi Takagi (Kyoto University Graduate School of Medicine).

Conflict of Interest Disclosures

None

References

1. Yamada S, Koizumi A, Iso Y, Wada Y, Watanabe Y, Date C, Yamamoto A, Kikuchi S, Inaba Y, Toyoshima H, Kondo T, Tamakoshi A, the JACC Study Group. Risk factors for fatal subarachnoid hemorrhage: the Japan collaborative cohort study. *Stroke*. 2003;34:2781-2787.

2. Stehbens WE. Familial intracranial aneurysms: an autopsy study. *Neurosurgery*. 1998;43:1258-1259.

3. Ronkainen A, Niskanen M, Piironen R and Hernesniemi J. Familial subarachnoid hemorrhage: outcome study. *Stroke*. 1999;30:1099-1102.

4. Ruigrok YM, Rinkel GJ, Algra A, Raaymakers TW and Van Gijn J. Characteristics of intracranial aneurysms in patients with familial subarachnoid hemorrhage. *Neurology*. 2004;23:891-894.

5. Kissela BM, Sauerbeck L, Woo D, Khoury J, Carrozzella J, Pancioli A, Jauch E, Moomaw CJ, Shukla R, Gebel J, Fontaine R and Broderick J. Subarachnoid hemorrhage: a preventable disease with a heritable component. *Stroke*. 2002;33:1321-1326.

6. Onda H, Kasuya H, Yoneyama T, Takakura K, Hori T, Takeda J, Nakajima T and Inoue I. Genome-wide linkage and haplotype-association syudies map intracranial aneurysm to chromosome 7q11. *Am J Hum Genet.* 2001;69:804-819.

7. Olson JM, Vongpunsawad S, Kuivaniemi H, Ronkainen A, Hernesniemi J, Ryynanen M, Kim

LL and Tromp G.. Search for intracranial aneurysm susceptibility gene(s) using Finnish families. *BMC Med Genet.* 2002;3:7.

8. Nahed BV, Seker A, Guclu B, Ozturk AK, Finberg K, Hawkins AA., DiLuna ML, State M, Lifton RP and Gunel M. Mapping a Mendelian form of intracranial aneurysm to 1p34.3-p36.13. *Am J Hum Genet.* 2005;76:172-179.

9. Yamada S, Utsunomiya M, Inoue K, Nozaki K, Inoue S, Takenaka K, Hashimoto N and Koizumi A. Genome-wide scan for Japanese familial intracranial aneurysms Linkage to several chromosomal regions. *Circulation.* 2004;110:3727-3733.

10. Mineharu Y, Inoue K, Inoue S, Yamada S, Takenaka K, Nozaki K, Hashimoto N and Koizumi A. Association Analysis of Common Variants of *ELN*, *NOS2A*, *APOE* and *ACE2* to Intracranial Aneurysm. *Stroke.* 2005;In Press.

11. Schievink WI. Genetics of intracranial aneurysms. *Neurosurgery.* 1997;40:651-62.

12. Jonathan CC, Robert SK, AlexanderP, Yves LM, Ruth M and Helen HH. Multiple Rare Alleles Contribute to Low Plasma Levels of HDL Cholesterol. *Science.* 2004;305:869-872.

13. Sunyaev S, Ramensky V, Koch I, Lathe W III, Kondrashov AS and Bork P. Prediction of deleterious human alleles. *Hum Mol Genet.* 2001;10:591-597.

14. Tregouet DA, Barbaux S, Escolano S, Tahri N, Golmard JL, Tiret L and Cambien F. Specific haplotypes of the P-selectin gene are associated with myocardial infarction.

Hum Mol Genet. 2002;11:2015–2023.

15. Carlson CS, Eberle MA, Rieder MJ, Yi O, Kruglyak L and Nickerson DA. Selecting a maximally informative set of single-nucleotide polymorphisms for association analyses using linkage disequilibrium. *Am J Hum Genet*. 2004;74:106-120.

16. Johnson GCL, Esposito L, Barratt BJ, Smith AN, Heward J, GenovaGD, Ueda H, Cordell HJ, Eaves IA, Dudbridge F, Twells RCB, Payne F, Hughes W, Nutland S, Stevens H, Carr P, Tuomilehto-Wolf E, Tuomilehto J, Gough SCL, Clayton1 DG and ToddJA. Haplotype tagging for the identification of common disease genes. *Nature Genetics*. 2001;29:233-237.

17. Hymowitz SG, Patel DR, Wallweber HJ, Runyon S, Yan M, Yin J, Shriver SK, Gordon NC, Pan B, Skelton NJ, Kelley RF and Starovasnik MA. Structures of APRIL-receptor complexes: like BCMA, TACI employs only a single cysteine-rich domain for high affinity ligand binding. *J Biol Chem*. 2005;280:7218-7227.

18. Locksley RM, Killeen N and Lenardo MJ. The TNF and TNF Receptor Review Superfamilies: Integrating Mammalian Biology. *Cell*. 2001;104:487–501.

19. Liu Y, Hong X, Kappler J, Jiang L, Zhang R, Xu L, Pan C, Martin WE, Murphy RC, Shu H, Dai S and Zhang G. Ligand–receptor binding revealed by the TNF family member TALL-1. *Nature*. 2003;423:49-56.

20. Yoneyama T, Kasuya H, Onda H, Akagawa H, Jinnai N, Nakajima T, Hori T and Inoue I.

Association of positional and functional candidate genes *FGF1, FBN2*, and *LOX* on 5q31 with intracranial aneurysm. *J Hum Genet*. 2003;48:309-314.

21. Yoneyama T, Kasuya H, Onda H, Akagawa H, Hashiguchi K, Nakajima T, Hori T and Inoue I. Collagen type I alpha2 (*COL1A2*) is the susceptible gene for intracranial aneurysms. *Stroke*. 2004;35:443-448.

22. Berthelemy-Okazaki N, Zhao Y, Yang Z, Camp NJ, Farnham J, Parker D, Tsuruda J, Macdonald J, Zhang K and Cannon-Albright LA. Examination of *ELN* as a candidate gene in the utah intracranial aneurysm pedigrees. *Stroke*. 2005;36:1283-1284.

23. Farnham JM, Camp NJ, Neuhausen SL, Tsuruda J, Parker D, MacDonald J and Cannon-Albright LA. Confirmation of chromosome 7q11 locus for predisposition to intracranial aneurysm. *Hum Genet*. 2004;114:250-255.

24. Wacholder S, Hartge P, Struewing JP, Pee D, McAdams M, Brody L and Tucker M . The kin-cohort study for estimating penetrance. *Am J Epidemiol*. 1998;148:623-630.

25. Gail MH, Pee D, Benichou J and Carroll R. Designing studies to estimate the penetrance of an identified autosomal dominant mutation: cohort, case-control, and genotyped-proband designs. *Genet Epidemiol*. 1999;16:15–39.

26. Salzer U, Chapel HM, Webster ADB, Pan-Hammarström Q, Schmitt-Graeff A, Schlesier M, Peter HH, Rockstroh JK, Schneider P, Schäffer AA, Hammarström L and Grimbacher B.

Mutations in *TNFRSF13B* encoding TACI are associated with common variable immunodeficiency in humans. *Nat Genet.* 2005;37:820-828.

27. Castigli E, Wilson SA, Garibyan L, Rachid R, Bonilla F, Schneider L and Geha RS. TACI is mutant in common variable immunodeficiency and IgA deficiency. *Nat Genet.* 2005;37:829-834.

28. Tchernitchko D, Goossens M and Wajcman H. In silico prediction of the deleterious effect of a mutation: proceed with caution in clinical genetics. *Clin Chem.* 2004;50:1974-8.

29. Mimori A, Suzuki T, Hashimoto M, Nara H, Yoshio T, Masuyama J-I, Okazaki H, Hirata D, Kano S and Minota S. Subarachnoid hermorrhage and systemic lupus erythematosus. *Lupus.* 2000;9:521-526.

30. Sanchez-Ojanguren J, Matias J, Misis M and Olive A. Systemic lupus erythematosus, berry aneurysm and subarachnoid haemorrhage. *Clin Rheumatol.* 1999;18:165-166.

Figure legend

Figure 1: Segregation of the *TNFRSF13B* deleterious change with the IA phenotype in pedigrees

Figure 2: Schema of the domain structure of full-length *TNFRSF13B*

Figure 3: LD (Linkage Disequilibrium) structure of *TNFRSF13B*

Table 1. Nine Genes First Sequenced in Chromosome 17 Centromere in 58 cases (the first cohort)

Gene Symbol	Gene name	MIM Number	Position	GenBank Accession Number (2006/02/23)	Genomic Region (kb)	mRNA Length (bp)	Number of Exons
TNFRSF13B	tumor necrosis factor receptor superfamily, member 13B	604907	16473152-16439349	NT_010718.15	33.804	879	4
M-RIP	myosin phosphatase-Rho interacting protein	-	16543056-16686620	NT_010718.15	143.565	3,114	29
COPS3	COP9 constitutive photomorphogenic homolog	604665	16782340-16747090	NT_010718.15	35.251	1,269	12
RAI1	retinoic acid induced 1	607642	17181736-17312516	NT_010718.15	130.781	5,718	8
SREBF1	sterol regulatory element binding transcription factor 1	184756	17338043-17312341	NT_010718.15	25.703	3,441	21
GRAP	GRB2-related adaptor protein	604330	18548021-18522034	NT_010718.15	25.988	651	6
MAPK7	mitogen-activated protein kinase 7	602521	18877883-18884469	NT_010718.15	6.587	2,448	7
MFAP4	microfibrillar-associated protein 4	600596	18888110-18883573	NT_010718.15	4.538	765	6
AKAP10	A kinase (PRKA) anchor protein 10	604694	19478745-19405569	NT_010718.15	73.177	1,986	15