Two Cases of Epstein-Barr Virus Infection Associated with Kawasaki Disease

Takaomi Minami[1], Tetsu Yoshimasu[1], Taketsugu Hama[1], Genkichi Izumi[1], Mina Booka[1], Nahoko Yata[1], Masakazu Miyawaki[1], Tomohiro Suenaga[2], Shoichi Shibuta[2], Takashi Takeuchi[2], Hiroyuki Suzuki[2] and Norishige Yoshikawa[2]

1) Department of Pediatrics, Social Insurance Kinan Hospital, Wakayama, Japan
2) Department of Pediatrics, Wakayama Medical University, Wakayama, Japan

We report two cases of Epstein-Barr virus (EBV) infection associated with Kawasaki disease (KD). Case 1 : 4-year-old boy. The IgM and IgG antibodies to viral capsid antigen (VCA) were detected, and anti-EBNA antibody was absent. These findings could confirm the diagnosis of acute EBV infection. Then he fulfilled the clinical criteria for typical KD, and was treated with single infusion of intravenous immunoglobulin (IVIG) (2 g/kg) on day 5. Though his fever fell on day 7, his left main coronary artery was dilated transiently. Case 2 : 3-year-old boy. The IgM antibody to VCA was also detected from his serum, and he fulfilled the KD criteria. He was treated with IVIG (2 g/kg) and aspirin (30 mg/kg/day) on day 5, and his fever fell on day 7 without coronary artery lesions. According to our cases, EBV possibly relates to the etiology of KD. Further studies are needed to elucidate the relationship between EBV infection and KD.

Positive association of genetic variants in the upstream region of *NKX2-3* with Crohn's disease in Japanese patients

K Yamazaki,[1] A Takahashi,[2] M Takazoe,[3] M Kubo,[4] Y Onouchi,[1] A Fujino,[5] N Kamatani,[3] Y Nakamura,[6,7] A Hata[1,8]

► Additional tables are published online only at http://gut.bmj.com/content/vol58/issue2

[1] Laboratory for Gastrointestinal Diseases, SNP Research Center, The Institute of Physical and Chemical Research (RIKEN), Kanagawa, Japan; [2] Laboratory for Statistical Analysis, SNP Research Center, RIKEN, Tokyo, Japan; [3] Department of Medicine, Division of Gastroenterology, Social Insurance Chuo General Hospital, Tokyo, Japan; [4] Laboratory for Genotyping, SNP Research Center, RIKEN, Kanagawa, Japan; [5] Department of Surgery, Keio University School of Medicine, Tokyo, Japan; [6] Laboratory of Molecular Medicine, Institute of Medical Science, The University of Tokyo, Japan; [7] SNP Research Center, RIKEN, Kanagawa, Japan; [8] Department of Public Health, Chiba University Graduate School of Medicine, Chiba, Japan

Correspondence to: Dr K Yamazaki, Laboratory for Genotyping Development, Center for Genomic Medicine, Institute of Physical and Chemical Research (RIKEN), 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan; kyama@src.riken.jp

Revised 4 September 2008
Accepted 16 September 2008
Accepted for Online First 17 October 2008

ABSTRACT

Background and aims: A number of genome-wide association studies have been performed as a robust means of identifying susceptibility loci for Crohn's disease (CD). The loci detected after the completion of the HapMap project are quite concordant among these studies, suggesting that the results are reliable. Recently, the Wellcome Trust Case Control Consortium (WTCCC) reported the primary scanning of 17 000 individuals for seven diseases, including CD, and a subsequent study has validated these susceptible genetic variants in independent UK sample sets. The purpose of this study was to study the possible association of the variants reported by the WTCCC with CD in a Japanese population.

Patients and methods: A total of 484 patients with CD and 470 healthy controls were examined. Seventeen genetic variants at eight newly identified loci, including *IRGM*, *NKX2-3* and *PTPN2*, were genotyped using the TaqMan assay or the invader assay.

Results: A positive association signal presumably common to different ethnic groups for rs10883365 was detected in the upstream region of *NKX2-3* ($p = 0.019$ under the genotypic model, $p = 0.0065$ under the allelic model, $p = 0.019$ under the recessive model, $p = 0.036$ under the dominant model). In addition to rs10883365, marginal associations for two single nucleotide polymorphisms (SNPs) were detected in the Japanese population; rs6887695 near *IL12B* and rs10761659 on 10q21. Further genotype–phenotype analysis found a significant association between rs6887695 and patients with pure ileal CD.

Conclusions: The results indicate that the three loci are possible candidates for conferring susceptibility to CD in people of different ethnicities.

Inflammatory bowel diseases (IBDs) such as Crohn's disease (CD; MIM 266600) and ulcerative colitis (MIM 191390) are chronic conditions characterised by remitting and relapsing inflammation of the small and/or large intestine. The onset of IBDs typically occurs in the second or third decade of life. In spite of the many studies performed in the past to identify the pathophysiological processes underlying these diseases, the aetiology of IBDs remains largely unknown. IBDs are thought to be multifactorial diseases and are assumed to be the result of an abnormal response of the mucosal immune system to unknown pathogen(s) and/or substance(s), and involving the aggregate effects of genetic, environmental and other processes.[1 2]

The involvement of genetic factors in the aetiology of IBD was suggested by epidemiological and linkage studies.[3] There are two main approaches to the identification of genes in complex genetic disorders: a hypothesis-free approach based on linkage analysis or an association study of candidate genes. The first susceptibility gene, *NOD2* (*CARD15*), was identified at the *IBD1* locus on chromosome 16 by a linkage and fine-mapping approach in 2001.[4 5] Recently, genome-wide association (GWA) studies have been used as a robust means of studying multigenic disorders. Now that the International HapMap Consortium has completed its comprehensive cataloguing of the human genome in various populations,[6] accurate information on genome-wide tagging single nucleotide polymorphisms (SNPs), which are representative of a unique haplotype, has become freely available on the internet. This makes GWA studies far more efficient and less costly than genotyping all SNPs in the genome.[7 8] The Wellcome Trust Case Control Consortium (WTCCC) recently reported on 1748 CD cases and 2938 controls from the UK using the Affymetrix GeneChip Human Mapping 500 k Array Set.[9] The results of Parkes *et al* supported these findings with an independent set of 1182 CD cases and 2024 controls of European descent.[10]

To evaluate whether the association identified in the UK population also holds in Japanese people, we selected 17 SNPs, including eight newly identified loci: two loci on chromosomes 1q24 and 1q31, and six nearby loci *BSN*, *IL12B*, *IRGM*, *NKX2-3*, *PTPN2* and *FLJ45139*. We have assessed the contribution of selected genetic variants in the Japanese population and have also investigated genotype–phenotype associations.

METHODS

Human subjects and phenotypic analysis

Blood samples were obtained with written informed consent from 484 Japanese patients with CD at the Social Insurance Chuo General Hospital and from 470 unaffected control individuals belonging to the Osaka-Midosuji Rotary Club or Keio University School of Medicine. The study protocol was approved by the local ethics board. All CD cases were diagnosed at the Inflammatory Bowel Unit of the Social Insurance Hospital by clinical, radiological, endoscopic and histological findings according to the Lennard–Jones criteria,[11] and patients with indeterminate colitis were excluded. Extensive clinical characterisation was available for 482 patients with CD. The clinical

characteristics of patients with CD were assessed at the time of diagnosis, and were categorised using the Vienna classification.[12] In addition, medical histories of surgery were obtained from the clinical records of 418 patients. The demographics of patients with CD and of controls are shown in table 1.

SNP analysis and genotyping

We selected a total of 17 SNPs for genotyping; nine of them were reported by the WTCCC[9] as showing the strongest association signals and 12 were confirmed in another sample set in the UK from subjects of European descent.[10] Seven SNPs were genotyped using the TaqMan assay (Applied Biosystems, Foster City, California, USA) and the remaining 10 SNPs by the invader assay (Third Wave Technologies, Madison, Wisconsin, USA).[13] The primers and typing methods used are described in Supplementary table 1. When the risk and minor allele frequencies of the SNPs were not given in the original reports,[10] the required information was obtained by calculating genotype data from information provided by the WTCCC (http://www.wtccc.org.uk/, ftp://ftp.sanger.ac.uk/pub/WTCCC/).

Statistical analysis

Departure from Hardy–Weinberg equilibrium was evaluated using χ^2 test. Allele frequencies were analysed with Fisher exact test and genotype distributions were analysed with the Cochran–Mantel–Haenszel χ^2 test; we used the test as implemented in R with the option that gives an exact p value (http://www.r-project.org/). Genotype–phenotype associations were analysed by the same method in CD susceptibility using R.

Meta-analyses were performed for the SNPs whose data were available for all three studies, an original WTCCC study,[9] a replication study[10] and the present study, by a DerSimonian–Laird random effects model.[14] We excluded a monomorphic SNP, rs17234567, in our Japanese sample from the analysis. p_H and p_E represented p values of heterogeneity and that of the combined odds ratio (OR), respectively (Supplementary table 2).

RESULTS

Association analysis with Japanese CD patients

The markers that showed significant deviation from Hardy–Weinberg equilibrium in controls were excluded before the analysis ($p>0.05$). The overall success rate of the genotyping assay was 99.1%. We evaluated the association of 17 SNPs in a Japanese population; these markers have been reported to be strongly associated with CD in the UK population. Three of the SNPs were found to be significantly associated: rs6887695 (5q33, $p = 0.035$; OR = 1.42 (1.03 to 1.95) under the recessive model), rs10761659 (10q21, $p = 0.040$; OR = 1.23 (1.01 to 1.50) under the allelic model, $p = 0.045$; OR = 1.30 (1.01 to 1.68) under the recessive model) and rs10883365 (10q24, $p = 0.019$ under the genotypic model, $p = 0.006$; OR = 1.29 (1.07 to 1.54) under the allelic model, $p = 0.036$; OR = 1.45 (1.04 to 2.02) under the dominant model, $p = 0.019$; OR = 1.45 (1.04 to 2.02) under the recessive model) (table 2). On the other hand, rs17234657 (5p13) was not polymorphic in our Japanese populations. Furthermore, rs11805303 in *IL23R* and rs10210302 in *ATG16L1*, which were recently identified by GWA studies, were not significant. We have already found other markers at these loci not significant in the previous report.[15]

Genotype association with CD phenotypes

To look for possible associations with CD phenotypes, we stratified samples based on the Vienna classification for location of disease, or disease behaviour. The subgroup of subjects with CD with the pure ileal disease (n = 70) appeared to be more strongly associated with rs6887695 (5q33) ($p<0.001$ under the genotypic model, $p<0.001$; OR = 2.21 (1.53 to 3.21) under the allelic model, $p = 0.0027$; OR = 2.76 (1.37 to 5.56) under the dominant model, $p<0.001$; OR = 2.90 (1.70 to 4.95) under the recessive model) (table 3). There was no significant association when compared with whole ileal disease regardless of colonic involvement (n = 265). The associations of other subgroups were not significant for the remaining 16 genetic variants (data not shown).

DISCUSSION

GWA studies, if available, have long been recognised as a powerful tool for identifying genetic variants related to common complex diseases. Recently developed resources from the International HapMap Project and the availability of low cost large-scale genotyping have finally made GWA studies a real robust tool.[6] A consortium of >50 British groups known as the WTCCC has performed GWA studies, the largest scale studies ever carried out, for seven common diseases; rheumatoid arthritis, hypertension, CD, coronary artery disease, bipolar disorder, and type 1 and type 2 diabetes.[9] Parkes *et al* replicated the result in CD with 37 SNPs from 31 loci in independent panels.[10]

Table 1 Demographic and clinical features of Crohn's disease patients and unaffected controls

		Crohn's disease (n = 482)	Controls (n = 470)
Sex (M/F/unknown)		351/129/2	236/233/1
Age at disease onset (years) (median (range))		22.4 (7–55)	38.7 (21–77)
Age at disease onset (n (%))*	<40 years	460 (95.8%)	
	⩾40 years	20 (4.2%)	
Disease location	Ileal disease (L1)	191 (39.6%)	
	Colonic disease (L2)	70 (14.5%)	
	Ileal and colonic disease (L3)	196 (40.7%)	
	Upper GI disease (L4)	22 (4.6%)	
	Uncertain	3 (0.6%)	
Disease behaviour	Inflammatory (B1)	126 (26.1%)	
	Stricturing (B2)	244 (50.6%)	
	Penetrating (B3)	107 (22.2%)	
	Uncertain	5 (1.0%)	
Need for surgery†		210 (69.3%)	

*Information on age at onset was available for 480 patients.
†Information on surgery was available for 418 patients.
F, female; GI, gastrointestinal; M, male.

Table 2 Genotype distribution of candidate single nucleotide polymorphisms (SNPs) positively associated with Crohn's disease (CD) in the Wellcome Trust Case Control Consortium

			Number of CD genotypes			Number of control genotypes			p value			
SNP marker	Chr*	Gene†	AA‡	Aa	aa	AA	Aa	aa	Genotype	Allelic	Dominant	Recessive
rs11805303	1p31	*IL23R*	151 (31.5%)	247 (51.5%)	82 (17.1%)	142 (30.4%)	226 (48.4%)	99 (21.2%)	0.27	0.27	0.12	0.78
rs12035082	1q24		445 (93.3%)	31 (6.5%)	1 (0.2%)	424 (90.8%)	43 (9.2%)	0 (0.0%)	0.13	0.24	1	0.19
rs10801047	1q31		8 (1.7%)	96 (20.2%)	371 (78.1%)	8 (1.7%)	87 (18.9%)	365 (79.3%)	0.89	0.72	0.69	1.00
rs10210302	2q37	*ATG16L1*	276 (57.1%)	184 (38.1%)	23 (4.8%)	270 (57.7%)	172 (36.8%)	26 (5.6%)	0.81	0.96	0.66	0.90
rs9858542	3p21	*BSN* etc.	447 (92.7%)	35 (7.3%)	0 (0.0%)	433 (92.5%)	33 (7.1%)	2 (0.4%)	0.54	0.72	0.24	0.90
rs17234657	5p13		0 (0.0%)	0 (0.0%)	484 (100.0%)	0 (0.0%)	0 (0.0%)	466 (100.0%)	1	1	1	1
rs9292777	5p13		16 (3.3%)	133 (27.6%)	333 (69.1%)	14 (3.0%)	105 (22.6%)	346 (74.4%)	0.18	0.10	0.07	0.85
rs10077785	5q23	*IBD5*	425 (89.7%)	47 (9.9%)	2 (0.4%)	428 (91.6%)	39 (8.4%)	0 (0.0%)	0.28	0.24	0.50	0.31
rs6887695	5q33	*IL12B*	110 (22.8%)	233 (48.2%)	140 (29.0%)	81 (17.2%)	245 (52.1%)	144 (30.6%)	0.10	0.12	0.62	0.035§
rs4958847	5q33	*IRGM*	116 (24.2%)	241 (50.2%)	123 (25.6%)	110 (23.5%)	244 (52.1%)	114 (24.4%)	0.83	0.93	0.71	0.82
rs13361189	5q33	*IRGM*	191 (40.9%)	220 (47.1%)	56 (12.0%)	199 (43.9%)	205 (45.3%)	49 (10.8%)	0.63	0.35	0.61	0.39
rs1000113	5q33		53 (11.0%)	222 (45.9%)	209 (43.2%)	43 (9.2%)	209 (44.7%)	216 (46.2%)	0.53	0.28	0.36	0.39
rs10761659	10q21		254 (52.7%)	188 (39.0%)	40 (8.3%)	217 (46.2%)	204 (43.4%)	49 (10.4%)	0.11	0.040	0.27	0.045§
rs10883365	10q24	*NKX2-3*	101 (21.0%)	254 (52.8%)	126 (26.2%)	73 (15.5%)	241 (51.3%)	156 (33.2%)	0.019§	0.0065§	0.019§	0.036§
rs17221417	16q12	*NOD2*	484 (100.0%)	0 (0.0%)	0 (0.0%)	466 (99.6%)	2 (0.4%)	0 (0.0%)	0.24	0.24	1	0.24
rs2542151	18p11	*PTPN2*	8 (1.7%)	94 (19.4%)	382 (78.9%)	4 (0.9%)	82 (17.6%)	381 (81.6%)	0.39	0.23	0.33	0.39
rs2836754	21q22	*FLJ45139*	24 (5.2%)	182 (39.3%)	257 (55.5%)	25 (5.4%)	170 (36.6%)	269 (58.0%)	0.69	0.59	0.47	1

*Chromosome (Chr) and region were identified from the original papers.[9 10]
†This column describes the nearest gene to the genotyping SNP.
‡ "A" denotes a risk allele.
§We defined the level of significance as ($p<0.05$) and no corrections were applied to the p values.

We analysed, in a Japanese population, a total of 17 candidate variants that had been reported to be strongly associated with CD in the WTCCC and follow-up experiment, and found that three markers, rs6887695, rs10761659 and rs10883365, were significant. Our further genotype–phenotype analysis identified a strong association signal between rs6887695 and the case of pure ileal disease. The first marker, rs6887695, was located 65 kb telomeric to the gene *IL12B* (interleukin 12B). The second marker, rs10761659, was a non-coding intergenic SNP between *ZNF365* (zinc finger protein 365) and *ATQL4* (antiquitin-like 4). The last marker, rs10883365, mapping to chromosome 10q24, was located in the 5 kb upstream region of *NKX2-3* (NK2 transcription factor-related, locus 3). Variants in the *IL12B* gene and the *IL23R* gene were recently reported to be associated with both psoriasis and CD.[16] The positive association of the intergenic region on chromosome 10q21 that includes rs10761659 was recently described in North American cohorts.[17] These two markers were replicated in a Japanese population, although the level of significance was relatively low possibly because of our small sample size compared with the original UK reports. Even so, simulation indicated that our study had a power of >0.70 in the case of SNPs with significant association in the original paper ($p<1\times10^{-7}$). Association of rs6887695 with pure ileal disease was significant even with the small sample size ($n=70$). However, we understand that a further study with a larger sample size will be necessary to confirm the association.

The association of rs10883365 is the most significant SNP that we have studied using information from European CD studies.[18 19] *NKX2-3*-deficient mice show severe defects of gut development, primarily in the epithelium of the small intestine.[20] In addition, the lymphoid organs of these mice, including the spleen and Peyer's patches, have abnormal tissue architecture and abnormalities in the migration and segregation of B and T cells.[21] Most of the CD susceptibility genes reported so far are related to epithelial barrier function or to innate and adaptive immunity. *NKX2-3* must have an important role in the pathogenesis of CD; however, its mechanism is unknown. More biological studies and replication analyses will be required.

The distribution of 17 genetic variants in UK and Japanese populations is summarised in table 4. In terms of minor allele frequencies and of risk alleles, most SNPs revealed significant differences between the two populations. The differences were most notable for rs17234657 and rs17221417, minor alleles of which were very rare or absent in the Japanese. On the other hand, three SNPs (rs6887695, rs10761659 and rs10883365) showed positive associations in both the UK and Japanese populations. Furthermore, in order to study heterogeneity between the Japanese and UK population, we performed a meta-analysis for 11 SNPs by using the random effect model among all three studies (shown in Supplementary table 2 as "All"). We also performed analysis restricted to the two replication studies in order to exclude the possible overestimation by the original study (designated "Rep." in Supplementary table 2). Although the heterogeneities were not significant ($p_H<0.05/30$), the combined OR of meta-analyses were shown to be significant in three SNPs for both "All" and "Rep." (rs9292777, rs6887695 and rs10883365). Two out of these three

Table 3 rs6887695 genotype distribution in CD cases stratified by disease location

	Number of genotypes			p Value			
	CC	CG	GG	Genotype	Allelic	Dominant	Recessive
Pure ileal disease	29 (41.4%)	31 (44.3%)	10 (14.3%)	<0.001	<0.001	0.0027	<0.001
Any ileal disease	68 (25.7%)	127 (47.9%)	70 (26.4%)	NS	NS	NS	NS

NS, not significant ($p\geq0.05$); no corrections were applied to the p values.

Table 4 Summary of genetic differences between UK and Japanese populations in candidate single nucleotide polymorphisms (SNPs) in relation to Crohn's disease susceptibility

		European population				Japanese population			
		Allele*		MAF†		Allele*		MAF†	
SNP	Chr.	Risk	Minor	Case	Control	Risk	Minor	Case	Control
rs11805303	1p31	T	T	0.39	0.32	T	C	0.57	0.55
rs12035082	1q24	C	C	0.44	0.39	C	T	0.97	0.95
rs10801047	1q31	A	A	0.09	0.07	A	A	0.12	0.11
rs10210302	2q37	T	C	0.40	0.48	C	T	0.238	0.239
rs9858542	3p21	A	A	0.33	0.28	G	A	0.036	0.04
rs17234657	5p13	G	G	0.18	0.12	G	G	0	0
rs9292777	5p13	T	C	0.32	0.39	T	T	0.83	0.86
rs10077785	5q23	G	A	0.19	0.23	G	A	0.05	0.04
rs6887695	5q33	C	C	0.34	0.32	C	C	0.47	0.43
rs4958847	5q33	A	A	0.15	0.11	G	G	0.51	0.50
rs13361189	5q33	C	C	0.10	0.07	C	C	0.36	0.33
rs1000113	5q33	T	T	0.10	0.07	T	T	0.34	0.32
rs10761659	10q21	G	A	0.41	0.46	G	A	0.28	0.32
rs10883365	10q24	G	G	0.54	0.48	G	G	0.47	0.41
rs17221417	16q12	G	G	0.36	0.29	C	G	0	0.002
rs2542151	18p11	G	G	0.21	0.16	G	G	0.11	0.10
rs2836754	21q22	T	T	0.40	0.35	C	C	0.75	0.76
rs11209026‡	*IL23R*	A	A	0.07	0.02	A	A	0	0
rs2241880‡	*ATG16L1*	G	A	0.41	0.47	A	G	0.76	0.74
rs1373692‡	5p13.1	C	A	0.59	0.68	C	C	0.83	0.84

Chr., chromosome; MAF, minor allele frequency.

*The minor allele is defined in the controls, and its frequency in that group as well as the case sample is reported. If information on risk and minor alleles in the UK population was unavailable in the original reports,[8 10] we calculated the data using information from the WTCCC database.

†MAFs were calculated as defined in the UK population.

‡The information on these three SNPs was from previous reports; rs11209026 in *IL23R*, rs2241880 in *ATG16L1* and rs1373692 on the chromosome 5p13 gene desert region.[15 23–25]

SNPs were also significant in our replication study, indicating them to be common risk factors of CD among the Japanese and UK population: rs6887695 (5q33, $p_E<0.001$; OR = 1.23 (1.11 to 1.35) under the allelic model designated "Rep.") and rs10883365 (10q24, $p_E<0.001$; OR = 1.21 (1.10 to 1.33) under the allelic model, $p_E<0.001$; OR = 1.34 (1.15 to 1.57) under the recessive model designated "Rep."). The third SNP, rs9292777, showing positive association in the present meta-analysis but negative in our replication study (5p13, $p_E<0.01$; OR = 1.32 (1.20 to 1.46) under the allelic model, $p_E<0.001$; OR = 1.38 (1.16 to 1.64) under the dominant model, $p_E<0.001$; OR = 1.46 (1.25 to 1.69) under the recessive model designated "Rep."), could still be a candidate for a common genetic factor in CD. The associations for four other SNPs (rs10801047, rs10077785, rs4958847 and rs2836754) were probably not present because the strength of their association was less than our proposed threshold and, furthermore, the OR and confidence intervals designated by "Rep." were not significant by all three models. However, we could not conclude for the remaining SNPs whether the association is present or not. We recognise a weakness in the present meta-analysis in the following two points: first, the number of studies included in the analysis was small and most of the samples were derived from a UK population. Secondly, we could not analyse the remaining six SNPs including rs10761659 that showed marginal association in Japanese samples because UK replication data were not available. Further genetic studies in various ethnic groups would be warranted to draw firm conclusions and to reveal the genetic architecture of CD.

Since *CARD15* was identified in 2001 as the first gene conferring susceptibility to CD by the positional cloning approach based on linkage studies,[4 5] a significant number of studies have been performed to identify CD susceptibility genes. In recent GWA studies, several genes have repeatedly been identified as susceptibility genes for CD: *TNFSF15*, *IL23R*, *ATG16L1* and the gene desert region on 5p13.1.[23–28] In addition to the newly identified loci, the WTCCC GWA study has detected previously defined loci such as *IL23R* by rs11805303, *ATG16L1* by rs10210302, the gene desert region on 5p13.1 by rs17234657 and rs9292777, *IBD5* by rs10077785 and *NOD2* by rs17221417. In our previous study, we failed to replicate the positive association of these variants.[15 18 29] In the present study, most of the newly identified loci except the three under discussion had also not been detected in the Japanese population. Thus, we again confirmed the existence of distinct ethnic differences in the genetic background of CD.

In conclusion, we examined whether 17 candidate SNPs identified by WTCCC-GWA scans were associated with susceptibility to CD in Japanese people, and detected positive association for three markers, the most significant of these being rs10883365 in the upstream region of *NKX2-3*. The WTCCC GWA scan has identified some of the common genes involved in susceptibility to CD. However, the major part of the genetic background of CD in Asian populations is still largely unknown, so it is desirable to repeat the GWA studies using the newly developed screening kit in Japanese people. In addition, further genetic studies are required to confirm our findings in other ethnically divergent populations, including other Asian populations.

Acknowledgements: This work was supported by a grant from the Japanese Millennium Project and in part by a Grant-in-Aid for Young Scientists (B) (18790484) from the Ministry of Education, Culture, Sports, Science and Technology (MEXT). We thank Ayumi Kemori, Rie Funahashi and the other members of the Laboratory for Gastrointestinal Diseases for their assistance.

Competing interests: None.

Ethics approval: The study protocol was approved by the local ethics board.

REFERENCES

1. **Podolsky DK.** Inflammatory bowel disease. *N Engl J Med* 2002;**347**:417–29.
2. **Xavier RJ,** Podolsky DK. Unravelling the pathogenesis of inflammatory bowel disease. *Nature* 2007;**448**:427–34.
3. **Vermeire S,** Rutgeerts P. Current status of genetics research in inflammatory bowel disease. *Genes Immun* 2005;**6**:637–45.
4. **Ogura Y,** Bonen DK, Inohara N et al. A frameshift mutation in *NOD2* associated with susceptibility to Crohn's disease. *Nature* 2001;**411**:603–6.
5. **Hugot JP,** Chamaillard M, Zouali H, *et al*. Association of *NOD2* leucine-rich repeat variants with susceptibility to Crohn's disease. *Nature* 2001;**411**:599–603.
6. **The International HapMap Consortium**. A haplotype map of the human genome. *Nature* 2005;**437**:1299–320.
7. **Hirschhorn JN,** Daly MJ. Genome-wide association studies for common diseases and complex traits. *Nat Rev Genet* 2005;**6**:95–108.
8. **Wang WYS,** Barratt BJ, Clayton DG, *et al*. Genome-wide association studies: theoretical and practical concerns. *Nat Rev Genet* 2005;**6**:109–18.
9. **The Wellcome Trust Case Control Consortium**. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 2007;**447**:661–78.
10. **Parkes M,** Barrett JC, Prescott NJ, *et al*. Sequence variants in the autophagy gene *IRGM* and multiple other replicating loci contribute to Crohn's disease susceptibility. *Nat Genet* 2007;**39**:830–2.
11. **Lennard-Jones JE.** Classification of inflammatory bowel disease. *Scand J Gastroenterol Suppl* 1989;**170**:2–6.
12. **Gasche C,** Scholmerich J, Brynskov J, *et al*. A simple classification of Crohn's disease: report of the Working Party for the World Congresses of Gastroenterology, Vienna 1998. *Inflamm Bowel Dis* 2000;**6**:8–15.
13. **Ohnishi Y,** Tanaka T, Ozaki K, *et al*. A high-throughput SNP typing system for genome-wide association studies. *J Hum Genet* 2001;**46**:471–7.
14. **DerSimonian R,** Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986;**7**:177–88.
15. **Yamazaki K,** Onouchi Y, Takazoe M, *et al*. Association analysis of genetic variants in *IL23R, ATG16L1* and 5p13. 1 loci with Crohn's disease in Japanese patients. *J Hum Genet* 2007;**52**:575–83.
16. **Cargill M,** Schrodi SJ, Chang M, *et al*. A large-scale genetic association study confirms *IL12B* and leads to the identification of *IL23R* as psoriasis-risk genes. *Am J Hum Genet* 2007;**80**:273–90.
17. **Rioux JD,** Xavier RJ, Taylor KD, *et al*. Genome-wide association study identifies new susceptibility loci for Crohn disease and implicates autophagy in disease pathogenesis. *Nat Genet* 2007;**39**:596–604.
18. **Yamazaki K,** Takazoe M, Tanaka T, *et al*. Absence of mutation in the *NOD2/CARD15* gene among 483 Japanese patients with Crohn's disease. *J Hum Genet* 2002;**47**:469–72.
19. **Yamazaki K,** Takazoe M, Tanaka T, *et al*. Association analysis of *SLC22A4, SLC22A5* and *DLG5* in Japanese patients with Crohn disease. *J Hum Genet* 2004;**49**:664–8.
20. **Pabst O,** Zweigerdt R, Arnold HH. Targeted disruption of the homeobox transcription factor Nkx2-3 in mice results in postnatal lethality and abnormal development of small intestine and spleen. *Development* 1999;**126**:2215–25.
21. **Pabst O,** Forster R, Lipp M, *et al*. NKX2.3 is required for MAdCAM-1 expression and homing of lymphocytes in spleen and mucosa-associated lymphoid tissue. *EMBO J* 2000;**19**:2015–23.
22. **Yamazaki K,** McGovern D, Ragoussis J, *et al*. Single nucleotide polymorphisms in *TNFSF15* confer susceptibility to Crohn's disease. *Hum Mol Genet* 2005;**14**:3499–506.
23. **Duerr RH,** Taylor KD, Brant SR, *et al*. A genome-wide association study identifies *IL23R* as an inflammatory bowel disease gene. *Science* 2006;**314**:1461–3.
24. **Hampe J,** Franke A, Rosenstiel P, *et al*. A genome-wide association scan of nonsynonymous SNPs identifies a susceptibility variant for Crohn disease in *ATG16L1*. *Nat Genet* 2007;**39**:207–11.
25. **Libioulle C,** Louis E, Hansoul S, *et al*. Novel Crohn disease locus identified by genome-wide association maps to a gene desert on 5p13.1 and modulates expression of *PTGER4*. *PLoS Genet* 2007;**3**:e58.

Editor's quiz: GI snapshot

Robin Spiller, *editor*

Very large gastric ulcer with a round lesion

CLINICAL PRESENTATION

A 56-year-old Japanese woman was admitted to our hospital with a 1-month history of abdominal fullness and a 1-year history of body weight loss of 10 kg. There was no significant family history nor history of previous abdominal surgery. There was no history of any use of alcohol, tobacco, steroids, aspirin or other non-steroidal anti-inflammatory drugs (NSAIDs).

Physical examination showed anaemic conjunctivae and upper abdominal tenderness without rigidity or rebound tenderness. Laboratory studies revealed that she had severe anaemia (haemoglobin (Hb) 6.7 g/dl, haematocrit (Ht) 23.6%).

Upper gastrointestinal endoscopy showed a very large ulcer extending from the lesser curvature of the angulus to the posterior wall of the lower body, with a mucosal bridge in the centre of the ulcer which seemed to divide the ulcer into two parts but under which there was enough room for an endoscope to pass through. A unique shaped elevated lesion was also seen on the ulcer base at the posterior wall (fig 1).

QUESTION

What is the likely diagnosis?

See page 293 for the answer.

This case is submitted by:

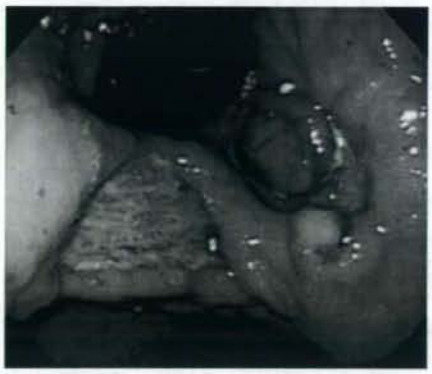


Figure 1 Upper endoscopy showed a very large ulcer with a mucosal bridge and a round lesion.

Y Tatsumi,[1] K Inoue,[1] H Kanemasa,[1] A Toma,[2] K Tsukamoto,[2] Y Kajita[3]

[1] Department of Gastroenterology, Nantan General Hospital, Kyoto, Japan; [2] Department of Surgery, Nantan General Hospital, Kyoto, Japan; [3] Department of Internal Medicine, Nantan General Hospital, Kyoto, Japan

Correspondence to: Dr Y Tatsumi, Department of Gastroenterology, Nantan General Hospital, 25 Yagi-ueno, Yagi-cho, Nantan City, Kyoto Prefecture, 629-0197, Japan; tatsumi-41@ked.biglobe.ne.jp

Competing interests: None.

Patient consent: Patient consent has been received for publication of the case details and the figures in this paper.

Gut 2009;**58**:232. doi:10.1136/gut.2008.159921

nature genetics

Susceptibility loci for intracranial aneurysm in European and Japanese populations

Kaya Bilguvar[1,2], Katsuhito Yasuno[3], Mika Niemelä[4], Ynte M Ruigrok[5], Mikael von und zu Fraunberg[6], Cornelia M van Duijn[7], Leonard H van den Berg[5], Shrikant Mane[8], Christopher E Mason[2,9], Murim Choi[2], Emília Gaál[1,2,4], Yasar Bayri[1,2], Luis Kolb[1,2], Zulfikar Arlier[1,2], Sudhakar Ravuri[8], Antti Ronkainen[6], Atsushi Tajima[3], Aki Laakso[4], Akira Hata[10], Hidetoshi Kasuya[11], Timo Koivisto[6], Jaakko Rinne[6], Juha Öhman[12], Monique M B Breteler[7], Cisca Wijmenga[13,14], Matthew W State[2,9], Gabriel J E Rinkel[5], Juha Hernesniemi[4], Juha E Jääskeläinen[6], Aarno Palotie[15,16], Ituro Inoue[3], Richard P Lifton[2,17] & Murat Günel[1,2]

Stroke is the world's third leading cause of death. One cause of stroke, intracranial aneurysm, affects ~2% of the population and accounts for 500,000 hemorrhagic strokes annually in mid-life (median age 50), most often resulting in death or severe neurological impairment[1]. The pathogenesis of intracranial aneurysm is unknown, and because catastrophic hemorrhage is commonly the first sign of disease, early identification is essential. We carried out a multistage genome-wide association study (GWAS) of Finnish, Dutch and Japanese cohorts including over 2,100 intracranial aneurysm cases and 8,000 controls. Genome-wide genotyping of the European cohorts and replication studies in the Japanese cohort identified common SNPs on chromosomes 2q, 8q and 9p that show significant association with intracranial aneurysm with odds ratios 1.24–1.36. The loci on 2q and 8q are new, whereas the 9p locus was previously found to be associated with arterial diseases, including intracranial aneurysm[2–5]. Associated SNPs on 8q likely act via *SOX17*, which is required for formation and maintenance of endothelial cells[6–8], suggesting a role in development and repair of the vasculature; *CDKN2A* at 9p may have a similar role[9]. These findings have implications for the pathophysiology, diagnosis and therapy of intracranial aneurysm.

Siblings of intracranial aneurysm probands are at ~fourfold increased risk of hemorrhage from intracranial aneurysm, suggesting a genetic component to risk[10]. Genome-wide linkage studies of familial cases[11] and rare apparently mendelian kindreds have not thus far identified robustly replicable loci, and no underlying mutations have been identified[12–14]. Similarly, examination of candidate genes in small case-control studies has failed to produce replicable results[12].

These considerations motivate the use of GWAS to identify common variants that contribute to intracranial aneurysm. We carried out a multistage intracranial aneurysm GWAS in three cohorts: a Finnish cohort of 920 cases and 985 controls, a Dutch cohort of 781 cases and 6,424 controls and a Japanese cohort of 495 cases and 676 controls (see **Supplementary Methods** online).

The study design consisted of a first stage of genome-wide genotyping of the European cohorts on the Illumina platform, careful matching of cases and controls, and identification of intervals harboring SNPs that surpassed a significance threshold of 5×10^{-7} for association with intracranial aneurysm[2]. This discovery phase had 80% power to detect common alleles that confer a genotype relative risk (GRR) of 1.31 and 50% power to detect a GRR of 1.25 (assuming an additive model in log-odds scale). Replication of association of SNPs in these intervals was tested in the Japanese cohort, setting $P < 0.05$ for significant replication. The replication study had

[1]Departments of Neurosurgery, Neurobiology and [2]Genetics, Yale Program on Neurogenetics, Yale Center for Human Genetics and Genomics, Yale University School of Medicine, New Haven, Connecticut 06510, USA. [3]Division of Molecular Life Science, School of Medicine, Tokai University, Shimokasuya 143, Isehara, Kanagawa 259-1193, Japan. [4]Department of Neurosurgery, Helsinki University Central Hospital, Helsinki, P.O. Box 266, FI-00029 HUS, Finland. [5]Department of Neurology, Rudolf Magnus Institute of Neuroscience, University Medical Center Utrecht, 3584 CX Utrecht, The Netherlands. [6]Department of Neurosurgery, Kuopio University Hospital, Kuopio FI-70211, Finland. [7]Genetic Epidemiology Unit, Department of Epidemiology and Biostatistics and Department of Clinical Genetics, Erasmus Medical Center, 2040, 3000 CA Rotterdam, The Netherlands. [8]Keck Foundation Biotechnology Resource Laboratory, Yale University, 300 George Street, New Haven, Connecticut 06510, USA. [9]Child Study Center, Yale University School of Medicine, New Haven, Connecticut 06510, USA. [10]Department of Public Health, School of Medicine, Chiba University, Chiba 260-8670, Japan. [11]Department of Neurosurgery, Medical Center East, Tokyo Women's University, Tokyo 116-8567, Japan. [12]Department of Neurosurgery, Tampere University Hospital, 33521 Tampere, Finland. [13]Complex Genetics Section, Department of Biomedical Genetics, University Medical Center Utrecht, 3508 AB Utrecht, The Netherlands. [14]Department of Genetics, University Medical Center Groningen and University of Groningen, 9700 RR Groningen, The Netherlands. [15]Biomedicum Helsinki, Research Program in Molecular Medicine, University of Helsinki, 00290 Helsinki, Finland. [16]Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1HH, UK. [17]Howard Hughes Medical Institute and Department of Internal Medicine, Yale University School of Medicine, New Haven, Connecticut 06510, USA. Correspondence should be addressed to R.P.L. (richard.lifton@yale.edu) or M.G. (murat.gunel@yale.edu).

Received 7 July; accepted 18 August; published online 9 November 2008; doi:10.1038/ng.240

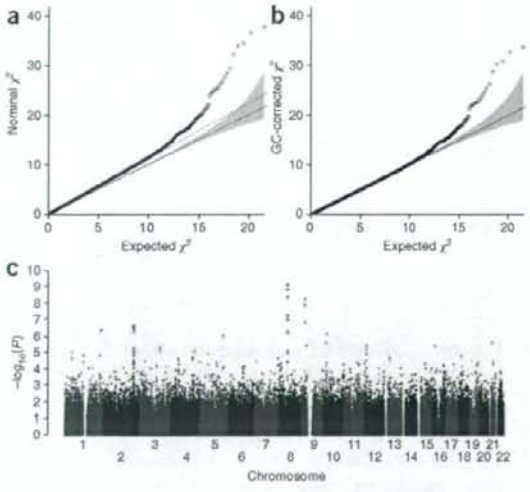


Figure 1 Genome-wide association of SNPs with intracranial aneurysm in the combined European cohort. (**a**) Quantile-quantile plot of the observed χ^2 values derived from the Mantel-extension test statistics versus the expected χ^2 distribution. The solid line represents concordance of observed and expected values. The slope of the dashed line represents the genomic inflation factor ($\lambda = 1.11$). (**b**) The plot of expected and observed χ^2 values for association of SNPs with intracranial aneurysm after correction for the genomic inflation factor ($\lambda = 1.0$). Significant deviation from the expected values suggests association of these SNPs with intracranial aneurysm phenotype. (**c**) The $-\log_{10}$ of uncorrected *P* values for association of each SNP and intracranial aneurysm is plotted according to its physical position on successive chromosomes. Green dots indicate SNPs yielding *P* values $< 1 \times 10^{-5}$, and red dots denote SNPs that surpass a significance level of 5×10^{-7}.

80% and 66% power to replicate SNPs with GRRs of 1.31 and 1.25, respectively. The utility of using a genetically diverse population for replication has been demonstrated by recent studies[15], thereby extending association results to a broad segment of the world's population.

Discovery phase genotypes were processed using rigorous quality controls; because Dutch controls and some Finnish controls were genotyped separately on Illumina chips of varying SNP density, particular attention was paid to ensuring consistent genotyping performance and excluding nonrandom genotyping error within and across cohorts (**Supplementary Methods** and **Supplementary Tables 1** and **2** online). To control for population stratification, we genetically matched cases and controls from each cohort[16], resulting in a dataset in which cases and controls are similarly distributed along axes of significant principal components (**Supplementary Table 1**).

We tested for association between each SNP and intracranial aneurysm by using the Cochran-Armitage trend test in each cohort and combined the results using the Mantel extension test. The distribution of test statistics for association of SNPs with intracranial aneurysm in the combined cohort is shown in **Figure 1a**. The genomic inflation factor (λ) was 1.043 and 1.136 for the Finnish and Dutch, respectively, and 1.114 combined, indicating well-matched populations[17] (**Fig. 1a**); further logistic regression including principal components as covariates[2] did not significantly change λ (**Supplementary Fig. 1** online), nor did exclusion of SNPs with call rates <99% in any case or control cohort (**Supplementary Methods**); in contrast, because genomic inflation factor increases with sample size[18], the large Dutch control sample was a major contributor to λ (**Supplementary Methods**). The association results reveal a number of SNPs whose *P* values exceed those expected under the null hypothesis; these persist after correction for λ (**Fig. 1b**). The *P* values across each chromosome are shown in **Figure 1c**. Four intervals (on 1q, 2q, 8q and 9p) harbored SNPs that surpassed the threshold for genome-wide significance; these include multiple SNPs with correlated *P* values and comprise 15 of the 16 SNPs with $P < 10^{-6}$ (**Fig. 1c**). Associated SNPs in each interval have very high call rates in every cohort and none violate HWE in any cohort (**Supplementary Table 2**). The first three loci have not previously shown association with intracranial aneurysm or other diseases, whereas SNPs on 9p are in the block of linkage disequilibrium (LD) that has previously been shown to be associated with myocardial infarction[2–4], abdominal aortic aneurysm and intracranial aneurysm[5]. Both Finnish and Dutch cohorts contributed to the significance of each locus, the risk alleles were identical and their odds ratios were not significantly different between cohorts (**Table 1**).

To attempt to replicate these four loci, we genotyped 15 SNPs from these intervals in the Japanese cohort (**Supplementary Tables 2** and **3** online). Eight of the 15 SNPs showed significant association with intracranial aneurysm; these included SNPs on 2q, 8q and 9p (**Table 1**). At each locus, SNPs in strong LD in the Japanese sample showed highly correlated *P* values (**Fig. 2**). For associated SNPs, risk alleles in Japan were identical to and showed similar odds ratios to those found in Europe (**Table 1** and **Supplementary Table 3**). Using the Mantel extension test to combine data from all three cohorts, we found the following *P* values and odds ratios for the SNPs showing the strongest evidence for association at each locus: 2q, $P = 4.4 \times 10^{-8}$ (odds ratio (OR) = 1.24); 8q, $P = 1.4 \times 10^{-10}$ (OR = 1.36); 9p, $P = 1.4 \times 10^{-10}$ (OR = 1.29) (**Table 1**). No locus showed significant deviation from an additive model (log-odds scale) (**Supplementary Table 3**).

We examined the distributions of *P* values in each significant interval. At 2q, association in Europe lies within a large block of LD (197.8–198.6 Mb; **Fig. 2** and **Supplementary Table 4** online). In Asian subjects, this segment is divided into two smaller blocks of LD and the association seen in Japan is confined to SNPs in the more telomeric block (198.2–198.5 Mb). This interval contains four known genes; the two most strongly associated SNPs, rs700651 and rs700675, lie in introns of adjacent genes, *BOLL* and *PLCL1*. PLCL1 is of interest because it has significant homology to phospholipase C, which lies downstream of VEGFR2 signaling[19]. VEGFR2 is a marker of endothelial progenitor cells and has a role in central nervous system angiogenesis[20].

The LD structure at 8q is also of interest (**Fig. 2** and **Supplementary Table 4**). SNP rs10958409 shows the most significant association; SNPs in high LD with rs10958409 show correlated *P* values. In addition, however, rs9298506, which lies 110 kb distally and shows virtually no LD with rs10958409 ($r^2 = 0.004$ in European HapMap subjects[21], 0.004 in Finnish cases and 0.0005 in Dutch cases) nonetheless also revealed significant association in Europeans; adjacent SNPs in LD showed correlated *P* values. This observation suggests the presence of two independent risk alleles. A conditional test of association demonstrated that after accounting for the association with rs9298506, rs10958409 still showed significant association with intracranial aneurysm (and vice versa), consistent with two risk loci (**Supplementary Table 4**). The Japanese cohort replicated association at rs10958409, but not rs9298506, despite having had 88% power to detect association of this latter SNP (**Supplementary Table 3**). Further work will be required to determine whether the European association with rs9298506 is a true positive result. This 8q interval contains a

Table 1 Summary of results for five SNPs that characterize the association with intracranial aneurysm on chromosomes 2, 8 and 9

Chr.	SNP	Pos. (Mb)	Risk allele	Dataset	RAF (control/case)	*P* value[a]	Odds ratio (95% CI)			Heterogeneity *P*[b]	Dom. *P*[c]	PAF(%)	RRF(%)
							Per allele	Heterozygous	Homozygous				
2	rs1429412	197.9	G	Finland	0.42/0.48	4.4×10^{-4}	1.27 (1.11–1.45)	1.33 (1.07–1.65)	1.60 (1.22–2.10)	–	0.60	–	–
				Netherlands	0.34/0.39	1.6×10^{-4}	1.25 (1.11–1.40)	1.35 (1.14–1.61)	1.48 (1.15–1.90)	–	0.21	–	–
				Europe	–	2.5×10^{-7}	1.26 (1.15–1.37)	1.34 (1.18–1.54)	1.54 (1.28–1.85)	0.85	0.21	–	–
				Japan	0.29/0.30	0.42	1.08 (0.90–1.30)	1.05 (0.82–1.35)	1.21 (0.78–1.86)	–	0.76	–	–
				All	–	5.8×10^{-7}	1.22 (1.13–1.32)	1.27 (1.13–1.43)	1.46 (1.24–1.73)	0.32	0.40	–	–
	rs700651	198.3	G	Finland	0.39/0.44	5.6×10^{-3}	1.21 (1.06–1.39)	1.19 (0.96–1.46)	1.48 (1.12–1.96)	–	0.81	14.1	0.7
				Netherlands	0.35/0.40	5.0×10^{-4}	1.23 (1.09–1.38)	1.22 (1.03–1.46)	1.50 (1.18–1.91)	–	0.99	14.1	0.7
				Europe	–	8.9×10^{-6}	1.22 (1.12–1.33)	1.21 (1.06–1.38)	1.49 (1.25–1.79)	0.89	0.87	–	–
				Japan	0.46/0.54	0.0011	1.30 (1.11–1.53)	1.05 (0.79–1.41)	1.70 (1.24–2.33)	–	0.08	14.8	1.6
				All	–	4.4×10^{-8}	1.24 (1.15–1.34)	1.18 (1.04–1.33)	1.56 (1.34–1.83)	0.77	0.30	–	–
8	*rs10958409*	55.5	A	Finland	0.18/0.22	1.4×10^{-3}	1.31 (1.11–1.55)	1.40 (1.15–1.71)	1.34 (0.81–2.21)	–	0.20	11.7	0.9
				Netherlands	0.15/0.20	1.4×10^{-7}	1.46 (1.27–1.68)	1.42 (1.20–1.69)	2.28 (1.51–3.45)	–	0.65	12.1	1.7
				Europe	–	1.6×10^{-9}	1.39 (1.25–1.55)	1.42 (1.25–1.61)	1.83 (1.31–2.55)	0.34	0.64	–	–
				Japan	0.25/0.30	0.016	1.26 (1.04–1.51)	1.21 (0.94–1.55)	1.66 (1.07–2.59)	–	0.66	10.9	0.8
				All	–	1.4×10^{-10}	1.36 (1.24–1.49)	1.37 (1.22–1.54)	1.79 (1.37–2.33)	0.40	0.77	–	–
	rs9298506	55.6	A	Finland	0.73/0.80	4.6×10^{-7}	1.50 (1.28–1.75)	1.26 (0.82–1.94)	1.99 (1.31–3.01)	–	0.41	38.5	2.0
				Netherlands	0.81/0.85	2.3×10^{-4}	1.34 (1.15–1.56)	1.50 (0.87–2.59)	1.97 (1.15–3.35)	–	0.66	–	–
				Europe	–	8.6×10^{-10}	1.41 (1.27–1.58)	1.37 (0.98–1.91)	1.95 (1.41–2.70)	0.32	0.84	–	–
				Japan	0.81/0.83	0.25	1.14 (0.92–1.41)	0.59 (0.30–1.13)	0.78 (0.41–1.47)	–	0.04	–	–
				All	–	1.8×10^{-9}	1.35 (1.22–1.49)	1.16 (0.87–1.56)	1.63 (1.23–2.16)	0.13	0.29	–	–
9	*rs1333040*	22.1	T	Finland	0.47/0.52	2.8×10^{-3}	1.22 (1.07–1.40)	1.23 (0.98–1.54)	1.50 (1.15–1.95)	–	0.96	18.5	0.7
				Netherlands	0.55/0.62	9.5×10^{-7}	1.33 (1.19–1.50)	1.38 (1.09–1.75)	1.80 (1.41–2.31)	–	0.73	30.2	1.4
				Europe	–	1.5×10^{-8}	1.29 (1.18–1.40)	1.30 (1.10–1.53)	1.66 (1.39–1.98)	0.34	0.89	–	–
				Japan	0.65/0.72	0.0024	1.32 (1.10–1.58)	1.26 (0.83–1.91)	1.69 (1.13–2.55)	–	0.81	29.3	1.1
				All	–	1.4×10^{-10}	1.29 (1.19–1.40)	1.29 (1.11–1.50)	1.67 (1.42–1.96)	0.61	1.00	–	–

SNPs that replicate with $P < 0.05$ in the Japanese cohort are italicized. Position shown is from NCBI build 36 coordinates. Risk allele is indexed to the forward strand of NCBI build 36.
[a]The Cochran-Armitage trend test *P* value for cohorts from Finland, Netherlands and Japan, respectively, and the Mantel extension test (uncorrected for λ) *P* value for combined cohorts. [b]*P* value of the test of heterogeneity in effect size among populations. [c]*P* value of the test of deviation from an additive model. PAF, population attributable fraction. RRF, recurrence risk fraction attributable to each SNP, assuming the overall sibling recurrence risk of 4 (ref. 10).

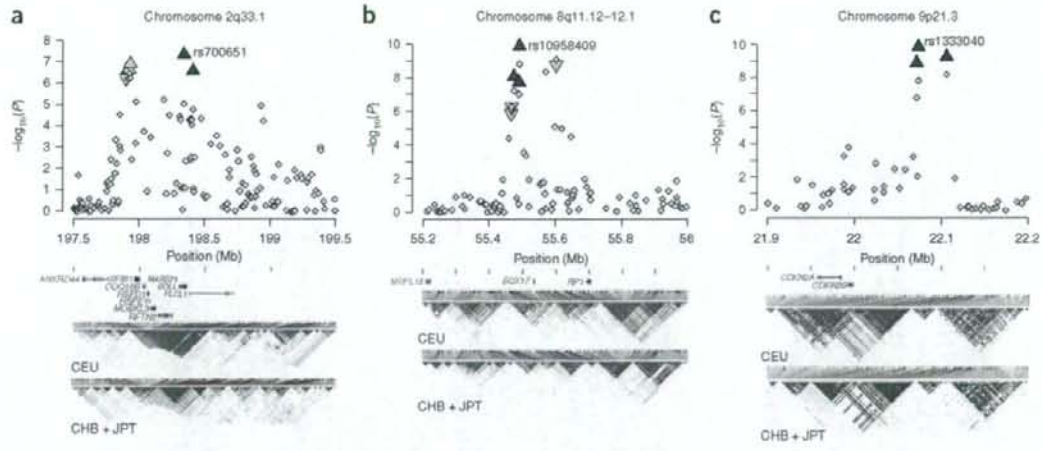


Figure 2 Regional association plots and linkage disequilibrium structure. (**a**–**c**) The $-\log_{10}$ of the *P* value for association of each SNP and intracranial aneurysm in discovery phase across segments of 2q (**a**), 8q (**b**) and 9q (**c**) are shown as small diamonds (using NCBI build 36 for map locations). Fifteen SNPs that were genotyped in the Japanese replication cohort are shown as triangles that show the combined (discovery + replication) *P* values: blue triangles represent SNPs that demonstrate replication in the Japanese cohort with $P < 0.05$, and gray triangles denote SNPs with $P > 0.05$ in the replication study. SNPs with the most significant *P* values in each interval in the combined analysis are marked with their SNP IDs. Known transcripts (RefSeq database) are represented as horizontal bars at the bottom of each panel. Population-specific LD structures based on D′ are shown for the HapMap European (CEU) and Asian (CHB + JPT) cohorts[21]. The results demonstrate that on chromosome 2, SNPs spanning an ~800-kb interval are in strong LD in the CEU population and show evidence of association with intracranial aneurysm in the Finnish and Dutch cohorts. In Asia, this segment is broken into two smaller blocks of LD that are not strongly correlated with one another, and significant association with intracranial aneurysm in Japanese cohort is seen only for the telomeric segment. For chromosome 8, SNPs in two blocks that are not in significant LD with one another both show significant association with intracranial aneurysm in the European cohort; only SNPs located within the proximal block replicate in Japan. Cohort-specific r^2 values among all SNPs genotyped in the replication studies are shown in **Supplementary Table 4**.

single gene, *SOX17*, which lies between these two association peaks, 43 kb from rs10958409 and 64 kb from rs9298506. The next closest genes lie 201 kb distal and 266 kb proximal to rs10958409. Sox17 has an important role in formation and maintenance of the endothelium (see below).

Finally, SNPs on 9p that showed association with intracranial aneurysm (22.07–22.10 Mb) (**Fig. 2** and **Supplementary Table 4**) were in LD with SNPs that have previously shown association with multiple arterial diseases[2–4]. Adjacent SNPs that are associated with type 2 diabetes mellitus[22–24] showed no significant association with intracranial aneurysm. The strongest association was with rs1333040, which lies 74 kb from the 5′ end of *CDKN2B* and 88 kb from *CDKN2A*. These genes encode the cyclin-dependent kinase inhibitors $p15^{INK4b}$ and $p16^{INK4a}$, as well as ARF, a regulator of p53 activity. In addition, a non-protein-coding transcript (*ANRIL*) lies within this interval[25]. Among these, $p16^{INK4a}$ is of particular interest (see below).

To determine whether the effects of these loci are influenced by known risk factors, we examined the odds ratios of the most significant allele at each locus after partitioning cases by gender, family history of intracranial aneurysm, age (older half versus younger) and ruptured versus unruptured aneurysm. The results showed no significant difference in odds ratios after any of these partitions, suggesting independent contributions to risk (**Supplementary Table 5** online).

Finally, to assess the combined effects of the three loci, we defined each subject's risk score by summing the logarithm of the odds ratio for each risk allele they harbor as determined in each cohort. The observed intracranial aneurysm risk showed a significant linear relationship with risk score in each cohort, with a more than threefold increase from lowest to highest strata (**Table 2** and **Supplementary Table 6** online).

This study provides the first results of a large GWAS of intracranial aneurysm or stroke. Three significant loci have been identified. These results cannot be explained by nonrandom genotyping error or population stratification and are robust to alternative analyses (**Supplementary Fig. 1**). We calculate that these loci collectively account for 38–46% of the population-attributable fraction of intracranial aneurysm and 2.3–3.8% of the sibling recurrence risk (**Table 1**). Additional common variants are likely to have a role in intracranial aneurysm, as the study was not well powered to find loci with GRR < 1.25. In addition, population-specific effects were not considered in this study design. Given the risk allele frequencies and odds ratios of the identified loci, future replication cohorts will require ~900 to 1,600 cases and controls to have 80% power for replication ($\alpha = 0.05$).

After genomic control correction and exclusion of SNPs at the four top loci, 37 SNPs remained with *P* values less than 10^{-4} (28 are expected by chance). Some of these may prove to be true risk alleles as additional cohorts are evaluated, as has occurred with type 2 diabetes[26]. In addition, rare variants with larger effects at these same loci may also contribute to the occurrence of intracranial aneurysm[27,28].

Intracranial aneurysms predominate at arterial branch points and sites of shear stress, locations that incur endothelial damage. Vascular injury mobilizes bone marrow–derived cells that localize to these sites and contribute to repair[29,30]. *SOX17*, a member of the Sry-related HMG box transcription factor family, is of particular interest because it is required for both endothelial formation and maintenance[6–8]. Sox17 plays a key role in the generation and maintenance of fetal and neonatal stem cells of both hematopoietic and endothelial lineages[8] and is expressed in adult endothelium[6]. $Sox17^{-/-}$ mice show multiple vascular abnormalities[7]; moreover, whereas $Sox18^{-/-}$ mice are normal, $Sox18^{-/-};Sox17^{+/-}$ mice show defective endothelial sprouting and

Table 2 Increased intracranial aneurysm risk with increased risk score based on genotypes for rs700651, rs10958409 and rs133040

	Japan			Netherlands			Finland		
No. of risk alleles	Frequency (control/case)	Average risk score (min–max)	OR (95% CI)	Frequency (control/case)	Average risk score (min–max)	OR (95% CI)	Frequency (control/case)	Average risk score (min–max)	OR (95% CI)
0 or 1	0.14/0.08	0.23 (0.00–0.28)	1	0.3/0.21	0.22 (0.00–0.38)	1	0.32/0.21	0.17 (0.00–0.27)	1
2	0.28/0.21	0.53 (0.46–0.55)	1.29 (0.81–2.05)	0.36/0.29	0.54 (0.41–0.75)	1.13 (0.9–1.41)	0.34/0.34	0.42 (0.38–0.54)	1.47 (1.15–1.88)
3	0.33/0.32	0.79 (0.72–0.82)	1.63 (1.05–2.55)	0.25/0.33	0.82 (0.70–1.04)	1.88 (1.51–2.34)	0.24/0.3	0.63 (0.59–0.74)	1.82 (1.41–2.36)
4	0.19/0.28	1.05 (0.99–1.08)	2.41 (1.51–3.83)	0.08/0.14	1.10 (0.98–1.33)	2.40 (1.82–3.16)	0.07/0.13	0.85 (0.79–0.94)	2.83 (1.98–4.04)
5 or 6	0.06/0.10	1.34 (1.26–1.54)	2.91 (1.62–5.21)	0.02/0.03	1.41 (1.36–1.74)	2.63 (1.56–4.41)	0.03/0.02	1.09 (1.06–1.33)	1.16 (0.62–2.19)
P value[a]		3.3×10^{-7}			8.4×10^{-16}			7.8×10^{-8}	
OR (95% CI)[b]		2.91 (1.92–4.41)			2.81 (2.19–3.61)			2.88 (1.95–4.25)	

Individuals with non-missing genotypes for all six alleles were included in the analysis.
[a]*P* value for linear relationship of risk score and logarithm of ORs (complete data set is listed in **Supplementary Table 6**). [b]Increase in odds ratio per one unit change in risk score.

vascular remodeling[6]. Similarly, $p16^{INK4a}$ has a role in regulation of stem (progenitor) cell populations, including bone marrow–derived cells of the vasculature[9]. These considerations suggest that intracranial aneurysm may result from defective stem (progenitor) cell–mediated vascular development and/or repair.

Finally, these findings have implications for identification of individuals with intracranial aneurysm before morbid events. The odds ratio of intracranial aneurysm increases greater than threefold in subjects with the highest versus the lowest risk (**Table 2** and **Supplementary Table 6**). Although we caution that further work is required, these findings advance the potential for preclinical diagnosis by combined assessment of inherited susceptibility with previously established risk factors.

METHODS

Cohorts. The study protocol was approved by the Yale Human Investigation Committee (HIC protocol 7680). In all cases, the diagnosis of intracranial aneurysm was made with computerized tomography angiogram, magnetic resonance angiogram or cerebral digital subtraction angiogram and confirmed at surgery, when applicable. Rupture of aneurysm was defined by identification of acute subarachnoid hemorrhage (via computerized tomography or magnetic resonance imaging) from a proven aneurysm. Cases with a first-degree relative with intracranial aneurysm were considered familial, and other cases were considered sporadic.

Three cohorts from independent studies in Finland, The Netherlands and Japan were collected and all participants provided informed consent. There were 960 Finnish cases and 1,017 controls; 786 Dutch cases and 6,424 controls; and 495 Japanese cases and 676 controls. Japanese controls were screened for not harboring intracranial aneurysm.

Genotyping and SNP quality control. Genome-wide genotyping in European cohorts was done on the Illumina platform according to the manufacturer's protocol (Illumina). We genotyped subjects on either the CNV370-Duo, HumanHap300 or HumanHap550 chips. SNPs shared across all platforms ($n = 314{,}125$) were extracted. We applied prespecified criteria to exclude samples and SNPs that performed poorly as well as samples that could not be genetically well matched (**Supplementary Table 1** and **Supplementary Methods**). The overall median genotype call rate was 99.7% and the mean heterozygosity of all SNPs was 35%. Seventy-two duplicate pairs of samples were genotyped and showed 99.91% genotype identity. We carried out detailed analysis of the performance of SNPs across cohorts and platforms to ensure that significant associations observed were not due to differences in SNP performance (**Supplementary Table 2**).

Cryptic relatedness. We determined the identity by state (IBS) similarity and estimated the degree of relatedness for each pair of samples in the GWAS (**Supplementary Methods**) and excluded inferred first- and second-degree relatives (**Supplementary Table 1**).

Analysis of population structure. In order to identify population outliers and cases whose genetic ancestry cannot be properly matched to controls (and vice versa), we used the Genetic Matching (GEM) method described previously[16] based on principal component analysis (PCA). After this matching process, three significant principal components remained in the Finnish cohort and none in the Dutch cohort, as previously observed (**Supplementary Methods**).

After quality control and analysis of population structure, there remained 874 cases and 944 controls in the Finnish cohort and 706 cases and 5,332 controls in the Dutch cohort. Among the Finnish cases, 57% were female; 73% had suffered ruptured aneurysm and 43% had positive family history; the median age at diagnosis was 50 years (those with rupture 49 years versus those without rupture 52 years). In the Dutch cohort, 69% were female, 92% had ruptured aneurysm, 15% had a positive family history and the median age was 49 years.

SNP association analysis. To test for association of each SNP with intracranial aneurysm, we assumed an additive (in log-odds scale) model. We used the Cochran-Armitage trend test for each cohort. For the combined sample of European descent or of European and Japanese cohorts, we used the Mantel extension test (**Supplementary Methods**).

We calculated the per-allele and genotype-specific ORs and their 95% confidence intervals by fitting 1-d.f. and 2-d.f. logistic models, respectively. We assessed heterogeneity of ORs among populations by considering the likelihood ratios of a logistic model with population by genotype interaction term(s) versus a linear model without the interaction term(s) and used a *P* value <0.05 as evidence of significant heterogeneity (**Supplementary Table 3**). To evaluate the degree of overdispersion of test statistics, we calculated the genomic inflation factor (λ) for each statistical test by the ratio of the mean of the lower 90% of observed test statistics to that of the expected χ^2 values[17]. We applied the genomic control method to correct for λ (**Fig. 1b**) and then compared a pairwise plot of *P* values for each SNP in the trend and corrected tests to determine the potential effect of any residual population stratification (**Supplementary Fig. 1a,b** and **Supplementary Methods**).

We also examined the validity of the assumption of additivity (in log-odds scale) in the association tests by comparing likelihood ratios assuming

alternative models of dominance and rejected additivity for $P < 0.05$ (ref. 2 and **Supplementary Table 3**).

For each chromosome segment showing significant association with intracranial aneurysm, we investigated whether more than one SNP had an independent marginal effect on intracranial aneurysm by the Mantel extension test conditioned on genotypes for SNPs within each interval (**Supplementary Table 4**).

To assess the robustness of our GWAS results, we also performed a weighted *Z*-score test and found that the results of this alternative analysis were highly correlated with the results of the Mantel extension test (**Supplementary Fig. 1c**).

Replication study in Japanese cohort. For the Japanese replication study, allelic discrimination assays were done with 15 SNPs on the Sequenom iPLEX genotyping platform according to the manufacturer's protocol. For SNPs that showed significant *P* values, genotypes were repeated and *P* values confirmed on the TaqMan platform (Applied Biosystems). Association tests were done as described above, using $P = 0.05$ (in the Cochran-Armitage trend test with the same allele found associated in Europe) as the threshold for significance (**Supplementary Table 3**).

Subset analysis. For SNPs with the most significant *P* values we investigated whether the association results were affected by potential confounding variables such as rupture status, family history or gender. We compared genotype distributions of cases stratified by these variables using the trend test (**Supplementary Table 5**).

Population-attributable fraction and proportion of genetic variance attributable to SNPs. We investigated two risk measures based on replicated SNPs: the population attributable fraction (PAF) and the proportion of the sibling recurrence risk attributable to a SNP ('recurrence risk fraction') as previously described (**Table 1** and **Supplementary Methods**). For these calculations we assumed intracranial aneurysm population prevalence of 2% and λ_{sib} of 4 (ref. 10). The combined contribution of SNPs was obtained by assuming the multiplicative model (**Supplementary Methods**).

Cumulative effects of risk alleles. We analyzed the cumulative effects of the risk alleles at the most significant SNP at 2q, 8q and 9p (rs700651, rs10958409 and rs1333040) by calculating the risk score for each individual by the weighted sum of the number of risk alleles as defined by

$$\text{Risk score} = \sum_i \psi[i] n[i]$$

where $\psi[i]$ is the logarithm of the calculated per-allele odds ratio at each locus and $n[i]$ is the number of risk alleles at the same locus. We then assessed the risk score for each of the 27 possible three-locus genotypes in each cohort (**Supplementary Table 6**). We fitted a simple linear logistic model with an additive effect (on log-odds scale) for each cohort and performed a likelihood-ratio test. For display purposes, the 27 strata of **Supplementary Table 6** are compressed into 5 strata shown in **Table 2** according to the absolute number of risk alleles, which closely parallels the risk score.

Note: Supplementary information is available on the Nature Genetics website.

ACKNOWLEDGMENTS

We are grateful to the participants who made this study possible. We thank A. Chamberlain, O. Törnwall, M. Alalahti, K. Helin, S. Malin and J. Budzinack for their technical help. This study was supported by the Yale Center for Human Genetics and Genomics and the Yale Program on Neurogenetics, the US National Institutes of Health grants R01NS057756 (M.G.) and U24 NS051869 (S.M.) and the Howard Hughes Medical Institute (R.P.L.). C.E.M. and M.W.S. are supported by a gift from the Lawrence Family and Y.M.R. by the Dr E. Dekker program of The Netherlands Heart Foundation (2005T014). K.Y. and I.I. were supported by the Core Research for Evolutional Science and Technology, Japan Science and Technology Corporation.

AUTHOR CONTRIBUTIONS

Cohort ascertainment, characterization and DNA preparation: M.N., E.G., A.L., A.P., J.Ö. and J.H. (Helsinki); M.v.u.z.F., A.R., T.K., J.R., A.P. and J.E.J. (Kuopio); Y.M.R., L.H.v.d.B., C.W. and G.J.E.R. (Utrecht); C.M.v.D. and M.M.B.B. (Rotterdam) and A.T., A.H., H.K. and I.I. (Japan). Genotyping: K.B., Y.B., L.K., Z.A., S.R., R.P.L., M.G. and S.M. Study design and analysis plan: R.P.L. and M.G. Data management and informatics: C.E.M., K.B., M.W.S. and M.G. Statistical analysis: K.Y., K.B., I.I., M.C., R.P.L. and M.G. Writing team: K.B., K.Y., M.W.S., M.G. and R.P.L.

Published online at http://www.nature.com/naturegenetics/
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/

1. Bederson, J.B. *et al.* Recommendations for the management of patients with unruptured intracranial aneurysms: a Statement for healthcare professionals from the Stroke Council of the American Heart Association. *Stroke* **31**, 2742–2750 (2000).
2. Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* **447**, 661–678 (2007).
3. McPherson, R. *et al.* A common allele on chromosome 9 associated with coronary heart disease. *Science* **316**, 1488–1491 (2007).
4. Helgadottir, A. *et al.* A common variant on chromosome 9p21 affects the risk of myocardial infarction. *Science* **316**, 1491–1493 (2007).
5. Helgadottir, A. *et al.* The same sequence variant on 9p21 associates with myocardial infarction, abdominal aortic aneurysm and intracranial aneurysm. *Nat. Genet.* **40**, 217–224 (2008).
6. Matsui, T. *et al.* Redundant roles of Sox17 and Sox18 in postnatal angiogenesis in mice. *J. Cell Sci.* **119**, 3513–3526 (2006).
7. Sakamoto, Y. *et al.* Redundant roles of Sox17 and Sox18 in early cardiovascular development of mouse embryos. *Biochem. Biophys. Res. Commun.* **360**, 539–544 (2007).
8. Kim, I., Saunders, T.L. & Morrison, S.J. Sox17 dependence distinguishes the transcriptional regulation of fetal from adult hematopoietic stem cells. *Cell* **130**, 470–483 (2007).
9. Janzen, V. *et al.* Stem-cell ageing modified by the cyclin-dependent kinase inhibitor p16INK4a. *Nature* **443**, 421–426 (2006).
10. Schievink, W.I., Schaid, D.J., Michels, V.V. & Piepgras, D.G. Familial aneurysmal subarachnoid hemorrhage: a community-based study. *J. Neurosurg.* **83**, 426–429 (1995).
11. Foroud, T. *et al.* Genome screen to detect linkage to intracranial aneurysm susceptibility genes: the Familial Intracranial Aneurysm (FIA) study. *Stroke* **39**, 1434–1440 (2008).
12. Nahed, B.V. *et al.* Mapping a mendelian form of intracranial aneurysm to 1p34.3-p36.13. *Am. J. Hum. Genet.* **76**, 172–179 (2005).
13. Ozturk, A.K. *et al.* Molecular genetic analysis of two large kindreds with intracranial aneurysms demonstrates linkage to 11q24–25 and 14q23–31. *Stroke* **37**, 1021–1027 (2006).
14. Ruigrok, Y.M. *et al.* Genomewide linkage in a large Dutch family with intracranial aneurysms: replication of 2 loci for intracranial aneurysms to chromosome 1p36.11-p36.13 and Xp22.2-p22.32. *Stroke* **39**, 1096–1102 (2008).
15. Gudbjartsson, D.F. *et al.* Variants conferring risk of atrial fibrillation on chromosome 4q25. *Nature* **448**, 353–357 (2007).
16. Luca, D. *et al.* On the use of general control samples for genome-wide association studies: genetic matching highlights causal variants. *Am. J. Hum. Genet.* **82**, 453–463 (2008).
17. Clayton, D.G. *et al.* Population structure, differential bias and genomic control in a large-scale, case-control association study. *Nat. Genet.* **37**, 1243–1246 (2005).
18. Freedman, M.L. *et al.* Assessing the impact of population stratification on genetic association studies. *Nat. Genet.* **36**, 388–393 (2004).
19. Shibuya, M. Differential roles of vascular endothelial growth factor receptor-1 and receptor-2 in angiogenesis. *J. Biochem. Mol. Biol.* **39**, 469–478 (2006).
20. Ziegler, B.L. *et al.* KDR receptor: a key marker defining hematopoietic stem cells. *Science* **285**, 1553–1558 (1999).
21. Frazer, K.A. *et al.* A second generation human haplotype map of over 3.1 million SNPs. *Nature* **449**, 851–861 (2007).
22. Zeggini, E. *et al.* Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. *Science* **316**, 1336–1341 (2007).
23. Scott, L.J. *et al.* A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. *Science* **316**, 1341–1345 (2007).
24. Saxena, R. *et al.* Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science* **316**, 1331–1336 (2007).
25. Pasmant, E. *et al.* Characterization of a germ-line deletion, including the entire INK4/ARF locus, in a melanoma-neural system tumor family: identification of ANRIL, an antisense noncoding RNA whose expression coclusters with ARF. *Cancer Res.* **67**, 3963–3969 (2007).
26. Zeggini, E. *et al.* Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. *Nat. Genet.* **40**, 638–645 (2008).
27. Romeo, S. *et al.* Population-based resequencing of ANGPTL4 uncovers variations that reduce triglycerides and increase HDL. *Nat. Genet.* **39**, 513–516 (2007).
28. Ji, W. *et al.* Rare independent mutations in renal salt handling genes contribute to blood pressure variation. *Nat. Genet.* **40**, 592–599 (2008).
29. Asahara, T. *et al.* Isolation of putative progenitor endothelial cells for angiogenesis. *Science* **275**, 964–967 (1997).
30. Purhonen, S. *et al.* Bone marrow-derived circulating endothelial precursors do not contribute to vascular endothelium and are not needed for tumor growth. *Proc. Natl. Acad. Sci. USA* **105**, 6620–6625 (2008).

Association of *TNFRSF4* gene polymorphisms with essential hypertension

Yoichi Mashimo[a], Yoichi Suzuki[a], Kazuko Hatori[a], Yasuharu Tabara[b], Tetsuro Miki[c], Katsushi Tokunaga[d], Tomohiro Katsuya[e], Toshio Ogihara[e], Michiko Yamada[f], Norio Takahashi[g], Yoshio Makita[h], Tomohiro Nakayama[i], Masayoshi Soma[j], Nobuhito Hirawa[k], Satoshi Umemura[k], Takayoshi Ohkubo[l], Yutaka Imai[m] and Akira Hata[a]

Background **Essential hypertension is a complex disorder that results from the interaction of a number of susceptibility genes and environmental factors. The *TNFRSF4* (tumor necrosis factor receptor superfamily, member 4) gene was one of the genes that showed altered renal expression in long-term salt loading in mice. Moreover, association of the *TNFRSF4* and *TNFSF4* (tumor necrosis factor (ligand) superfamily, member 4) genes with myocardial infarction was recently reported. Since essential hypertension is a well-known risk factor for myocardial infarction, we hypothesized that *TNFRSF4* could be a susceptibility gene for essential hypertension.**

Methods **We performed a case–control study of *TNFRSF4* in two independent population.**

Results **Extensive investigation of single nucleotide polymorphisms of the entire gene suggested that it resided in one linkage disequilibrium block, and four single nucleotide polymorphisms in the 5′ flanking region sufficiently represented major haplotypes. In the combined population, the frequency of the most frequent haplotype, C-C-A-A, was significantly lower ($P = 8.07 \times 10^{-5}$) and that of the second most frequent haplotype, C-T-G-A, was significantly higher ($P = 6.07 \times 10^{-4}$) in hypertensive subjects than in control subjects. This difference was observed only in female patients. The C-T-G-A haplotype showed a lower promoter activity than other haplotypes, suggesting a relationship with disease susceptibility.**

Conclusion **Our results suggest that *TNFRSF4* is a female-specific susceptible gene for essential hypertension.** ***J Hypertens*** **26:902–913 © 2008 Wolters Kluwer Health | Lippincott Williams & Wilkins.**

Journal of Hypertension 2008, 26:902–913

Keywords: association studies, essential, haplotype, hypertension, single nucleotide polymorphisms, *TNFRSF4* gene

Abbreviations: 95% CI, 95% confidence interval; *Agt*, mouse angiotensinogen gene; ANOVA, analysis of variance; APC, antigen-presenting cell; CRP, C-reactive protein; DBP, diastolic blood pressure; EH, essential hypertension; HT, hypertensive; LD, linkage disequilibrium; MI, myocardial infarction; NT, normotensive; OR, odds ratio; PCR, polymerase chain reaction; SBP, systolic blood pressure; SD, standard deviation; SNP, single nucleotide polymorphism; TaqMan-ASA, TaqMan allele-specific amplification; TNFRSF4 (OX40), tumor necrosis factor receptor superfamily, member; 4TNFSF4 (OX40L), tumor necrosis factor (ligand) superfamily, member 4; UTR, untranslated region

[a]Department of Public Health, Graduate School of Medicine, Chiba University, Chiba, [b]Department of Medical Genomics, [c]Department of Geriatric Medicine, School of Medicine, Ehime University, Ehime, [d]Department of Human Genetics, Graduate School of Medicine, University of Tokyo, Tokyo, [e]Department of Geriatric Medicine, Graduate School of Medicine, Osaka University, Osaka, [f]Department of Clinical Studies, [g]Department of Genetics, Radiation Effects Research Foundation, Hiroshima, [h]Department of Pediatrics, Asahikawa Medical College, Asahikawa, [i]Division of Molecular Diagnostics, Department of Advanced Medical Science, [j]Division of Nephrology and Endocrinology, Department of Medicine, Nihon University School of Medicine, Tokyo, [k]Second Department of Internal Medicine, Graduate School of Medicine, Yokohama City University, Kanagawa, [l]Department of Planning for Drug Development and Clinical Evaluation and [m]Department of Clinical Pharmacology and Therapeutics, Graduate School of Pharmaceutical Sciences and Medicine, Tohoku University, Sendai, Japan

Correspondence to Akira Hata MD, PhD, Department of Public Health, Graduate School of Medicine, Chiba University, 1-8-1 Inohana, Chuo-ku, Chiba 260-8670, Japan
Tel: +81 43 2262067; fax: +81 43 2262070; e-mail: ahata@faculty.chiba-u.jp

Received 20 June 2007 **Revised** 30 November 2007
Accepted 4 January 2008

Introduction

Hypertension affects more than 25% of the adult population worldwide [1]. Essential hypertension (EH) accounts for more than 90% of hypertension cases and is a multifactorial disorder resulting from the interaction of a number of susceptibility genes and environmental factors. It is estimated that the genetic contribution to blood pressure variation ranges from 30 to 50% [2]. Identification of susceptibility genes for hypertension would provide a clue to the pathophysiology of the disease.

Several approaches exists for genetic causes of EH: candidate–gene linkage studies, genome–scanning linkage studies, candidate–gene association studies, genetic studies in animal models, and gene expression profiling in animal models [3]. Each approach has its own strengths and weaknesses, and some argue that integration of the approaches is a more efficient way forward [4]. The Millennium Genome Project for Hypertension in Japan has adopted the candidate–gene association strategy because of its relatively higher statistical power and

convenience of collecting samples [5]. Candidate genes are selected on the basis of the accumulation of experimental evidence (expression profiling in animal models) and information in the literature. As a first step in this project, we performed DNA microarray experiments in mice to screen genes whose renal expression was changed by long-term salt loading, because genes that showed salt sensitivity were considered to be candidate genes for EH. The results showed that more than 300 genes were either downregulated or upregulated. For the genetic association study, from these 300 genes, we nominated 121 that had been reported in the literature as candidate genes. To date, 70 genes have been screened, 10 of which showed significant association with EH on haplotype-based analysis. Three of these 10 genes were positive in both the expression profiling and genetic association studies. The tumor necrosis factor receptor superfamily, member 4 (*TNFRSF4*) gene was one of the three.

TNFRSF4 (OX40) is a member of the tumor necrosis factor receptor (TNFR) superfamily, and is primarily expressed as a transmembrane protein on activated $CD4^+$ T cells after antigen recognition [6–9]. Tumor necrosis factor (ligand) superfamily, member 4 (TNFSF4, also called OX40L) [10], the ligand for TNFRSF4 on activated $CD4^+$ T cells, is expressed on antigen-presenting cells (APCs) including activated B cells, macrophages, and dendritic cells, as well as on endothelial cells and some activated T cells [11–14]. The TNFRSF4–TNFSF4 interaction between T cell and APC contributes to proinflammatory T-cell function. In particular, TNFRSF4–TNFSF4 interactions are crucial for the generation of memory $CD4^+$ T cells by promoting the survival of effector T cells [15–18]. Thus, it is suggested that the TNFRSF4–TNFSF4 pathway is involved in inflammation and immune response.

T lymphocyte activation involving several receptor–ligand pairs such as TNFRSF4–TNFSF4 is suggested to promote atherosclerosis [12,19,20], which is now considered to be an inflammatory disease [21]. Recently, *TNFSF4* was identified as a susceptibility gene for atherosclerosis and a genetic variation in *TNFSF4* was reported to be associated with myocardial infarction (MI) and severity of coronary artery disease [22]. Genetic variation in *TNFRSF4* was also shown to be associated with MI [23]. These reports suggested that the TNFRSF4–TNFSF4 pathway plays an important role in the pathogenesis of atherosclerosis and MI in humans. It is generally believed that hypertension is one of the major risk factors for atherosclerosis and MI [24]; however, MI and hypertension often coexist, as seen in the SHEEP study cohort in which MI patients were significantly associated with hypertension [25]. Thus, the association between MI and TNFRSF4/TNFSF4 in human subjects may be due to not only atherogenesis but also hypertension itself. We hypothesized that *TNFRSF4* and/or *TNFSF4* were potential candidate genes for EH.

The aim of the present study was to investigate the association between genetic variations of the *TNFRSF4* gene and EH in the Japanese population. We performed a case–control study using two independent population of Japanese patients with EH.

Methods

Study subjects

Initial screening of candidate genes involved 1035 subjects with EH (762 men and 273 women) and 1058 age-matched controls (792 men and 266 women) who were recruited through the study group of the Millennium Genome Project for Hypertension [5]. Six medical institutes took part in the collaborative study and recruited subjects in Japan. Recruitment procedures, case–control criteria, and clinical characteristics are described in detail elsewhere [5].

The clinical characteristics of the subjects included in this study for *TNFRSF4* gene analyses are shown in Tables 1 and 2. Subjects in population 1 were part of the population recruited through the study group of the Millennium Genome Project for Hypertension [5]. Subjects in population 2 were recruited from Ohasama, a cohort in a rural community of northern Japan [26].

Each subject was assigned to one of the blood pressure diagnostic categories defined by the criteria of the 1999 WHO/ISH guidelines for the management of hypertension [27]. Hypertensive (HT) subjects had systolic blood pressure (SBP) of at least 140 mmHg or diastolic blood

Table 1 Characteristics of subjects in population 1

	Total subjects			Male subjects			Female subjects		
Parameters	NT	HT	*P*	NT	HT	*P*	NT	HT	*P*
No. of subjects	562	587		301	316		261	271	
Age (years)	61.6 ± 9.2	60.1 ± 11.2	0.011*	59.9 ± 9.0	58.5 ± 11.1	0.083	63.6 ± 9.1	62.0 ± 11.1	0.056
BMI (kg/m^2)	22.2 ± 2.8	23.9 ± 3.3	< 0.001*	22.1 ± 2.9	23.8 ± 3.1	< 0.001*	22.3 ± 2.7	24.0 ± 3.6	< 0.001*
SBP (mmHg)	111.7 ± 8.9	163.7 ± 21.1	< 0.001*	111.8 ± 8.8	162.1 ± 18.4	< 0.001*	111.5 ± 9.1	166.1 ± 24.4	< 0.001*
DBP (mmHg)	68.9 ± 7.3	98.3 ± 14.8	< 0.001*	69.4 ± 7.3	98.7 ± 14.0	< 0.001*	68.0 ± 7.3	97.6 ± 16.0	< 0.001*
TC (mg/dl)	205.5 ± 38.0	207.1 ± 34.9	0.596	195.8 ± 35.7	198.7 ± 33.0	0.294	216.2 ± 37.6	216.0 ± 34.7	0.970
HDL-C (mg/dl)	57.3 ± 15.1	58.3 ± 17.2	0.314	55.3 ± 15.0	56.8 ± 17.5	0.263	59.5 ± 15.0	60.0 ± 16.7	0.697
TG (mg/dl)	123.8 ± 87.4	141.7 ± 84.7	0.003*	132.1 ± 106.2	147.0 ± 94.5	0.133	116.7 ± 67.4	135.4 ± 71.3	0.007*

Values are mean ± SD. BMI, body mass index; DBP, diastolic blood pressure; HDL-C, HDL cholesterol; HT, hypertensive patient; NT, normotensive patient; SBP, systolic blood pressure; TC, total cholesterol; TG, triglyceride. *Difference was statistically significant.

Table 2 Characteristics of subjects in population 2

	Total subjects			Male subjects			Female subjects		
Parameters	NT	HT	P	NT	HT	P	NT	HT	P
No. of subjects	925	732		317	323		608	409	
Age (years)	54.6 ± 11.5	61.6 ± 9.7	< 0.001*	55.8 ± 11.1	61.5 ± 10.2	< 0.001*	54.0 ± 11.6	61.7 ± 9.3	< 0.001*
BMI (kg/m²)	23.4 ± 3.1	24.2 ± 3.3	< 0.001*	23.5 ± 3.0	23.6 ± 3.1	0.506	23.4 ± 3.1	24.6 ± 3.4	< 0.001*
SBP (mmHg)	123.9 ± 9.8	142.2 ± 12.1	< 0.001*	125.4 ± 8.6	143.9 ± 11.4	< 0.001*	123.1 ± 10.3	140.8 ± 12.5	< 0.001*
DBP (mmHg)	70.3 ± 7.1	80.2 ± 9.1	< 0.001*	71.6 ± 6.9	81.9 ± 9.4	< 0.001*	69.6 ± 7.2	78.8 ± 8.6	< 0.001*
TC (mg/dl)	193.4 ± 34.2	195.0 ± 33.8	0.358	186.4 ± 33.8	183.9 ± 34.0	0.352	197.1 ± 33.9	203.8 ± 31.0	0.001*
HDL-C (mg/dl)	55.3 ± 14.1	53.8 ± 14.6	0.028*	51.2 ± 14.1	52.8 ± 14.4	0.180	57.4 ± 13.7	54.4 ± 14.7	0.001*
TG (mg/dl)	128.9 ± 73.4	142.5 ± 89.8	0.001*	139.1 ± 85.1	145.3 ± 103.0	0.340	123.7 ± 66.1	139.5 ± 77.9	0.001*

Values are mean ± SD. BMI, body mass index; DBP, diastolic blood pressure; HDL-C, HDL cholesterol; HT, hypertensive patient; NT, normotensive patient; SBP, systolic blood pressure; TC, total cholesterol; TG, triglyceride. *Difference was statistically significant.

pressure (DBP) of at least 90 mmHg or were patients currently taking chronic antihypertensive medication. Normotensive (NT) subjects had SBP/DBP lower than 140/90 mmHg and had never been treated with antihypertensive medication. Informed consent was obtained from each individual as per the protocol approved by each institution's ethics committee.

DNA microarray experiments in mice

In DNA microarray experiments, we used two lines of mice having different numbers of the functional mouse angiotensinogen gene (*Agt*) [28,29], kindly donated by Professor Oliver Smithies (Department of Pathology, University of North Carolina, Chapel Hill, North Carolina, USA). To observe distinct effects by long-term salt loading, *Agt* 2/2 mice (with four wild-type copies of the *Agt* gene) were fed a high-salt diet containing 8% NaCl for 6 months, whereas *Agt* 0/1 mice (with one wild-type copy of the *Agt* gene) were fed a low-salt diet containing 0.3% NaCl. Total RNA was isolated from the kidneys of mice and differences in gene expression were examined using mouse cDNA microarray (Incyte Genomics Inc., Palo Alto, California, USA), which contains 9222 mouse cDNA clones.

Screening of candidate genes

We selected a total of 121 candidate genes (Supplemental Table S1) based on the following criteria: (1) genes reported as candidates in the literature or with functions relevant to the blood pressure regulation and (2) human homologue of genes in which renal expression was changed by long-term salt loading in mice. For an initial screening of these candidate genes, some of the available single nucleotide polymorphisms (SNPs) per gene were selected from the Japanese SNP database (http://snp.ims.u-tokyo.ac.jp/) or dbSNP database (http://www.ncbi.nlm.nih.gov/SNP/) and were genotyped in 1035 patients and 1058 controls using the PCR–SSP–FCS method [30]. Haplotype-based association analyses were performed using SNPAlyze v4.1 Pro software (DYNACOM, Mobara, Japan) based on an expectation/maximization (EM) algorithm. *P* values for overall distribution of haplotypes were calculated by permutation test at 1000 iterations. *P* values less than 0.05 were considered statistically significant.

Screening for polymorphisms in *TNFRSF4*

To identify genetic variants of the human *TNFRSF4* gene, we sequenced all seven exons, the adjacent intronic sequence, 4 kb of the 5′ flanking region, and 1.5 kb of the 3′ flanking region in 32 control subjects. Nineteen primer sets were designed on the basis of the *TNFRSF4* genomic and mRNA sequences from the GenBank database (accession numbers NT_004350.18 and NM_003327, respectively). All polymerase chain reaction (PCR) products were sequenced using the BigDye Terminator v3.1 Cycle Sequencing kit and an ABI PRIZM 3100 Genetic Analyzer (Applied Biosystems, Foster City, California, USA). The sequences were analyzed and polymorphisms identified using the Genetyx program (Genetyx Corp., Tokyo, Japan).

Genotyping of polymorphisms in *TNFRSF4*

Genotyping of four SNPs in the *TNFRSF4* gene (P1: −3948C>T, P2: −3606C>T, P8: −1725A>G and P12: −530A>G) was performed using either the TaqMan allele-specific amplification (TaqMan-ASA) method [31] or the Custom TaqMan Genomic Assays kit (Applied Biosystems). In the TaqMan-ASA method, specific primers were designed on the basis of the *TNFRSF4* genomic sequence from the GenBank database (accession number NT_004350.18). The primer sequences are shown in Table 3. The PCR mixture for the TaqMan-ASA method contained 5 μl of 2× TaqMan Universal Mix (Applied Biosystems), 0.4 μmol/l of each PCR primer, 0.12 μmol/l of TaqMan probe, and 5 ng of template DNA in a final volume of 10 μl. The samples were analyzed with an ABI PRIZM 7000 Sequence Detection system (Applied Biosystems). The thermoprofiles were 50°C for 2 min, 95°C for 10 min, followed by 40 cycles of 95°C for 15 s and 60°C for 1 min.

Luciferase assay

TNFRSF4 reporter constructs of 3970 bp (nt −3968 to +2) were created by means of PCR amplification of genomic DNA from homozygous subjects who had alternative haplotypes with the use of following primers: forward, 5′-GGGGTACCGTGCCACATGGCTGGAATTTAC-3′ (including *KpnI* site) and reverse, 5′-TCTAGCTAGC GTCTCTGCTGTCGCCAGAGTC-3′ (including *NheI*

Table 3 Primer sequence (5′→3′) for TaqMan-ASA genotyping

SNP	Allele-specific primer	Common primer	TaqMan probe*
P1	CACATGGCTGGAATTTACCATC CACATGGCTGGAATTTACCTCT	CTCAGCAGTGGGAGAAAAACAA	CCTCTGAAGCGTTTTCACTGGTATCATGTGT
P2	GTCGCCTTTCCCCCTCCG GTCGCCTTTCCCCCTCCA	GCTGCAGCCAATAGGCACCTT	AATAGCCACTTCGTGCGGCTGG
P8	GTCACAGGTCCAAGAAAGCCGT GTCACAGGTCCAAGAAAGCCGC	GCAGGCTGCCTTACAGACCTT	TGAGCTCTGGGTCAGTGTCCA
P12	GGTCAGGAGTTCAAGACCAGTGT GGTCAGGAGTTCAAGACCAGTTC	CCACGCCCGAATAATTTTTG	AGTAGAGACGGGATTTCGCCATGTTAGC

*TaqMan probes contained a 5′ FAM (6-carboxyfluorescein) reporter fluorophore and a 3′ TAMRA (6-carboxytetramethylrhodamine) quencher.

site). Amplicons of three haplotypes (Pr-H1, Pr-H2, and Pr-H5) were cloned into the pGL4.10[luc2] vector (Promega, Madison, Wisconsin, USA). Promoter constructs that contained one polymorphic change (Pr-P2-T, Pr-P3-T, Pr-P4-del, Pr-P6-G, Pr-P8-G, Pr-P9-G, Pr-P10-T, and Pr-P11-G) were created by site-directed mutagenesis carried out in the Pr-H1 plasmid using the GeneEditor *in vitro* site-directed mutagenesis system (Promega). All constructs were verified by sequencing. COS-7 cells (monkey kidney, SV40 T antigen transformed) were cultured in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum and antibiotics. HEK293 cells (human embryonic kidney) were cultured in minimum essential medium supplemented with 2 mmol/l L-glutamine, 1% nonessential amino acids, 10% fetal bovine serum, and antibiotics. Cells in 12-well plates at 50–70% confluence were transfected with 500 ng of each construct and 10 ng of pGL4.74[hRluc/TK] *Renilla* luciferase vector (Promega) as an internal control for transfection efficiency, using 1.5 μl of FuGENE 6 transfection reagent (Roche Diagnostics, Basel, Switzerland). After 24 h of transfection, the cells were harvested, and firefly and *Renilla* luciferase activities were measured using the Dual-Luciferase Reporter Assay System and a TD-20/20 luminometer (Promega). Each experiment was repeated five or six times, and each sample was studied in triplicate.

Statistical analysis

Haploview version 3.32 (http://www.broad.mit.edu/mpg/haploview/index.php) was used to analyze and visualize the linkage disequilibrium (LD) and haplotypic patterns. Hardy–Weinberg equilibrium was assessed by χ^2 analysis. Overall distributions of the genotypes or alleles were analyzed by χ^2 analysis using 2×3 or 2×2 contingency tables between NT controls and HT patients. Haplotype frequencies were estimated using SNPAlyze v4.1 Pro software. The distributions of each haplotype between NT controls and HT patients were calculated both by χ^2 tests of one haplotype against the others (haplotype-wise test) and by permutation tests with 1000 iterations using SNPAlyze software. We calculated odds ratios (ORs) with 95% confidence intervals (CIs) using logistic regression analyses with or without clinical covariates (age, BMI, total cholesterol, high-density lipoprotein cholesterol, and triglyceride). To estimate the contribution of the gene to the total variance of blood pressure, the variance component procedure with the analysis of variance (ANOVA) type III variance estimates was used. Comparisons in reporter assays were performed using Student's *t*-test or ANOVA. All statistical analyses were performed with SPSS software (SPSS Japan Inc., Tokyo, Japan) unless otherwise stated. *P* values less than 0.05 were considered statistically significant.

Results

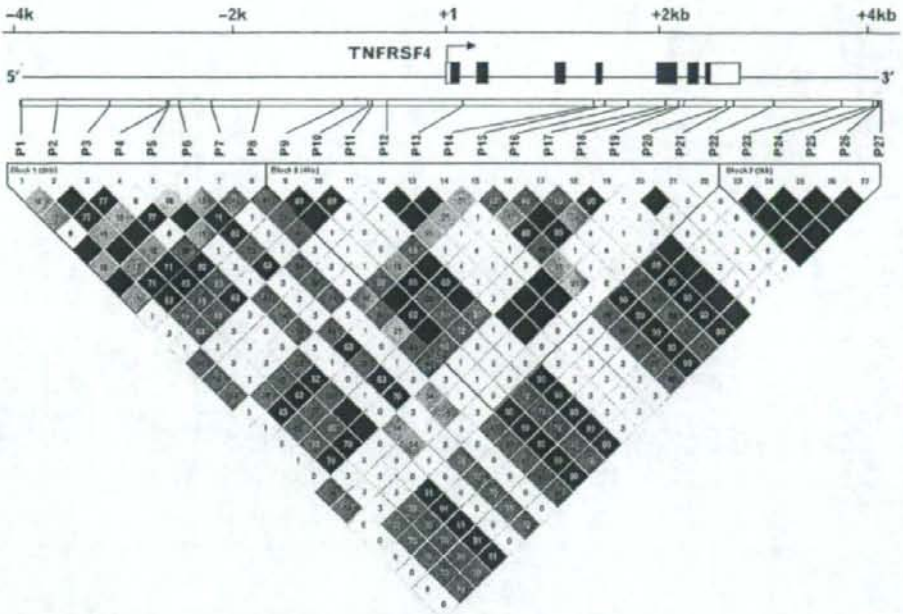
DNA microarray experiments in mice

We used cDNA microarray analyses to compare the expression profiles of 9222 genes in the kidneys of *Agt* 2/2 mice (with four wild-type copies of the *Agt* gene) with a high-salt diet versus those of *Agt* 0/1 mice (with one wild-type copy of the *Agt* gene) with a low-salt diet. Differential expression values greater than 1.3 based on internal quality control data are summarized in Supplemental Tables S2 and S3. We found that 119 genes were downregulated in the kidneys of *Agt* 2/2 mice by 1.3–3.1-fold compared with *Agt* 0/1 mice and 192 genes were upregulated by 1.3–1.9-fold. Murine *TNFRSF4* gene (*Tnfrsf4*) was the gene downregulated 1.3-fold.

Screening of candidate genes by haplotype association study

We selected a total of 121 candidate genes (Supplemental Table S1) on the basis of the following criteria: (1) genes reported as candidates in the literature or with possible involvement of blood pressure regulation and (2) human homologue of genes in which renal expression was changed by long-term salt loading in mice. We excluded genes whose genotype data were not available due to the following reasons: no SNP data was available in the databases; minor allele frequencies of SNPs in Japanese were too low (<5%); or the genotyping of some SNPs was difficult. So far, 191 SNPs in 70 genes have been successfully genotyped for genetic association tests, and the genotyping of only a single SNP was completed in eight genes. A haplotype-based association test was performed in 62 genes and a single SNP association study in eight genes. *P* values for difference in overall distribution of the haplotype or genotype frequencies between normotension and hypertension in total (men + women), male,

Fig. 1



Haplotype block structure of the *TNFRSF4* gene. (Top) Organization of the *TNFRSF4* gene. Exons are indicated by boxes (black, coding sequences; white, untranslated sequences). (Bottom) Linkage disequilibrium structure of polymorphisms across the *TNFRSF4* gene region using data from 32 Japanese controls. Haplotype blocks were defined by the solid spine of LD method in Haploview. The number in each cell represents the LD parameter r^2 (100×), blank cells denote $r^2 = 1$. Each cell is painted with graduated color relative to the strength of LD between markers, which is defined by the r^2 value.

and female subjects are shown in Supplemental Table S4. Significant *P* values were observed for 10 genes: *Aquaporin-2* (*AQP2*), *Estrogen receptor 2* (*ESR2*), *Glycogen synthase 1* (*GYS1*), *Kallikrein 1* (*KLK1*), *Nephrin* (*NPHN*), *Solute carrier family 1 (glial high affinity glutamate transporter), member 2* (*SLC1A2*), *Solute carrier family 9 (sodium/hydrogen exchanger), isoform 3* (*SLC9A3*), *Steroidogenic acute regulatory protein* (*STAR*), *Syntaxin binding protein 1* (*STXBP1*), and *TNFRSF4*. Three genes (*STAR*, *STXBP1*, and *TNFRSF4*) are the human homologues to the mouse genes that showed changes in renal expression in the salt-loading experiment. The *P* value for overall distribution of the haplotype of *TNFRSF4* was significant only in female subjects.

Identification of polymorphisms in *TNFRSF4*

We searched for polymorphisms in the *TNFRSF4* gene, including 4 kb of the 5′ flanking region and 1.5 kb of the 3′ flanking region. By direct sequencing in 32 Japanese individuals, a total of 44 polymorphisms were identified; 20 in the 5′ flanking region, four in exons, seven in introns, and 13 in the 3′ flanking region. Of those, 27 polymorphisms (P1–P27) with minor allele frequencies (MAF) 5% of higher (in 32 DNA samples) are presented in Table 4. A graphical overview of the structure of the human *TNFRSF4* gene showing the location of the 27 polymorphisms identified in this study is shown in Fig. 1. Pairwise LD measuring r^2 between polymorphisms and defined haplotype block structures in this region was evaluated using the solid spine of LD method in Haploview (Fig. 1). Three haplotype blocks (blocks 1, 2, and 3) were defined in the *TNFRSF4* gene region with this method. Blocks 1 and 2 appear to be separated because P8 showed low LD to other polymorphisms and blocks 2 and 3 were separated by P22 for the same reason. Strong LDs, however, were observed among certain blocks, such as between P4 and P27 ($r^2 = 0.91$). In addition, multiallelic D′ values between these blocks were high (0.86 between blocks 1 and 2; 1.0 between blocks 2 and 3). Thus, we decided to handle an entire gene region as one block, which could be analyzed by tag SNPs from the entire region. Four SNPs in the 5′ flanking region (P1: −3948C>T, P2: −3606C>T, P8: −1725A>G, and P12: −530A>G) were employed for further analysis. The four SNP haplotypes constructed

Table 4 Polymorphisms with minor allele frequencies 5% or higher detected in the *TNFRSF4* genomic region in 32 Japanese controls

Name	Polymorphism[a]	Location	Amino acid change	MAF[b]	dbSNP ID	JSNP ID
P1	−3943C>T	5′ Flanking		0.06		
P2	−3601C>T	5′ Flanking		0.27	rs12036216	
P3	−3119G>T	5′ Flanking		0.27	rs11721	
P4	−2577delA	5′ Flanking		0.22		
P5	−2568C>G	5′ Flanking		0.06		
P6	−2461C>G	5′ Flanking		0.27		
P7	−2167C>T	5′ Flanking		0.06		
P8	−1720A>G	5′ Flanking		0.30	rs3813201	JST-IMS173304
P9	−936A>G	5′ Flanking		0.19	rs34115518	
P10	−699C>T	5′ Flanking		0.16	rs35339498	
P11	−669C>G	5′ Flanking		0.19	rs35659545	
P12	−525A>G	5′ Flanking		0.11	rs35107976	
P13	150+47G>C	Intron1		0.11	rs35737009	
P14	376-16C>G	Intron3		0.11	rs34108055	
P15	442+32ins35bp	Intron4		0.25		
P16	442+248C>T	Intron4		0.19	rs9661697	
P17	539G>A	Exon5	Glu178Glu	0.25	rs17568	
P18	639+25C>T	Intron5		0.19	rs2298212	JST-IMS053053
P19	640-31T>G	Intron5		0.20	rs2298211	JST-IMS053052
P20	921C>T	Exon7 (3′ UTR)		0.11	rs2298210	JST-IMS053051
P21	989C>G	Exon7 (3′ UTR)		0.11	rs2298209	JST-IMS053050
P22	1067+308G>A	3′ Flanking		0.08	rs2298208	JST-IMS053049
P23	1067+941G>C	3′ Flanking		0.20	rs34067070	
P24	1067+1224delTT	3′ Flanking		0.20		
P25	1067+1240G>C	3′ Flanking		0.20	rs34279802	
P26	1067+1266T>C	3′ Flanking		0.20	rs35916760	
P27	1067+1296C>T	3′ Flanking		0.20	rs36057244	

[a] Numbering according to the cDNA sequence of *TNFRSF4* (accession number NM_003327). [b] Minor allele frequency (MAF) on the basis of the sequencing of 32 DNA samples.

from these SNPs covered more than 85% of haplotype diversity of the entire *TNFRSF4* gene when P22 was not included for analysis.

Case–control study for *TNFRSF4* polymorphisms

The clinical characteristics of the NT and HT subjects in population 1 are summarized in Table 1. Difference in age between the NT and HT subjects was significant when men and women were jointly compared ($P=0.011$), whereas it was not significant when men and women were separately compared.

In population 1, four SNPs (P1, P2, P8, and P12) were genotyped in 562 NT controls and 587 HT patients. All

Table 5 Genotype and allele frequencies among normotensive (NT) and hypertensive (HT) subjects in population 1

		Total subjects			Male subjects			Female subjects		
	Genotype	NT ($n=562$)	HT ($n=587$)	P[a]	NT ($n=301$)	HT ($n=316$)	P[a]	NT ($n=261$)	HT ($n=271$)	P[a]
P1	CC	448 (0.799)	455 (0.776)		238 (0.793)	245 (0.775)		210 (0.805)	210 (0.778)	
	CT	109 (0.194)	124 (0.212)		60 (0.200)	67 (0.212)		49 (0.188)	57 (0.211)	
	TT	4 (0.007)	7 (0.012)	0.524	2 (0.007)	4 (0.013)	0.691	2 (0.008)	3 (0.011)	0.722
	Allele									
	C	1005 (0.896)	1034 (0.882)		536 (0.893)	557 (0.881)		469 (0.898)	477 (0.883)	
	T	117 (0.104)	138 (0.118)	0.305	64 (0.107)	75 (0.119)	0.506	53 (0.102)	63 (0.117)	0.429
P2	CC	324 (0.578)	319 (0.544)		166 (0.553)	170 (0.538)		158 (0.605)	149 (0.552)	
	CT	208 (0.371)	220 (0.375)		121 (0.403)	117 (0.370)		87 (0.333)	103 (0.381)	
	TT	29 (0.052)	47 (0.080)	0.129	13 (0.043)	29 (0.092)	0.055	16 (0.061)	18 (0.067)	0.455
	Allele									
	C	856 (0.763)	858 (0.732)		453 (0.755)	457 (0.723)		403 (0.772)	401 (0.743)	
	T	266 (0.237)	314 (0.268)	0.089	147 (0.245)	175 (0.277)	0.203	119 (0.228)	139 (0.257)	0.263
P8	AA	284 (0.506)	280 (0.478)		144 (0.480)	154 (0.487)		140 (0.536)	126 (0.467)	
	AG	238 (0.424)	248 (0.423)		137 (0.457)	125 (0.396)		101 (0.387)	123 (0.456)	
	GG	39 (0.070)	58 (0.099)	0.182	19 (0.063)	37 (0.117)	0.044	20 (0.077)	21 (0.078)	0.250
	Allele									
	A	806 (0.718)	808 (0.689)		425 (0.708)	433 (0.685)		381 (0.730)	375 (0.694)	
	G	316 (0.282)	364 (0.311)	0.129	175 (0.292)	199 (0.315)	0.376	141 (0.270)	165 (0.306)	0.202
P12	AA	401 (0.716)	393 (0.671)		209 (0.699)	215 (0.680)		192 (0.736)	178 (0.659)	
	AG	148 (0.264)	175 (0.299)		84 (0.281)	88 (0.278)		64 (0.245)	87 (0.322)	
	GG	11 (0.020)	18 (0.031)	0.179	6 (0.020)	13 (0.041)	0.318	5 (0.019)	5 (0.019)	0.144
	Allele									
	A	950 (0.848)	961 (0.820)		502 (0.839)	518 (0.820)		448 (0.858)	443 (0.820)	
	G	170 (0.152)	211 (0.180)	0.069	96 (0.161)	114 (0.180)	0.355	74 (0.142)	97 (0.180)	0.093

[a] Significant P value after Bonferroni's correction for four loci is 0.0125 (0.05/4).

Table 6 Four SNP haplotypes (P1, P2, P8, and P12) frequency among normotensive (NT) and hypertensive (HT) subjects in population 1

		Male subjects				Female subjects			
	Haplotype*	NT (n=298)	HT (n=316)	P^b	Permutation P	NT (n=261)	HT (n=270)	P^b	Permutation P
H1	C-C-A-A	404 (0.677)	413 (0.653)	0.371	0.363	376 (0.720)	356 (0.659)	0.031	0.021*
H2	C-T-G-A	81 (0.136)	96 (0.152)	0.419	0.420	60 (0.116)	75 (0.138)	0.267	0.259
H3	T-T-G-G	63 (0.106)	73 (0.115)	0.584	0.559	52 (0.099)	62 (0.115)	0.405	0.402
H4	C-C-G-G	16 (0.026)	21 (0.033)	0.484	0.470	17 (0.033)	14 (0.027)	0.578	0.617
H5	C-C-A-G	17 (0.029)	18 (0.029)	0.967	0.958	1 (0.002)	19 (0.036)	6.78×10^{-5}*	< 0.001*
H6	C-C-G-A	12 (0.021)	5 (0.008)	0.063	0.066	9 (0.018)	12 (0.021)	0.646	0.682
	Others	3 (0.005)	6 (0.010)			7 (0.013)	2 (0.004)		
Entire distribution				0.722[c]				0.003[c]*	

* Four loci are P1, P2, P8, and P12, and six predominant haplotypes are listed; 'others' category includes minor haplotypes with less than 1% frequency. [b] Significant P value after Bonferroni's correction for major six haplotypes is 0.0083 (0.05/6). [c] P value for the entire distribution with permutation test. * Difference was statistically significant.

of these SNPs were in Hardy–Weinberg equilibrium in the NT group. Table 5 shows the distribution of genotypic and allelic frequencies of the four SNPs in each group. The overall distribution of genotype and allele did not significantly differ between the HT and NT groups for total, male, or female subjects. The P value of χ^2 test for the difference in the genotypic frequency of P8 between male HT and NT groups was 0.044, which was not significant after Bonferroni's correction (multiplied by 4).

We next analyzed the four SNP haplotypes in population 1 (Table 6). Six common haplotypes (H1–H6) covered approximately 99% of the subjects in the HT and NT groups. The frequencies of each haplotype in men did not differ between the HT and NT groups. In women, the frequency of the major C-C-A-A haplotype (H1) of the HT subjects was significantly lower than that of the NT subjects (P=0.031). Multiple logistic regression in women revealed that the association of the H1/H1 diplotype with hypertension remained significant (P=0.006) after adjustment for age, BMI, total cholesterol, high-density lipoprotein cholesterol, and triglyceride. The OR of the H1/H1 diplotype against the others was 0.56 with a 95% CI of 0.37–0.85. The frequency of the minor C-C-A-G haplotype (H5) of the HT subjects was significantly higher than that of the NT subjects ($P=6.78\times10^{-5}$). H5 haplotype was significantly associated with hypertension in a dominant model (P=0.004) after adjustment for the above factors. The OR of the H5/H5+H5/other diplotype against the others was 6.93 with a 95% CI of 1.88–25.5.

To confirm an association of the four SNP haplotypes in women with EH, we genotyped them using the second case–control population (population 2) comprising 925 NT controls and 732 HT patients. Table 2 presents the clinical features of the NT controls and HT patients in population 2. All genotype results of four SNPs in each group were consistent with Hardy–Weinberg equilibrium. Table 7 shows the distribution of genotypic and allelic frequencies of four SNPs in each group of population 2. The overall distribution of genotype and allele of all four SNPs did not significantly differ between the HT and NT groups for total or male subjects. Among women, however, significant differences were observed in the allelic frequencies of P2 (P=0.005) and the genotypic and allelic frequencies of P8 (P=0.005 and 0.003, respectively) between the HT and NT subjects even after Bonferroni's correction (multiplied by 4). P2 was still significantly associated with hypertension in women in both a dominant (P=0.007) and recessive model (P=0.038) after adjustment for age, BMI, total cholesterol, high-density lipoprotein cholesterol, and triglyceride. The OR of T/T+C/T against C/C (dominant model) was 1.22 with a 95% CI of 1.05–1.40, and the OR of T/T against C/T+C/C (recessive model) was 1.94 with a 95% CI of 1.04–3.62. P22 was also significantly associated with hypertension in women in both a dominant model (P=0.011) and recessive model (P=0.002) after adjustment for the above factors. The OR of G/G+A/G against A/A (dominant model) was 1.20 with a 95% CI of 1.04–1.38, and the OR of G/G against A/G+G/G (recessive model) was 1.49 with a 95% CI of 1.16–1.92.

Table 8 shows the frequency of four SNP haplotypes in population 2. Among women, the HT subjects showed a significantly lower frequency of H1 (C-C-A-A) ($P=8.48\times10^{-4}$) and a significantly higher frequency of H2 (C-T-G-A) ($P=6.46\times10^{-4}$) than the NT subjects, whereas in men no significant difference in frequencies of haplotypes between the HT and NT groups was observed. Multiple logistic regression in women revealed that the association of H1 haplotype with hypertension remained significant in both a dominant (P=0.006) and recessive model (P=0.005) after adjustment for age, BMI, total cholesterol, high-density lipoprotein cholesterol, and triglyceride. The OR of the H1/H1+H1/other diplotype against the others (dominant model) was 0.52 with a 95% CI of 0.32–0.83, and the OR of the H1/H1 diplotype against the others (recessive model) was 0.67 with a 95% CI of 0.50–0.89. The H2 haplotype was also significantly associated with hypertension in women in a dominant model (P=0.001) after adjustment for the above factors. The OR of the H2/H2+H2/other diplotype against the others was 1.40 with a 95% CI of 1.18–1.65. In population 2, the frequency of H5

Table 7 Genotype and allele frequencies among normotensive (NT) and hypertensive (HT) subjects in population 2

		Total subjects			Male subjects			Female subjects		
	Genotype	NT ($n=925$)	HT ($n=732$)	P^a	NT ($n=317$)	HT ($n=323$)	P^a	NT ($n=608$)	HT ($n=409$)	P^a
P1	CC	729 (0.792)	573 (0.786)		253 (0.801)	249 (0.778)		476 (0.788)	324 (0.792)	
	CT	181 (0.197)	147 (0.202)		58 (0.184)	66 (0.206)		123 (0.204)	81 (0.198)	
	TT	10 (0.011)	9 (0.012)	0.929	5 (0.016)	5 (0.016)	0.770	5 (0.008)	4 (0.010)	0.949
	Allele									
	C	1639 (0.891)	1293 (0.887)		564 (0.892)	564 (0.881)		1075 (0.890)	729 (0.891)	
	T	201 (0.109)	165 (0.113)	0.721	68 (0.108)	76 (0.119)	0.530	133 (0.110)	89 (0.109)	0.927
P2	CC	550 (0.598)	403 (0.553)		176 (0.555)	182 (0.567)		374 (0.620)	221 (0.542)	
	CT	323 (0.351)	282 (0.387)		118 (0.372)	123 (0.383)		205 (0.340)	159 (0.390)	
	TT	47 (0.051)	44 (0.060)	0.176	23 (0.073)	16 (0.050)	0.488	24 (0.040)	28 (0.069)	0.017
	Allele									
	C	1423 (0.773)	1088 (0.746)		470 (0.741)	487 (0.759)		953 (0.790)	601 (0.737)	
	T	417 (0.227)	370 (0.254)	0.069	164 (0.259)	155 (0.241)	0.477	253 (0.210)	215 (0.263)	0.005*
P8	AA	464 (0.508)	342 (0.472)		146 (0.465)	157 (0.489)		318 (0.530)	185 (0.458)	
	AG	384 (0.420)	316 (0.436)		139 (0.443)	143 (0.445)		245 (0.408)	173 (0.428)	
	GG	66 (0.072)	67 (0.092)	0.189	29 (0.092)	21 (0.065)	0.436	37 (0.062)	46 (0.114)	0.005*
	Allele									
	A	1312 (0.718)	1000 (0.690)		431 (0.688)	457 (0.712)		881 (0.734)	543 (0.672)	
	G	516 (0.282)	450 (0.310)	0.080	197 (0.314)	185 (0.288)	0.321	319 (0.266)	265 (0.328)	0.003*
P12	AA	630 (0.686)	479 (0.659)		214 (0.677)	208 (0.650)		416 (0.691)	271 (0.666)	
	AG	265 (0.289)	220 (0.303)		93 (0.294)	100 (0.313)		172 (0.286)	120 (0.295)	
	GG	23 (0.025)	28 (0.039)	0.213	9 (0.028)	12 (0.038)	0.690	14 (0.023)	16 (0.039)	0.301
	Allele									
	A	1525 (0.831)	1178 (0.810)		521 (0.824)	516 (0.806)		1004 (0.834)	662 (0.813)	
	G	311 (0.169)	276 (0.190)	0.128	111 (0.176)	124 (0.194)	0.405	200 (0.166)	152 (0.187)	0.231

*Significant P value after Bonferroni's correction for four loci is 0.0125 (0.05/4). *Difference was statistically significant.

did not significantly differ between the HT and NT groups for women.

Although trends of frequency changes in the H1 and H2 haplotypes among women in the two independent population were the same, the frequency of H2 showed a significant difference not in population 1 but in population 2. This discrepancy could have been caused by difference in the sample size. When we analyzed the differences in frequencies of each haplotype between the HT and NT groups in combined samples of the two population (Table 9), female HT subjects showed a significantly lower frequency of H1 ($P=8.07\times10^{-5}$) and a significantly higher frequency of H2 ($P=6.07\times10^{-4}$) than the NT subjects. The frequency of H5 of female HT subjects was still significantly higher than that of NT subjects ($P=0.003$). No significant difference in haplotype frequencies between male HT and NT groups was observed.

Variance component estimation of *TNFRSF4*

The variance estimates of the *TNFRSF4* diplotype and the residual in SBP of the control women of population 1 were 5.5 and 79.6, respectively. Therefore, the *TNFRSF4* gene explains 6.5% of the variation of SBP in this group. The values in DBP were 2.8 and 52.1, respectively, with the gene contributing 5.2% of the variation.

Transcriptional effects of polymorphisms in the promoter region

To study transcriptional effects of the polymorphisms, we transfected COS-7 cells and HEK293 cells with promoter constructs containing the haplotypes in the *TNFRSF4* gene (Pr-H1, Pr-H2, and Pr-H5). In COS-7 cells, promoter activity of the Pr-H2 construct was significantly lower than that of the Pr-H1 or Pr-H5 construct (0.89 for Pr-H2/Pr-H1, $P=0.008$ and 0.91 for Pr-H2/Pr-H5, $P=0.026$; Fig. 2a). The same results were observed in HEK293 cells (0.92

Table 8 Four SNP haplotypes (P1, P2, P8, and P12) frequency among normotensive (NT) and hypertensive (HT) subjects in population 2

		Male subjects				Female subjects			
	Haplotype*	NT ($n=303$)	HT ($n=299$)	P^b	Permutation P	NT ($n=584$)	HT ($n=388$)	P^b	Permutation P
H1	C-C-A-A	403 (0.665)	403 (0.674)	0.743	0.714	839 (0.718)	502 (0.647)	8.48×10^{-4}*	<0.001*
H2	C-T-G-A	86 (0.142)	75 (0.125)	0.400	0.388	115 (0.098)	116 (0.149)	6.46×10^{-4}*	0.001*
H3	T-T-G-G	66 (0.109)	71 (0.119)	0.592	0.593	125 (0.107)	84 (0.108)	0.939	0.926
H4	C-C-G-G	30 (0.049)	24 (0.040)	0.429	0.443	47 (0.040)	45 (0.058)	0.067	0.074
H5	C-C-A-G	11 (0.018)	20 (0.034)	0.095	0.113	21 (0.018)	18 (0.023)	0.434	0.451
H6	C-C-G-A	8 (0.014)	4 (0.007)	0.265	0.374	15 (0.013)	8 (0.010)	0.606	0.583
	Others	2 (0.003)	1 (0.002)			6 (0.005)	3 (0.004)		
Entire distribution				0.533[c]				0.026[c,]*	

*Four loci are P1, P2, P8, and P12, and six predominant haplotypes are listed; 'others' category includes minor haplotypes with less than 1% frequency. [b] Significant P value after Bonferroni's correction for major six haplotypes is 0.0083 (0.05/6). [c] P value for the entire distribution with permutation test. * Difference was statistically significant.