Identification of six polymorphisms as novel susceptibility loci for ischemic or hemorrhagic stroke by exome-wide association studies

YOSHIJI YAMADA[1,2], JUN SAKUMA[2-4], ICHIRO TAKEUCHI[2,4,5], YOSHIKI YASUKOCHI[1,2], KIMIHIKO KATO[1,6], MITSUTOSHI OGURI[1,7], TETSUO FUJIMAKI[8], HIDEKI HORIBE[9], MASAAKI MURAMATSU[10], MOTOJI SAWABE[11], YOSHINORI FUJIWARA[12], YU TANIGUCHI[12], SHUICHI OBUCHI[13], HISASHI KAWAI[13], SHOJI SHINKAI[14], SEIJIRO MORI[15], TOMIO ARAI[16] and MASASHI TANAKA[17]

[1]Department of Human Functional Genomics, Advanced Science Research Promotion Center, Mie University, Tsu 514-8507; [2]CREST, Japan Science and Technology Agency, Kawaguchi 332-0012; [3]Computer Science Department, College of Information Science, University of Tsukuba, Tsukuba 305-8573; [4]RIKEN Center for Advanced Intelligence Project, Tokyo 103-0027; [5]Department of Computer Science, Nagoya Institute of Technology, Nagoya 466-8555; [6]Department of Internal Medicine, Meitoh Hospital, Nagoya 465-0025; [7]Department of Cardiology, Kasugai Municipal Hospital, Kasugai 486-8510; [8]Department of Cardiovascular Medicine, Inabe General Hospital, Inabe 511-0428; [9]Department of Cardiovascular Medicine, Gifu Prefectural Tajimi Hospital, Tajimi 507-8522; [10]Department of Molecular Epidemiology, Medical Research Institute, Tokyo Medical and Dental University, Tokyo 101-0062; [11]Section of Molecular Pathology, Graduate School of Health Care Sciences, Tokyo Medical and Dental University, Tokyo 113-8510; [12]Research Team for Social Participation and Community Health, [13]Research Team for Promoting Support System for Home Care, and [14]Research Team for Social Participation and Health Promotion, Tokyo Metropolitan Institute of Gerontology; [15]Center for Promotion of Clinical Investigation, [16]Department of Pathology, and [17]Department of Clinical Laboratory, Tokyo Metropolitan Geriatric Hospital, Tokyo 173-0015, Japan

Received February 3, 2017; Accepted April 20, 2017

DOI: 10.3892/ijmm.2017.2972

Abstract. In this study, we performed exome-wide association studies (EWASs) to identify genetic variants that confer susceptibility to ischemic stroke, intracerebral hemorrhage (ICH), or subarachnoid hemorrhage (SAH). EWAS for ischemic stroke was performed using 1,575 patients with this condition and 9,210 controls, and EWASs for ICH and SAH were performed using 673 patients with ICH, 265 patients with SAH and 9,158 controls. Analyses were performed with Illumina HumanExome-12 DNA Analysis BeadChip or Infinium Exome-24 BeadChip arrays. The relation of allele frequencies for 41,339 or 41,332 single nucleotide polymorphisms (SNPs) that passed quality control to ischemic or hemorrhagic stroke, respectively, was examined with Fisher's exact test. Based on Bonferroni's correction, a P-value of $<1.21 \times 10^{-6}$ was considered statistically significant. EWAS for ischemic stroke revealed that 77 SNPs were significantly associated with this condition. Multivariable logistic regression analysis with adjustment for age, sex and the prevalence of hypertension and diabetes mellitus revealed that 4 of these SNPs [rs3212335 of *GABRB3* (P=0.0036; odds ratio, 1.29), rs147783135 of *TMPRSS7* (P=0.0024; odds ratio, 0.37), rs2292661 of *PDIA5* (P=0.0054; odds ratio, 0.35) and rs191885206 of *CYP4F12* (P=0.0082; odds ratio, 2.60)] were related (P<0.01) to ischemic stroke. EWASs for ICH or SAH revealed that 48 and 12 SNPs, respectively, were significantly associated with these conditions. Multivariable logistic regression analysis with adjustment for age, sex and the prevalence of hypertension revealed that rs138533962 of *STYK1* ($P<1.0 \times 10^{-23}$; odds ratio, 111.3) was significantly ($P<2.60 \times 10^{-4}$) associated with ICH and that rs117564807 of *COL17A1* (P=0.0009; odds ratio, 2.23×10^{-8}) was significantly (P<0.0010) associated with SAH. *GABRB3*, *TMPRSS7*, *PDIA5* and *CYP4F12* may thus be novel susceptibility loci for ischemic stroke, whereas *STYK1* and *COL17A1* may be such loci for ICH and SAH, respectively.

Correspondence to: Professor Yoshiji Yamada, Department of Human Functional Genomics, Advanced Science Research Promotion Center, Mie University, 1577 Kurima-machiya, Tsu, Mie 514-8507, Japan
E-mail: yamada@gene.mie-u.ac.jp

Key words: ischemic stroke, intracerebral hemorrhage, subarachnoid hemorrhage, genetics, exome-wide association studies

Introduction

Stroke is a common and serious condition, with approximately 795,000 individuals having experienced a new or recurrent stroke and 128,978 stroke-related deaths having occurred in the United States in 2013. The prevalence of stroke in the United States was approximately 6.6 million in 2012, with 87% of

these events being ischemic stroke, 10% intracerebral hemorrhage (ICH) and 3% subarachnoid hemorrhage (SAH) (1). Despite recent advances in acute therapy, stroke remains the leading cause of severe disability and the fifth leading cause of mortality (after heart disease, cancer, chronic lower respiratory disease and accidents) in the United States (1). Given that stroke is a life-threatening condition, its prevention is an important goal for reducing its overall burden. The identification of biomarkers for stroke risk is key both for risk prediction and for potential intervention to avert future cerebrovascular events.

Ischemic stroke is a complex multifactorial disorder that is thought to result from an interaction between a person's genetic background and various environmental factors (2). The heritability of ischemic stroke has been estimated to be 40.3% for large-vessel disease, 32.6% for cardioembolic stroke, 16.1% for small-vessel disease and 37.9% for ischemic stroke as a whole (3). Genome-wide association studies (GWASs) in European ancestry populations have identified various genes and loci that confer susceptibility to ischemic stroke (4-11), with one recent large-scale study having identified *HDAC9* and chromosome 1p13.2 (near *TSPAN2*) as susceptibility loci for large-vessel disease, *PITX2* and *ZFHX3* as such loci for cardioembolic stroke, and 12q24 (near *ALDH2*) as a susceptibility locus for small-vessel disease, indicating that ischemic stroke-related loci are subtype specific (12).

ICH accounts for a large proportion of severe or fatal cases of stroke, with its most important risk factors being advanced age and hypertension (13). ICH is usually attributed to hypertensive small-vessel disease, with the most common sites of hemorrhage including the basal ganglia, cerebellum and pons (13,14). In some cases, however, the hemorrhage is lobar in location, such as in the frontal, parietal, temporal or occipital cortex, and such patients often do not have hypertension (14). The occurrence of lobar ICH has been shown to be associated with the ε2 and ε4 alleles of the apolipoprotein E gene (*APOE*) (15,16). This relation of lobar ICH to *APOE* is presumably due to the association of this gene with cerebral amyloid angiopathy (17). The heritability of deep or lobar ICH has been estimated to be 34 and 73%, respectively (18), although genetic factors may influence, not only the development of ICH, but also hypertension (19). A previous meta-analysis of GWASs for ICH in European ancestry populations identified chromosome 12q21.1 (near *TRHDE*) as a susceptibility locus for lobar ICH and 1q22 (near *PMF1-SCL25A44*) as such a locus for nonlobar ICH (20).

SAH is commonly caused by the rupture of an aneurysm in an intracranial artery (21,22). Although the incidence of aneurysmal SAH in the general population is low (~8/100,000 person-years) (23), a young age at onset and poor prognosis result in the loss of productive life-years similar to that for ischemic stroke (24). Given that a family history is an important risk factor for the development of intracranial aneurysm, genetics may play an important role in the development of this condition (25). GWASs have implicated several loci and genes as conferring susceptibility to intracranial aneurysm (26-31), with a meta-analysis of such studies having identified 19 genetic variants related to this condition (32).

Most genetic variants identified by GWASs for ischemic stroke, ICH, or intracranial aneurysm have a minor allele frequency (MAF) of >5% and a small individual effect size. Given that these common variants explain only a fraction of the heritability of ischemic and hemorrhagic stroke, it is expected that low-frequency (MAF of 0.5-5%) or rare (MAF of <0.5%) variants with larger effect sizes contribute to the genetic architecture of these conditions (33). Although several polymorphisms have been found to be significantly associated with ischemic stroke (34-36) or intracranial aneurysm (30) in Japanese individuals, genetic variants, including low-frequency and rare variants, that contribute to genetic susceptibility to ischemic stroke, ICH or SAH in Japanese individuals remain to be identified definitively.

In this study, we performed exome-wide association studies (EWASs) with the use of exome array-based genotyping methods to identify single nucleotide polymorphisms (SNPs) and in particular, low-frequency or rare coding variants with moderate to high effect sizes, that confer susceptibility to ischemic stroke, ICH, or SAH in Japanese individuals. Given that most of the known low-frequency or rare variants were not included in the arrays adopted in previous GWASs for these conditions, we applied Illumina arrays that provide coverage of functional SNPs in entire exons, including such variants.

Materials and methods

Study subjects. For EWAS of ischemic stroke, 1,575 patients with ischemic stroke and 9,210 control individuals were examined, whereas for EWASs of hemorrhagic stroke, 673 patients with ICH, 265 patients with SAH and 9,158 controls were examined. The majority of the control individuals were the same for the studies of ischemic and hemorrhagic stroke. The subjects were recruited from individuals who either visited outpatient clinics of or were admitted to participating hospitals (Gifu Prefectural Tajimi Hospital, Tajimi; Gifu Prefectural General Medical Center, Gifu; Japanese Red Cross Nagoya First Hospital, Nagoya; Inabe General Hospital, Inabe; Hirosaki University Hospital and Hirosaki Stroke and Rehabilitation Center, Hirosaki, Japan) due to various symptoms or for an annual health checkup between 2002 and 2014; were community-dwelling individuals recruited to a population-based cohort study in Inabe between 2010 and 2014 or in Tokyo or Kusatsu between 2011 and 2015; or were cases of autopsy performed at the Tokyo Metropolitan Geriatric Hospital from 1995 to 2012.

The diagnosis of ischemic stroke, ICH, or SAH was based on the occurrence of a new and abrupt focal neurological deficit, with neurological symptoms and signs persisting for >24 h, and it was confirmed by positive findings in computed tomography or magnetic resonance imaging (or both) of the head. The type of stroke was determined according to the Classification of Cerebrovascular Diseases III (37). Given that susceptibility loci for ischemic stroke are subtype-specific (12), we examined subjects with atherothrombotic cerebral infarction (large-vessel disease).

For the study of ischemic stroke, subjects with cardiogenic embolic stroke, lacunar infarction alone, transient ischemic attack, hemorrhagic stroke, cerebrovascular malformations, moyamoya disease, cerebral venous sinus thrombosis, brain tumors, or traumatic cerebrovascular diseases were excluded from enrollment. For the studies of hemorrhagic stroke,

individuals with ischemic stroke, lacunar infarction, transient ischemic attack, intracranial hemorrhage resulting from cerebrovascular malformations, moyamoya disease, cerebral venous sinus thrombosis, brain tumors, traumatic cerebrovascular diseases, or subdural hematoma were excluded. The control individuals had no history of ischemic or hemorrhagic stroke; of aortic, coronary, or peripheral artery disease; or of other thrombotic, embolic or hemorrhagic disorders. Individuals with unruptured intracranial aneurysm were also excluded from the controls. The absence of stroke history was evaluated with a detailed questionnaire and was confirmed by the absence of a history of neurological deficits. Autopsy cases were excluded from the controls.

Body mass index was calculated as follows: body mass index = body weight (kg)/[body height (m)]^2. Blood pressure was measured at least twice with subjects having first rested in the sitting position for >5 min; the measurements were taken by a skilled physician or nurse. Venous blood was collected in the early morning after the subjects had fasted overnight. Plasma glucose level, blood hemoglobin A_{1c}, and serum concentrations of triglycerides, high density lipoprotein (HDL)-cholesterol, low density lipoprotein (LDL)-cholesterol, and creatinine were measured with automated analyzers at the clinical laboratory of each hospital. Estimated glomerular filtration rate (eGFR) was calculated as follows: eGFR (ml/min/1.73 m^2) = 194 x [age (years)]$^{-0.287}$ x [serum creatinine (mg/dl)]$^{-1.094}$ x [0.739 if female]. Chronic kidney disease was defined as an eGFR of <60 ml/min/1.73 m^2.

The study protocol complied with the Declaration of Helsinki and was approved by the Committees on the Ethics of Human Research of Mie University Graduate School of Medicine, Tokyo Metropolitan Institute of Gerontology, Hirosaki University Graduate School of Medicine, and participating hospitals. Written informed consent was obtained from each participant or the families of the deceased subjects.

EWASs. Venous blood (5 or 7 ml) was collected into tubes containing 50 mmol/l ethylenediaminetetraacetic acid (disodium salt), peripheral blood leukocytes were isolated and genomic DNA was extracted from these cells either with a DNA extraction kit (Genomix; Talent, Trieste, Italy; or SMITEST EX-R&D; Medical and Biological Laboratories, Nagoya, Japan) or by standard protocols based on phenol-chloroform extraction and spin columns. In autopsy cases, genomic DNA was extracted from the kidneys. EWASs were performed with the use of a HumanExome-12 v1.1 or v1.2 DNA Analysis BeadChip or Infinium Exome-24 v1.0 BeadChip (Illumina, San Diego, CA, USA), both of which include putative functional exonic variants selected from >12,000 individual exome and whole-genome sequences. The exonic content of ~244,000 SNPs represents diverse populations, including European, African, Chinese and Hispanic individuals (38). SNPs contained in only one of the exome arrays (~3.6% of all SNPs) were excluded from the analysis. We performed quality control (39) as follows: i) genotyping data with a call rate of <97% were discarded, with the mean call rate for the remaining data being 99.9%; ii) sex specification was examined for all samples, and those for which sex phenotype in the clinical records was inconsistent with genetic sex were discarded; iii) duplicated samples and cryptic relatedness were examined by calculation of identity by descent; all pairs of DNA samples showing identity by descent of >0.1875 were inspected, and one sample from each pair was excluded; iv) heterozygosity of SNPs was calculated for all samples, with those showing extremely low or high heterozygosity (>3 standard deviations from the mean) being discarded; v) SNPs in sex chromosomes or in mitochondrial DNA were excluded from the analysis, as were non-polymorphic SNPs or SNPs with a MAF of <0.001; vi) SNPs whose genotype distributions deviated significantly (P<0.001) from Hardy-Weinberg equilibrium in control individuals were discarded; vii) genotype data were examined for population stratification by principal components analysis (40), and population outliers were excluded from the analysis. A total of 41,339 and 41,332 SNPs that passed quality control for the studies of ischemic and hemorrhagic stroke, respectively, were subjected to analysis.

Statistical analysis. For the analysis of characteristics of the study subjects, quantitative data were compared between subjects with ischemic stroke, ICH or SAH and controls with the Mann-Whitney U test, given that variables exhibited skewed distribution (P<0.01 by the Kolmogorov-Smirnov and Lilliefors test). Categorical data were compared between two groups with Fisher's exact test. Allele frequencies were estimated by the gene counting method, and Fisher's exact test was applied to identify departure from Hardy-Weinberg equilibrium. Allele frequencies of SNPs were compared between subjects with ischemic stroke, ICH or SAH and the controls with Fisher's exact test. Given that the Fisher's exact test gives exact P-values, it is appropriate to use this test to examine the relation of low frequency or rare variants to phenotypes. To compensate for multiple comparisons of genotypes with ischemic stroke, ICH or SAH, we applied Bonferroni's correction for statistical significance of association. Given that 41,339 or 41,332 SNPs were analyzed for ischemic and hemorrhagic stroke, respectively, the significance level was set at P<1.21x10^{-6} (0.05/41,339 or 0.05/41,332) for each EWAS. Quantile-quantile plots for P-values of allele frequencies in EWASs for ischemic stroke, ICH and SAH are shown in Fig. 1. The inflation factor (λ) was 1.30 for ischemic stroke, 1.52 for ICH, and 1.72 for SAH. Multivariable logistic regression analysis was performed with ischemic stroke as a dependent variable and independent variables including age, sex (0, woman; 1, man), the prevalence of hypertension and diabetes mellitus (0, no history of these conditions; 1, positive history) and genotype of each SNP. Similar analysis was performed with ICH or SAH as a dependent variable and independent variables including age, sex, the prevalence of hypertension and genotype of each SNP. Genotypes of each SNP were assessed according to dominant [0, AA; 1, AB + BB (A, major allele; B, minor allele)], recessive (0, AA + AB; 1, BB) and additive genetic models, and the P-value, odds ratio and 95% confidence interval were calculated. Additive models comprised additive 1 (0, AA; 1, AB; 0, BB) and additive 2 (0, AA; 0, AB; 1, BB) scenarios, which were analyzed simultaneously with a single statistical model. The relation of genotypes of SNPs to intermediate phenotypes was examined with Fisher's exact test (2x2) or Pearson's Chi-square test (2x3). Bonferroni's correction was also applied to other statistical analysis as indicated. Statistical tests were performed with

Table I. Characteristics of the 10,785 subjects in the exome-wide association study for ischemic stroke.

Characteristic	Ischemic stroke	Controls	P-value
No. of subjects	1,575	9,210	
Age (years)	71.8±12.2	58.8±13.8	<0.0001
Sex (male/female, %)	58.9/41.1	50.4/49.6	<0.0001
Body mass index (kg/m^2)	23.5±3.5	23.1±3.5	<0.0001
Current or former smoker (%)	33.9	37.2	0.0450
Hypertension (%)	79.2	42.4	<0.0001
Systolic blood pressure (mmHg)	147±27	125±20	<0.0001
Diastolic blood pressure (mmHg)	82±16	75±12	<0.0001
Diabetes mellitus (%)	44.0	14.5	<0.0001
Fasting plasma glucose (mmol/l)	7.09±2.88	5.80±1.95	<0.0001
Blood hemoglobin A1c (%)	6.34±1.42	5.70±0.93	<0.0001
Dyslipidemia (%)	58.7	57.6	0.5082
Serum triglycerides (mmol/l)	1.47±0.96	1.38±0.96	<0.0001
Serum HDL-cholesterol (mmol/l)	1.28±0.42	1.62±0.44	<0.0001
Serum LDL-cholesterol (mmol/l)	3.10±0.98	3.13±0.80	0.0208
Chronic kidney disease (%)	36.8	18.8	<0.0001
Serum creatinine (μmol/l)	90.3±115.5	72.9±6804	<0.0001
eGFR (ml m^{-1} 1.73 m^{-2})	67.7±22.7	74.0±18.1	<0.0001
Hyperuricemia (%)	16.9	16.1	0.4794
Serum uric acid (μmol/l)	329±96	323±90	0.1073

Quantitative data are the means ± SD and were compared between subjects with ischemic stroke and controls with the Mann-Whitney U test. Categorical data were compared between the 2 groups with Fisher's exact test. Based on Bonferroni's correction, a P-value of <0.0026 (0.05/19) was considered statistically significant. HDL, high density lipoprotein; LDL, low density lipoprotein.

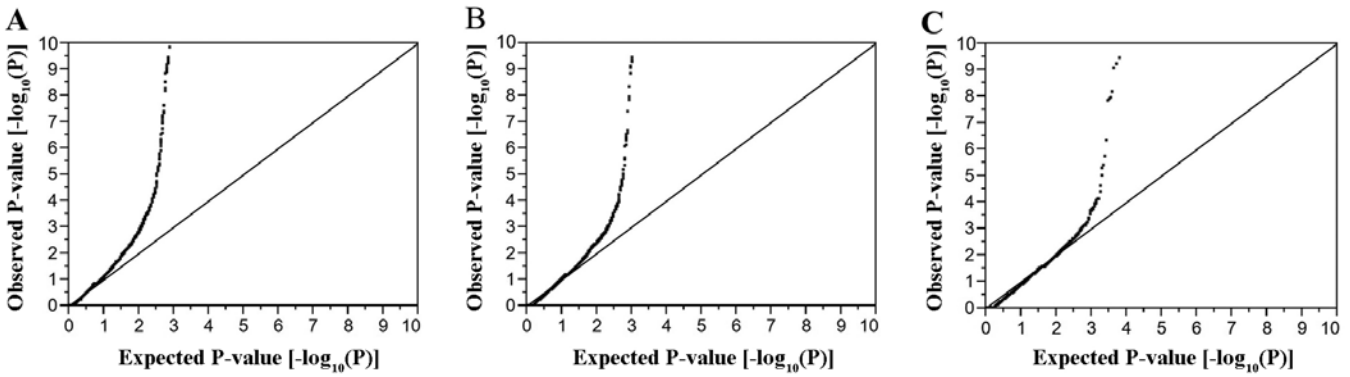

Figure 1. Quantile-quantile plots for P-values of allele frequencies in the EWASs for (A) ischemic stroke, (B) ICH, or (C) SAH. The observed P-values (y-axis) are compared with the expected P-values (x-axis) under the null hypothesis, with the values being presented as -log$_{10}$(P).

JMP Genomics version 6.0 software (SAS Institute, Cary, NC, USA).

Results

EWAS for ischemic stroke. The characteristics of the subjects enrolled in the study are shown in Table I. Age, the frequency of the male gender, body mass index, and the prevalence of hypertension, diabetes mellitus and chronic kidney disease, as well as systolic and diastolic blood pressure, fasting plasma glucose levels, blood glycosylated hemoglobin (hemoglobin A_{1c}) content and serum concentrations of triglycerides and creatinine were greater, whereas the serum concentrations of HDL-cholesterol and estimated glomerular filtration rate were lower, in the subjects with ischemic stroke than in the controls.

We examined the relation of allele frequencies for 41,339 SNPs that passed quality control to ischemic stroke with Fisher's exact test. A Manhattan plot of EWAS for ischemic stroke is shown in Fig. 2. After Bonferroni's correction, 77 SNPs were found to be significantly ($P<1.21\times10^{-6}$) associated with ischemic stroke (Table II). The genotype distributions of these SNPs were in Hardy-Weinberg equilib-

Table II. The 77 SNPs significantly ($P<1.21x10^{-6}$) associated with ischemic stroke in the exome-wide association study.

Gene	dbSNP	Nucleotide (amino acid) substitution[a]	Chromosome: position	MAF (%)	P-value (allele)	Allele OR	95% CI
TTLL5	rs2833270	C/T	21: 3115036	4.1	$5.45x10^{-112}$	1.05	0.87-1.27
	rs605066	C/T	6: 139508529	25.6	$1.93x10^{-89}$	1.02	0.94-1.11
	rs146036604	A/C (L879F)	14: 75783181	1.4	$1.09x10^{-73}$	0.90	0.65-1.25
	rs12615742	T/C	2: 37768584	47.3	$3.80x10^{-62}$	1.04	0.96-1.12
TCEB3B	rs2010834	A/C (F254C)	18: 47034504	24.5	$2.40x10^{-61}$	1.02	0.94-1.11
TCF19	rs3130933	G/A	6: 31164308	5.2	$1.05x10^{-53}$	1.18	1.00-1.39
GABRB3	rs3212335	G/A	15: 26766994	38.9	$8.21x10^{-53}$	1.02	0.94-1.11
NRXN3	rs6574433	G/A	14: 78319816	26.5	$6.14x10^{-49}$	0.98	0.90-1.06
UBASH3A	rs11203203	G/A	21: 42416077	3.9	$7.70x10^{-44}$	1.13	0.93-1.37
TBC1D32	rs79221470	T/A (C505S)	6: 121281639	17.8	$5.10x10^{-37}$	1.01	0.92-1.11
DBT	rs140308307	T/C (N13D)	1: 100249784	0.3	$2.69x10^{-36}$	1.59	0.86-2.94
IL20RA	rs191996643	T/C (I43V)	6: 137017065	0.8	$2.63x10^{-31}$	0.74	0.47-1.18
MMP28	rs117651561	G/A	17: 35766644	13.6	$1.80x10^{-29}$	1.01	0.90-1.12
C15orf57	rs3803354	T/C	15: 40564790	8.9	$3.88x10^{-27}$	1.11	0.98-1.27
TMPRSS7	rs147783135	C/T (R692*)	3: 112081004	0.8	$3.41x10^{-26}$	0.54	0.31-0.94
	rs7752978	A/G	6: 114869897	47.4	$1.29x10^{-24}$	1.00	0.92-1.08
HMGCS2	rs181428774	G/T (L59M)	1: 119764556	0.3	$1.56x10^{-24}$	0.49	0.21-1.14
DPF3	rs757572	T/C	14: 72665321	41.2	$1.28x10^{-21}$	1.07	0.99-1.15
SLC6A4	rs56316081	T/C (I108V)	17: 30221637	0.1	$2.54x10^{-21}$	ND	ND
OSGEPL1	rs75321854	C/T (A98T)	2: 189755490	4.8	$2.31x10^{-20}$	1.10	0.93-1.31
S100A3	rs36022742	C/T (R3K)	1: 153548478	1.6	$4.16x10^{-20}$	0.77	0.55-1.09
ALMS1	rs3820700	G/A (S2576N)	2: 73489683	26.1	$6.70x10^{-20}$	0.98	0.90-1.08
OR2D2	rs1965209	A/G (S148P)	11: 6892059	12.2	$9.24x10^{-20}$	1.11	0.99-1.24
	rs9494145	T/C	6: 135111414	31.8	$9.77x10^{-20}$	0.96	0.88-1.04
AXDND1	rs41267592	C/T (T627M)	1: 179468524	0.3	$1.92x10^{-19}$	0.61	0.28-1.33
SCMH1	rs182666831	A/T (S48C)	1: 41151649	0.1	$5.23x10^{-19}$	0.58	0.18-1.92
ELMO1	rs11984075	A/G	7: 37397251	1.1	$3.61x10^{-18}$	0.78	0.53-1.16
OTOL1	rs12633334	G/A	3: 161452823	45.1	$2.34x10^{-17}$	0.97	0.89-1.04
	rs5026743	T/G	6: 32472187	29.7	$4.27x10^{-17}$	1.12	1.03-1.22
MUC16	rs12459532	G/A (P9422L)	19: 8948505	1.8	$4.29x10^{-17}$	1.42	1.10-1.85
TMPRSS6	rs2235321	G/A	22: 37066886	43.7	$7.08x10^{-17}$	1.09	1.01-1.18
	rs11624336	G/A	14: 96727175	14.9	$2.40x10^{-16}$	0.99	0.88-1.10
KIAA1324	rs1052878	C/T (P922L)	1: 109202996	4.9	$6.66x10^{-16}$	1.05	0.89-1.24
MMP28	rs79742527	G/A	17: 35766662	13.6	$1.04x10^{-15}$	1.00	0.89-1.11
SERINC1	rs11964202	G/A	6: 122444993	0.6	$1.64x10^{-15}$	0.87	0.51-1.47
KANK4	rs2258470	C/T (R822H)	1: 62263166	19.3	$8.84x10^{-15}$	1.13	1.03-1.24
PRR22	rs201451364	C/T (R78H)	19: 5784014	0.1	$1.42x10^{-14}$	1.54	0.57-4.13
PLCB1	rs6140742	C/G	20: 8838465	36.3	$3.52x10^{-14}$	1.01	0.94-1.09
AHNAK2	rs150385420	A/G (S4485P)	14: 104941998	2.6	$6.43x10^{-14}$	0.98	0.77-1.23
FBXO16	rs3735726	C/T (R63Q)	8: 28463730	3.3	$4.02x10^{-13}$	1.14	0.93-1.39
CALR3	rs3810198	T/C	19: 16490383	28.6	$1.45x10^{-12}$	1.03	0.95-1.12
PDGFD	rs10895547	C/T	11: 103937424	41.0	$1.74x10^{-12}$	0.99	0.92-1.07
LOC101928877	rs6844558	T/C	4: 185060282	38.5	$4.86x10^{-11}$	1.02	0.94-1.10
MBD2	rs1145315	T/C	18: 54162585	33.3	$1.47x10^{-10}$	0.95	0.88-1.03
LRTM1	rs182902370	C/T	3: 54918891	0.9	$3.41x10^{-10}$	1.10	0.74-1.62
SEC16B	rs202187751	G/C (D41H)	1: 177967861	0.2	$4.18x10^{-10}$	0.53	0.16-1.72
NEURL4	rs117553236	T/G (M1048L)	17: 7321417	0.3	$5.35x10^{-10}$	0.97	0.41-2.33
PPFIA1	rs546502	G/A (V71I)	11: 70272383	15.9	$6.09x10^{-10}$	1.18	1.08-1.30
SLC18A1	rs1390938	G/A (T136I)	8: 20179202	25.5	$6.98x10^{-10}$	1.06	0.97-1.15
DDOST	rs74526704	C/T (R396Q)	1: 20652655	0.2	$9.25x10^{-10}$	0.86	0.34-2.20

Table II. Continued.

Gene	dbSNP	Nucleotide (amino acid) substitution[a]	Chromosome: position	MAF (%)	P-value (allele)	Allele OR	95% CI
ZNF16	rs139521477	C/T (R669Q)	8: 144930781	0.1	$9.85x10^{-10}$	0.73	0.22-2.43
TTC16	rs142193455	G/A (R450Q)	9: 127726328	1.1	$1.05x10^{-9}$	0.60	0.38-0.93
SEMA4D	rs13295305	C/T (R713K)	9: 89363482	12.7	$1.24x10^{-9}$	1.07	0.96-1.20
TRH	rs13306057	G/A (G3S)	3: 129975823	0.2	$1.47x10^{-9}$	0.44	0.14-1.41
MAST4	rs56337909	G/A (S1863N)	5: 67164767	2.7	$3.19x10^{-9}$	1.01	0.79-1.28
PTPRR	rs10784867	C/T	12: 70740947	48.3	$4.83x10^{-9}$	0.98	0.91-1.06
PREX1	rs6095241	G/A	20: 48692260	41.5	$4.95x10^{-9}$	1.09	1.01-1.18
	rs12402711	G/A	1: 41413409	27.1	$6.95x10^{-9}$	1.00	0.92-1.09
	rs10277516	C/A	7: 22869155	3.3	$2.30x10^{-8}$	1.08	0.88-1.33
ROR2	rs200805854	C/A (E703D)	9: 91724385	0.2	$3.93x10^{-8}$	0.89	0.35-2.27
DUS4L	rs4730250	A/G	7: 107567250	9.4	$4.04x10^{-8}$	0.91	0.80-1.04
	rs11185362	A/G	1: 104018866	24.2	$4.37x10^{-8}$	1.04	0.96-1.14
IL7R	rs3194051	A/G (I356V)	5: 35876172	7.4	$6.06x10^{-8}$	1.01	0.87-1.16
FAM200A	rs75129401	T/C (H35R)	7: 99548304	1.5	$6.08x10^{-8}$	0.64	0.44-0.93
PLXNC1	rs75674989	G/T (R614S)	12: 94226656	3.8	$6.12x10^{-8}$	0.84	0.68-1.03
	rs7131744	G/A	12: 4387114	44.7	$7.92x10^{-8}$	1.09	1.01-1.18
PDIA5	rs2292661	C/T (T391M)	3: 123150263	0.7	$1.03x10^{-7}$	0.47	0.26-0.88
GOLGB1	rs3732407	G/C (S911T)	3: 121697776	3.5	$1.11x10^{-7}$	1.15	0.94-1.41
ARPC1B	rs1045012	G/C (K37N)	7: 99386731	1.7	$1.97x10^{-7}$	0.94	0.70-1.27
ALMS1	rs138921247	G/C (V807L)	2: 73448943	1.8	$2.71x10^{-7}$	0.87	0.65-1.16
TRIM67	rs1998027	A/G	1: 231190340	33.7	$2.80x10^{-7}$	0.98	0.90-1.06
C16orf95	rs3748393	A/C (S26A)	16: 87317167	41.7	$3.24x10^{-7}$	1.06	0.98-1.14
CES5A	rs145397395	C/T (V220M)	16: 55866010	0.9	$4.76x10^{-7}$	0.85	0.76-1.79
ACTR5	rs3752289	C/T (P580L)	20: 38771731	1.4	$5.15x10^{-7}$	0.98	0.71-1.36
	rs3095354	A/G	6: 30868334	26.7	$5.75x10^{-7}$	1.12	1.03-1.22
CYP4F12	rs191885206	T/C (C402R)	19: 15696024	0.3	$6.57x10^{-7}$	2.11	1.23-3.57
CELSR2	rs117684956	G/A (V598M)	1: 109251871	1.3	$9.21x10^{-7}$	1.25	0.93-1.69

Allele frequencies were analyzed with Fisher's exact test. [a]Major allele/minor allele. SNP, single nucleotide polymorphisms; MAF, minor allele frequency; OR, odds ratio; CI, confidence interval; ND, not determined.

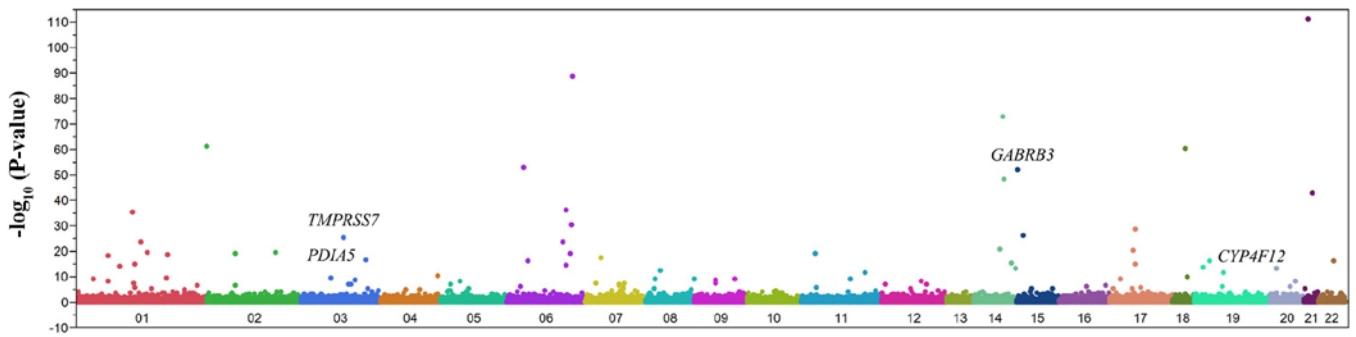

Figure 2. Manhattan plot for P-values of allele frequencies in the EWAS of ischemic stroke. The P-values (y-axis) are plotted as $-\log_{10}(P)$ with respect to the physical chromosomal positions of the corresponding SNPs (x-axis). The 4 SNPs (*GABRB3*, *TMPRSS7*, *PDIA5* and *CYP4F12*) ultimately found to be related to ischemic stroke are indicated.

rium (P>0.001) among both subjects with ischemic stroke and the controls (data not shown).

Multivariable logistic regression analysis of the relation of SNPs to ischemic stroke. The relation of the 77 SNPs iden-

Table III. Relation of SNPs to ischemic stroke as determined by multivariable logistic regression analysis.

SNP		Dominant		Recessive		Additive 1		Additive 2	
		P-value	OR (95% CI)	P-value	OR (95% CI)	P-value	OR (95% CI)	P-value	OR (95% CI)
rs3212335	G/A	0.6428		0.0036	1.29 (1.09-1.53)	0.5896		0.0150	1.26 (1.05-1.52)
rs147783135	C/T (R692*)	0.0024	0.37 (0.16-0.72)	0.4322		0.0029	0.38 (0.17-0.74)	0.4296	
rs2292661	C/T (T391M)	0.0054	0.35 (0.14-0.76)	ND		0.0054	0.35 (0.14-0.76)	ND	
rs191885206	T/C (C402R)	0.0082	2.60 (1.30-4.93)	ND		0.0082	2.60 (1.30-4.93)	ND	

Multivariable logistic regression analysis was performed with adjustment for age, sex and the prevalence of hypertension and diabetes mellitus. Based on Bonferroni's correction, a P-value of $<1.62 \times 10^{-4}$ (0.05/308) was considered statistically significant. SNPs, single nucleotide polymorphisms; OR, odds ratio; CI, confidence interval; ND, not determined.

tified by EWAS to ischemic stroke was examined further by multivariable logistic regression analysis with adjustment for age, sex and the prevalence of hypertension and diabetes mellitus (Table III). Four SNPs, rs3212335 (G/A) of γ-aminobutyric acid type A receptor β3 subunit (*GABRB3*), rs147783135 [C/T (R692*)] of transmembrane protease, serine 7 gene (*TMPRSS7*), rs2292661 [C/T (T391M)] of protein disulfide isomerase family A member 5 gene (*PDIA5*), and rs191885206 [T/C (C402R)] of cytochrome P450 family 4 subfamily F member 12 gene (*CYP4F12*) were related (P<0.01 in at least one genetic model) to ischemic stroke, although no SNP was significantly [$P<1.62 \times 10^{-4}$ (0.05/308)] associated with this condition. The minor A and C alleles of rs3212335 and rs191885206, respectively, were risk factors for ischemic stroke, whereas the minor T alleles of rs147783135 and rs2292661 were protective against this condition.

EWAS for ICH. The characteristics of the subjects for EWAS for ICH are shown in Table IV. Age, the frequency of the male gender, and the prevalence of hypertension and diabetes mellitus were greater in the subjects with ICH than in the controls.

We examined the relation of allele frequencies for 41,332 SNPs to ICH with Fisher's exact test. A Manhattan plot for EWAS for ICH is shown in Fig. 3A. After Bonferroni's correction, 48 SNPs were significantly ($P<1.21 \times 10^{-6}$) associated with ICH (Table V). The genotype distributions of these SNPs were in Hardy-Weinberg equilibrium (P>0.001) both among subjects with ICH and among controls (data not shown).

Multivariable logistic regression analysis of the relation of SNPs to ICH. The relation of the 48 SNPs identified by EWAS to ICH was examined further by multivariable logistic regression analysis with adjustment for age, sex and the prevalence of hypertension. A total of 6 SNPs were related (P<0.05 in at least one genetic model) to ICH (Table VI). Among these SNPs, rs138533962 [G/A (R379C)] of serine/threonine/tyrosine kinase 1 gene (*STYK1*) (dominant and additive 1 models) was significantly [$P<2.60 \times 10^{-4}$ (0.05/192)] associated with ICH, with the minor A allele representing a risk factor for this condition.

EWAS for SAH. The characteristics of the subjects in the EWAS for SAH are shown in Table IV. The prevalence of hypertension was greater, whereas that of dyslipidemia was lower, in the subjects with SAH than in the controls.

We examined the relation of allele frequencies for 41,332 SNPs to SAH with Fisher's exact test. A Manhattan plot for EWAS for SAH is shown in Fig. 3B. After Bonferroni's correction, 12 SNPs were significantly ($P<1.21 \times 10^{-6}$) associated with SAH (Table VII). The genotype distributions of these SNPs were in Hardy-Weinberg equilibrium (P>0.001) both among subjects with SAH and among controls (data not shown).

Multivariable logistic regression analysis of the relation of SNPs to SAH. The relation of the 12 SNPs identified by EWAS to SAH was examined further by multivariable logistic regression analysis with adjustment for age, sex and the prevalence of hypertension. Three SNPs were related (P<0.05) to SAH (Table VIII). Among these SNPs, rs117564807

Table IV. Characteristics of the subjects in the exome-wide association studies for ICH or subarachnoid hemorrhage.

Characteristic	Controls	ICH	P-value	SAH	P-value
No. of subjects	9158	673		265	
Age (years)	58.8±13.9	69.4±13.9	<0.0001	59.4±13.2	0.8515
Sex (male/female, %)	50.4/49.6	62.6/37.4	<0.0001	41.1/58.9	0.0033
Body mass index (kg/m^2)	23.1±3.5	23.1±3.8	0.9986	23.1±3.3	0.9672
Current or former smoker (%)	37.1	33.3	0.1836	31.5	0.1590
Hypertension (%)	42.4	81.4	<0.0001	60.5	<0.0001
Diabetes mellitus (%)	14.5	38.1	<0.0001	19.6	0.0345
Dyslipidemia (%)	57.5	51.5	0.0152	41.2	<0.0001
Chronic kidney disease (%)	18.8	21.2	0.2656	28.0	0.0052
Hyperuricemia (%)	16.0	15.6	0.8915	9.4	0.0060

Quantitative data are the means ± SD and were compared between subjects with ICH or SAH and controls with the Mann-Whitney U test. Categorical data were compared between 2 groups with Fisher's exact test. Based on Bonferroni's correction, a P-value of <0.0028 (0.05/18) was considered statistically significant. ICH, intracerebral hemorrhage; SAH, subarachnoid hemorrhage.

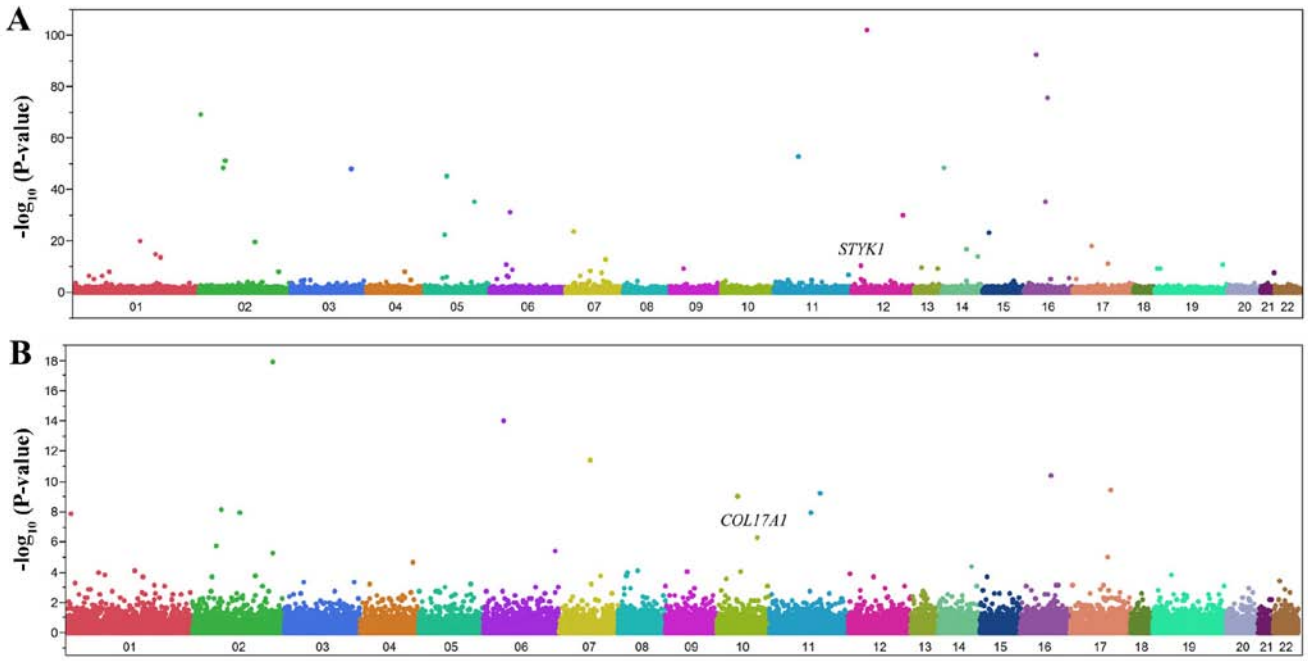

Figure 3. Manhattan plots for P-values of allele frequencies in the EWASs of (A) ICH or (B) SAH. The P-values (y-axis) are plotted as -log$_{10}$(P) with respect to the physical chromosomal positions of the corresponding SNPs (x-axis). The two SNPs ultimately found to be significantly associated with ICH (*STYK1*) or SAH (*COL17A1*) are indicated.

[C/T (D919N)] of collagen type XVII α1 chain gene (*COL17A1*) (dominant and additive 1 models) was significantly [P<0.0010 (0.05/48)] associated with SAH, with the minor T allele being protective against this condition.

Relation of SNPs to intermediate phenotypes. We examined the relation of 6 SNPs (rs3212335 of *GABRB3*, rs147783135 of *TMPRSS7*, rs2292661 of *PDIA5*, rs191885206 of *CYP4F12*, rs138533962 of *STYK1* and rs117564807 of *COL17A1*) to intermediate phenotypes of ischemic or hemorrhagic stroke, including hypertension, diabetes mellitus, hypertriglyceridemia, hypo-HDL-cholesterolemia, hyper-LDL-cholesterolemia, chronic kidney disease, obesity, and hyperuricemia. The rs138533962 SNP of *STYK1* was significantly [P<0.0010 (0.05/48)] associated with the prevalence of hypertension, diabetes mellitus, hypertriglyceridemia and hypo-HDL-cholesterolemia, whereas the other 5 SNPs were not related to any of these intermediate phenotypes (Table IX).

Relation of genes and SNPs identified in the present study to phenotypes previously examined in GWASs. We examined the 6 genes and SNPs identified in the present study to phenotypes previously examined by GWASs available in public databases [GWAS Catalog (http://www.ebi.ac.uk/gwas) and GWAS

Table V. The 48 SNPs significantly ($P<1.21x10^{-6}$) associated with intracerebral hemorrhage in the exome-wide association study.

Gene	dbSNP	Nucleotide (amino acid) substitution[a]	Chromosome: position	MAF (%)	P-value (allele)	Allele OR	95% CI
RECQL	rs146924988	A/G (M1T)	12: 21499569	0.3	$8.57x10^{-103}$	0.93	0.29-3.00
ATF7IP2	rs13335336	A/G	16: 10465406	17.0	$3.91x10^{-93}$	1.15	1.00-1.33
SHCBP1	rs11545690	A/C (M60R)	16: 46618297	5.3	$3.58x10^{-76}$	0.87	0.67-1.13
	rs1405262	T/C	2: 5994808	41.3	$6.36x10^{-70}$	0.98	0.88-1.10
CAT	rs7943316	A/T	11: 34438925	33.7	$1.49x10^{-53}$	1.02	0.91-1.15
EXOC6B	rs1517182	C/A	2: 72508184	1.7	$8.40x10^{-52}$	0.91	0.59-1.41
RNASE10	rs202109789	G/A (G87S)	14: 20510730	0.2	$3.52x10^{-49}$	1.00	0.31-3.25
GKN2	rs146849599	C/T (V130M)	2: 68946388	2.7	$6.04x10^{-49}$	0.82	0.57-1.18
IGSF10	rs78090556	G/A (T981M)	3: 151447039	0.2	$1.49x10^{-48}$	2.57	1.08-6.25
GFM2	rs77099085	G/T (H92N)	5: 74759397	6.4	$5.68x10^{-46}$	0.93	0.74-1.17
ADRB2	rs1042713	G/A (G16R)	5: 148826877	49.1	$4.57x10^{-36}$	1.02	0.92-1.15
SPN	rs3764276	C/T	16: 29661882	37.3	$5.46x10^{-36}$	1.13	1.01-1.26
	rs3135365	T/G	6: 32421478	18.9	$4.03x10^{-32}$	0.99	0.85-1.14
NAA25	rs12231744	C/T (R876K)	12: 112039251	35.1	$1.00x10^{-30}$	1.14	1.02-1.28
DNAH11	rs78763603	G/A	7: 21698150	15.3	$3.69x10^{-24}$	0.99	0.85-1.15
C15orf57	rs3803354	T/C	15: 40564790	8.9	$6.29x10^{-24}$	0.92	0.76-1.12
BDP1	rs34529158	C/A (P1669Q)	5: 71522303	0.5	$3.02x10^{-23}$	0.94	0.44-2.04
FLG	rs2184953	G/A (H2194Y)	1: 152308306	36.2	$9.10x10^{-21}$	0.98	0.88-1.10
	rs563694	T/G	2: 168917561	2.1	$1.75x10^{-20}$	1.24	0.85-1.79
LOC100996813	rs2453589	G/A	17: 19585538	26.0	$8.53x10^{-19}$	1.09	0.96-1.23
DPF3	rs757572	T/C	14: 72665321	41.2	$2.14x10^{-17}$	1.03	0.93-1.16
CCDC18	rs3820059	G/A (S172F)	1: 169421916	7.1	$2.24x10^{-15}$	0.98	0.79-1.22
	rs11624336	G/A	14: 96727175	14.9	$1.56x10^{-14}$	0.99	0.84-1.15
AXDND1	rs41267592	C/T (T627M)	1: 179468524	0.3	$1.65x10^{-14}$	0.62	0.19-1.97
	rs13234712	G/A	7: 119939419	38.8	$1.39x10^{-13}$	0.98	0.88-1.10
MAPT	rs3785879	C/A	17: 45908270	40.8	$8.89x10^{-12}$	0.95	0.85-1.06
NLRP13	rs17711239	T/C (N781S)	19: 55907897	9.3	$1.57x10^{-11}$	1.05	0.87-1.27
AIF1	rs2857697	A/G	6: 31617442	34.1	$1.84x10^{-11}$	1.19	1.06-1.34
STYK1	rs138533962	G/A (R379C)	12: 10620278	2.0	$3.90x10^{-11}$	401.7	100.0-995.0
CCDC169	rs9546897	T/C (K120R)	13: 36254100	34.6	$3.74x10^{-10}$	1.04	0.93-1.16
HMHA1	rs150294461	G/A (G654E)	19: 1080682	1.4	$4.63x10^{-10}$	1.12	0.69-1.79
TJP3	rs1046268	C/T (T898M)	19: 3750617	29.5	$5.56x10^{-10}$	1.00	0.88-1.13
PAX5	rs2297105	A/C	9: 37020625	48.3	$7.69x10^{-10}$	0.98	0.88-1.09
	rs4996815	G/T	13: 105999312	12.6	$8.66x10^{-10}$	1.03	0.88-1.22
HLA-DPB1	rs9277471	A/G	6: 33085905	43.6	$1.55x10^{-9}$	1.01	0.90-1.12
SSC4D	rs10227141	C/T (R505G)	7: 76390273	14.4	$4.91x10^{-9}$	0.94	0.80-1.10
ZCCHC11	rs138145860	A/G (I1270T)	1: 52445800	0.2	$1.04x10^{-8}$	0.93	0.29-2.99
MOGAT1	rs35959734	G/A (A13T)	2: 222671822	1.2	$1.11x10^{-8}$	0.82	0.49-1.39
	rs6534076	C/T	4: 118030971	38.4	$1.38x10^{-8}$	0.96	0.85-1.07
SPATC1L	rs113710653	C/T (E231K)	21: 46161921	1.9	$3.75x10^{-8}$	3.64	2.43-5.44
LRRC17	rs3800939	A/G (K119E)	7: 102934268	14.0	$3.90x10^{-8}$	0.97	0.82-1.13
PATE1	rs2114084	A/G (Q47R)	11: 125747715	36.4	$2.19x10^{-7}$	1.10	0.98-1.23
LY6G6C	rs117894946	G/C (G75A)	6: 31719250	9.5	$2.77x10^{-7}$	1.06	0.88-1.28
VWA5B1	rs139281890	G/A (R92Q)	1: 20312971	0.2	$3.26x10^{-7}$	0.44	0.06-3.23
PABPC4	rs4660293	A/G	1: 39562508	15.2	$3.99x10^{-7}$	1.10	0.95-1.28
ELMO1	rs11984075	A/G	7: 37397251	1.1	$5.49x10^{-7}$	1.04	0.62-1.74
C2	rs511294	A/C	6: 31921092	0.6	$7.29x10^{-7}$	1.20	0.63-2.30
	rs2823962	G/A	21: 16673913	32.8	$8.40x10^{-7}$	0.99	0.88-1.11

Allele frequencies were analyzed with Fisher's exact test. [a]Major allele/minor allele. SNPs, single nucleotide polymorphisms; MAF, minor allele frequency; OR, odds ratio; CI, confidence interval.

Table VI. Relation of SNPs to intracerebral hemorrhage as determined by multivariable logistic regression analysis.

SNP		Dominant		Recessive		Additive 1		Additive 2	
		P-value	OR (95% CI)	P-value	OR (95% CI)	P-value	OR (95% CI)	P-value	OR (95% CI)
rs3764276	C/T	0.0377	1.21 (1.01-1.45)	0.3181		0.0634		0.1001	
rs3820059	G/A (S172F)	0.6828		0.0091	$2.6x10^{-8}$ (0-0.55)	0.4378		0.0095	$2.6x10^{-8}$ (0-$5.5x10^{-7}$)
rs2857697	A/G	0.0444	1.20 (1.00-1.43)	0.0176	1.36 (1.06-1.74)	0.1876		0.0073	1.46 (1.11-1.90)
rs138533962	G/A (R379C)	**$<1.0x10^{-23}$**	111.3 (33.0-694.6)	ND		**$<1.0x10^{-23}$**	111.3 (33.0-694.6)	ND	
rs113710653	C/T (E231K)	0.0020	2.35 (1.39-3.78)	0.6760		0.0018	2.38 (1.41-3.83)	0.6804	
rs3800939	A/G (K119E)	0.8551		0.0414	0.44 (0.15-0.97)	0.5338		0.0466	0.44 (0.16-0.99)

Multivariable logistic regression analysis was performed with adjustment for age, sex, and the prevalence of hypertension. Based on Bonferroni's correction, P-values of $<2.60x10^{-4}$ (0.05/192) were considered statistically significant and are shown in bold. SNPs, single nucleotide polymorphisms; OR, odds ratio; CI, confidence interval; ND, not determined.

Central (http://www.gwascentral.org/browser)]. None of these genes or SNPs was found to be associated with ischemic or hemorrhagic stroke in previous GWASs (data not shown).

Discussion

The pathogenesis of ischemic stroke is predominantly attributed to atherothrombosis of the cerebral arteries, with its treatable risk factors including hypertension, diabetes mellitus and chronic kidney disease (1). In addition to these conventional risk factors, genetic variants are important in the development of ischemic stroke (2). ICH is typically a manifestation of underlying small-vessel disease. Long-standing hypertension can thus lead to hypertensive vasculopathy and consequent microscopic degenerative changes in the walls of small to medium penetrating vessels (13). Cerebral amyloid angiopathy is characterized by the deposition of the β-amyloid peptide in the walls of small leptomeningeal and cortical vessels, which can also ultimately result in degenerative changes in the vessel wall including loss of smooth muscle cells, wall thickening, luminal narrowing, microaneurysm formation and microhemorrhages (14-17). Spontaneous SAH usually results from a ruptured intracranial aneurysm. Blood in the subarachnoid space gives rise to chemical meningitis that commonly increases intracranial pressure for days or a few weeks. Secondary vasospasm may then cause focal brain ischemia. Brain edema accelerates vasospasm and subsequent infarction (21,22). Given the serious nature of both ischemic and hemorrhagic stroke, prediction of the risk for these conditions on the basis of genetic variants would be of benefit for decision-making with regard to how aggressively to target the clinical risk factors that are currently amenable to treatment.

In this study, we demonstrated that rs3212335 (G/A) of *GABRB3*, rs147783135 [C/T (R692*)] of *TMPRSS7*, rs2292661 [C/T (T391M)] of *PDIA5* and rs191885206 [T/C (C402R)] of *CYP4F12* were related to ischemic stroke in Japanese individuals. The minor A allele of rs3212335 and the C allele of rs191885206 were risk factors for ischemic stroke, whereas the minor T alleles of rs147783135 and rs2292661 were protective against this condition. We also found that rs138533962 [G/A (R379C)] of *STYK1* was significantly associated with ICH, with the minor A allele representing a risk factor for ICH, and that rs117564807 [C/T (D919N)] of *COL17A1* was significantly associated with SAH, with the minor T allele being protective against this condition.

SNPs associated with ischemic stroke. The *GABRB3* gene is located at chromosomal region 15q12 (NCBI Gene, https://www.ncbi.nlm.nih.gov/gene) and is highly expressed in the brain (The Human Protein Atlas, http://www.proteinatlas.org). The GABRB3 protein is a component of a multisubunit Cl^- channel that serves as a receptor for γ-aminobutyric acid (GABA), a major inhibitory neurotransmitter of the mammalian nervous system (41). Mutations of *GABRB3* have been associated with several disorders, including Angelman syndrome (42), Prader-Willi syndrome (43), non-syndromic orofacial clefts (44), epilepsy (45) and autism (46). In this study, we demonstrated that rs3212335 (G/A) of *GABRB3* was related to ischemic stroke, with the minor A allele being a risk factor

Table VII. The 12 SNPs significantly ($P<1.21x10^{-6}$) associated with subarachnoid hemorrhage in the exome-wide association study.

Gene	dbSNP	Nucleotide (amino acid) substitution[a]	Chromosome: position	MAF (%)	P-value (allele)	Allele OR	95% CI
CDL6	rs2972146	A/C	2: 226235982	8.7	$1.28x10^{-18}$	0.98	0.72-1.33
	rs3135365	T/G	6: 32421478	18.9	$9.71x10^{-15}$	0.86	0.68-1.09
	rs2282978	T/C	7: 92635096	10.6	$3.91x1^{0-12}$	1.23	0.95-1.59
	rs2639889	A/G	16: 61089243	32.3	$4.23x10^{-11}$	0.88	0.72-1.06
ANKFN1	rs12449568	T/C	17: 56352794	42.0	$3.64x10^{-10}$	0.92	0.77-1.09
CHRDL2	rs79893604	G/A (P395L)	11: 74697229	0.8	$5.70x10^{-10}$	2.45	1.23-4.76
CTNNA3	rs12256826	C/T	10: 66214832	13.6	$8.96x10^{-10}$	0.83	0.63-1.09
SLC4A5	rs10177833	A/C	2: 74230591	46.8	$7.03x10^{-9}$	0.99	0.84-1.18
TEX41	rs2381683	A/G	2: 144981989	2.7	$1.15x10^{-8}$	1.14	0.67-1.91
PYGM	rs589691	T/C	11: 64757744	42.2	$1.25x10^{-8}$	1.06	0.89-1.27
	rs34429154	A/C	1: 4303675	40.2	$1.47x10^{-8}$	1.04	0.87-1.24
COL17A1	rs117564807	C/T (D919N)	10: 104040357	1.1	$4.83x10^{-7}$	0.00	ND

Allele frequencies were analyzed with Fisher's exact test. [a]Major allele/minor allele. SNPs, single nucleotide polymorphisms; MAF, minor allele frequency; OR, odds ratio; CI, confidence interval; ND, not determined.

for this condition, although the molecular mechanism underlying this association remains unclear.

TMPRSS7 is located at chromosomal region 3q13.2 (NCBI Gene) and is highly expressed in the testis (The Human Protein Atlas). The TMPRSS7 protein belongs to the type II transmembrane serine protease (TTSP) family, the 17 human members of which degrade components of the extracellular matrix (47) and play physiological and pathological roles in digestion, cardiac function, blood pressure regulation, hearing, iron metabolism, and epithelial homeostasis (48,49). They have also been implicated in tumor growth, invasion and metastasis (48,49), and the genetic variants of *TMPRSS7* have been associated with the risk for and prognosis of breast cancer (50). In this study, we demonstrated that rs147783135 [C/T (R692*)] of *TMPRSS7* was related to ischemic stroke, with the minor T allele being protective against this condition. Given the potential roles of *TMPRSS7* in tumor growth and blood pressure regulation (48,49) the association of this gene with ischemic stroke may reflect an effect on atherosclerosis or blood pressure.

PDIA5 is located at chromosomal region 3q21.1 (NCBI Gene) and is expressed widely including in the brain and vascular smooth muscle (The Human Protein Atlas). *PDIA5* encodes one of the protein disulfide isomerases that catalyze protein folding and thiol-disulfide interchange reactions in the endoplasmic reticulum (ER). The PDIA5 protein contains an NH_2-terminal ER signal sequence, three catalytically active thioredoxin domains, a thioredoxin-like domain, and a COOH-terminal ER retention sequence. The thioredoxin-like domain is the primary binding site for the major ER chaperone calreticulin (51,52). Recent GWASs indicated that a SNP in *PDIA5* was associated with plasma fibrinogen concentration (53) and platelet count (54). In this study, we demonstrated that rs2292661 [C/T (T391M)] of *PDIA5* was related to ischemic stroke, with the minor T allele being protective against this condition. Given its potential role as a determinant of fibrinogen concentration and platelet count, both of which are important in the development of atherosclerotic thrombosis, the association of *PDIA5* with ischemic stroke may reflect an effect of this gene on arterial thrombosis.

CYP4F12 is located at chromosomal region 19p13.12 (NCBI Gene) and is expressed in various tissues and organs including the brain and vascular smooth muscle (The Human Protein Atlas). *CYP4F12* encodes a member of the cytochrome P450 superfamily of monooxygenases that catalyze many reactions including those related to drug metabolism, as well as to the synthesis of cholesterol, steroids and other lipids. The CYP4F12 protein is likely localized to the ER (55,56). Given that human CYP4F enzymes play a role in the metabolism of endogenous compounds such as inflammatory mediators, they likely contribute to regulation of inflammatory processes (57). Arachidonic acid and prostaglandin H_2 (PGH_2) have been found to serve as substrates of CYP4F12, which also metabolizes to a lesser extent PGE_2, $PGF_{2\alpha}$ and leukotriene B4 (57,58). In this study, we demonstrated that rs191885206 [T/C (C402R)] of *CYP4F12* was related to ischemic stroke, with the minor C allele representing a risk factor for this condition. The association of *CYP4F12* with ischemic stroke may reflect an effect of this gene on vascular inflammation.

A SNP associated with ICH. The *STYK1* gene is located at chromosomal region 12p13.2 (NCBI Gene) and is expressed in various tissues and organs, including the brain (The Human Protein Atlas). *STYK1* plays important roles in diverse cellular and developmental processes, including cell proliferation, differentiation and survival (59,60). The upregulation of *STYK1*

Table VIII. Relation of SNPs to subarachnoid hemorrhage as determined by multivariable logistic regression analysis.

SNP		Dominant		Recessive		Additive 1		Additive 2	
		P-value	OR (95% CI)	P-value	OR (95% CI)	P-value	OR (95% CI)	P-value	OR (95% CI)
rs2639889	A/G	0.0203	0.74 (0.58-0.96)	0.9267		0.0129	0.71 (0.55-0.93)	0.5252	
rs12256826	C/T	0.5019		0.0024	$2.2x10^{-8}$ ($0-6.1x10^{-4}$)	0.8687		0.0024	$2.2x10^{-8}$ ($0-1.1x10^{-4}$)
rs117564807	C/T (D919N)	**0.0009**	$2.2x10^{-8}$ ($0-3.7x10^{-4}$)	0.6615		**0.0009**	$2.2x10^{-8}$ ($0-1.1x10^{-4}$)	0.6583	

Multivariable logistic regression analysis was performed with adjustment for age, sex, and the prevalence of hypertension. Based on Bonferroni's correction, P-values of <0.0010 (0.05/48) were considered statistically significant and are shown in bold. SNPs, single nucleotide polymorphisms; OR, odds ratio; CI, confidence interval.

expression has been detected in many types of tumor, including breast cancer (61), lung cancer (62), ovarian cancer (63), prostate cancer (64), colorectal cancer (65) and hepatocellular carcinoma (66). The STYK1 protein has been shown to promote cell transformation, tumorigenesis and metastasis by activating the phosphoinositide 3-kinase (PI3K)-AKT signaling pathway and inactivating glycogen synthase kinase 3β signaling (67). It has also been found to enhance angiogenesis and to change the morphology of blood vessels during tumor growth (68). In this study, we demonstrated that rs138533962 [G/A (R379C)] of *STYK1* was significantly associated with ICH, with the minor A allele representing a risk factor for this condition. The relation of *STYK1* to ICH may be attributable to an effect of this gene on the remodeling of blood vessels in the brain, although the molecular mechanism underlying this association remains to be determined.

A SNP associated with SAH. The *COL17A1* gene is located at chromosomal region 10q25.1 (NCBI Gene) and is highly expressed in skin (The Human Protein Atlas). Collagen XVII is a homotrimer of three 180-kDa α1 chains, each of which possesses a long intracellular NH_2-terminal domain, a short transmembrane region, and an extracellular COOH-terminus. It is a structural component of hemidesmosomes, which mediate the adhesion of epidermal keratinocytes and certain other types of epithelial cell to the underlying basement membrane (69,70). Mutations in *COL17A1* that result in the generation of premature stop codons are associated with non-Herlitz junctional epidermolysis bullosa (69,70), which is characterized by generalized blistering of the skin and mucous membranes together with atrophic scarring and nail dystrophy. The majority of patients with such mutations lack type XVII collagen mRNA as a result of nonsense-mediated mRNA decay (71). In addition, an autoimmune response to collagen XVII is responsible for epidermal impairment in individuals with bullous pemphigoid (72). A population-based study revealed an association of bullous pemphigoid with neurological disorders, in particular stroke, for which the odds ratio was 3.3 (73). This association is likely due to the fact that the inflammatory state present in individuals with bullous pemphigoid is not confined to the skin but also involves the vascular endothelium (74,75). It has been suggested that intracranial aneurysm is a chronic inflammatory disease at bifurcation sites of cerebral arterial walls (76). Indeed, many inflammatory cells have been detected in specimens of surgically dissected intracranial vessel walls affected by aneurysm (77). Chronic vascular inflammation, acting together with the structural properties of the cerebral arterial wall and hemodynamic factors, may therefore accelerate the development of intracranial aneurysm and subsequent aneurysm rupture. We have now shown that rs117564807 [C/T (D919N)] of *COL17A1* was significantly associated with SAH, with the minor T allele being protective against this condition. The genotype distribution of rs117564807 was 265/265 (100%) CC in subjects with SAH and 8958/9158 (97.82%) CC, 197/9158 (2.15%) CT, and 3/9158 (0.03%) TT in control individuals, giving an allele odds ratio of 0. The association of *COL17A1* with SAH may be attributable to the effect of this gene on cerebrovascular inflammation, although the molecular mechanism remains to be determined.

Table IX. Relation of SNPs to intermediate phenotypes of ischemic or hemorrhagic stroke.

SNP		Hypertension	DM	Hyper-TG	Hypo-HDL	Hyper-LDL	CKD	Obesity	HU
Related to ischemic stroke									
rs3212335	G/A	0.1945	0.7378	0.3226	0.3391	0.2901	0.8774	0.1276	0.0084
rs147783135	C/T (R692*)	0.5725	0.0957	0.6348	0.9054	0.4731	0.6812	0.5797	0.8922
rs2292661	C/T (T391M)	0.2069	0.2699	0.9279	0.8970	0.4782	0.2028	0.7069	0.3671
rs191885206	T/C (C402R)	1.0000	0.8703	0.1862	0.4469	0.5154	0.8764	1.0000	0.1129
Related to ICH									
rs138533962	G/A (R379C)	**4.55×10^{-13}**	**7.12×10^{-5}**	**0.0002**	**1.81×10^{-13}**	0.0765	1.0000	0.5188	1.0000
Related to SAH									
rs117564807	C/T (D919N)	0.7052	0.5715	0.2072	0.2934	0.1943	0.1056	0.5019	0.1484

Data are shown as P-values. The prevalence of each condition was compared among genotypes with Fisher's exact test (2x2) or Pearson's Chi-square test (2x3). Based on Bonferroni's correction, P-values of <0.0010 (0.05/48) were considered statistically significant and are shown in bold. SNPs, single nucleotide polymorphisms; DM, diabetes mellitus; hyper-TG, hypertriglyceridemia; hypo-HDL, hypo-HDL-cholesterolemia; hyper-LDL, hyper-LDL-cholesterolemia; CKD, chronic kidney disease; HU, hyperuricemia; ICH, intracerebral hemorrhage; SAH, subarachnoid hemorrhage.

General considerations. In previous meta-analyses of GWASs for ischemic stroke, the MAF of most SNPs ranged from 6 to 49% and the odds ratio from 0.8 to 2.0 (4-12,34-36). In our study, we identified four SNPs related to ischemic stroke, with the allele odds ratio (MAF, %) of rs3212335 of *GABRB3*, rs147783135 of *TMPRSS7*, rs2292661 of *PDIA5*, and rs191885206 of *CYP4F12* being 1.02 (38.9%), 0.54 (0.8%), 0.47 (0.7%), and 2.11 (0.3%), respectively. Whereas rs3212335 of *GABRB3* was thus a common variant with a small effect size, the other three SNPs were low-frequency or rare variants with a moderate effect size.

A meta-analysis of GWASs identified rs11179580 at 12q21.1 and rs156197380 at 1q22 as susceptibility loci for lobar and nonlobar ICH, respectively, in European ancestry populations (20). The odds ratio (MAF, %) was 1.56 (24%) for rs11179580 and 1.44 (32%) for rs156197380. Previous GWASs (26-31) and a meta-analysis of GWASs (32) also identified SNPs associated with intracranial aneurysm. The MAFs of these SNPs ranged from 15 to 48% and the odds ratios from 0.47 to 2.22. Another recent GWAS of intracranial aneurysm identified low-frequency genetic variants with MAFs of 3 to 5% and odds ratios of ~2.0 (78). We have now identified two SNPs significantly associated with hemorrhagic stroke, with the odds ratio (MAF, %) of rs138533962 of *STYK1* and rs117564807 of *COL17A1* being 111.3 (2.0%) and 2.23×10^{-8} (1.1%), respectively. Both of these SNPs were thus low-frequency variants with a large effect size.

There are several limitations to the present study: i) The inflation factors for ischemic stroke, ICH, and SAH were relatively high, which may be attributable to the relatively small numbers of subjects with ischemic and hemorrhagic stroke. Given that our results were not replicated, they will require validation in other independent subject panels or in other ethnic groups; ii) it is possible that rs3212335 of *GABRB3*, rs147783135 of *TMPRSS7*, rs2292661 of *PDIA5*, rs191885206 of *CYP4F12*, rs138533962 of *STYK1* or rs117564807 of *COL17A1* is in linkage disequilibrium with other polymorphisms in the same gene or in other nearby genes that are actually responsible for the development of ischemic or hemorrhagic stroke; iii) the functional relevance of these SNPs to the pathogenesis of ischemic stroke, ICH, or SAH remains to be elucidated.

In conclusion, our results suggest that rs3212335 (G/A) of *GABRB3*, rs147783135 [C/T (R692*)] of *TMPRSS7*, rs2292661 [C/T (T391M)] of *PDIA5*, and rs191885206 [T/C (C402R)] of *CYP4F12* may be novel susceptibility loci for ischemic stroke, whereas rs138533962 [G/A (R379C)] of *STYK1* and rs117564807 [C/T (D919N)] of *COL17A1* may be such loci for ICH and SAH, respectively, in Japanese individuals. Determination of genotypes for these SNPs may prove informative for assessment of the genetic risk for ischemic or hemorrhagic stroke in Japanese individuals.

Acknowledgements

This study was supported by the CREST (grant number, JPMJCR1302), the Japan Science and Technology Agency (to Y.Y., J.S. and I.T.) and by the Japan Society for the Promotion of Science KAKENHI grants JP15H04772 (to Y.Y.), JP25242062 (to M.T.) and JP16H01872 (to M.T.).

References

1. Mozaffarian D, Benjamin EJ, Go AS, Arnett DK, Blaha MJ, Cushman M, Das SR, de Ferranti S, Després JP, Fullerton HJ, *et al*; Writing Group Members; American Heart Association Statistics Committee; Stroke Statistics Subcommittee: Heart Disease and Stroke Statistics-2016 Update: A Report From the American Heart Association. Circulation 133: e38-e360, 2016.
2. Falcone GJ, Malik R, Dichgans M and Rosand J: Current concepts and clinical applications of stroke genetics. Lancet Neurol 13: 405-418, 2014.
3. Bevan S, Traylor M, Adib-Samii P, Malik R, Paul NL, Jackson C, Farrall M, Rothwell PM, Sudlow C, Dichgans M, *et al*: Genetic heritability of ischemic stroke and the contribution of previously reported candidate gene and genomewide associations. Stroke 43: 3161-3167, 2012.
4. Gretarsdottir S, Thorleifsson G, Manolescu A, Styrkarsdottir U, Helgadottir A, Gschwendtner A, Kostulas K, Kuhlenbäumer G, Bevan S, Jonsdottir T, *et al*: Risk variants for atrial fibrillation on chromosome 4q25 associate with ischemic stroke. Ann Neurol 64: 402-409, 2008.

5. Gudbjartsson DF, Holm H, Gretarsdottir S, Thorleifsson G, Walters GB, Thorgeirsson G, Gulcher J, Mathiesen EB, Njølstad I, Nyrnes A, *et al*: A sequence variant in ZFHX3 on 16q22 associates with atrial fibrillation and ischemic stroke. Nat Genet 41: 876-878, 2009.
6. Ikram MA, Seshadri S, Bis JC, Fornage M, DeStefano AL, Aulchenko YS, Debette S, Lumley T, Folsom AR, van den Herik EG, *et al*: Genomewide association studies of stroke. N Engl J Med 360: 1718-1728, 2009.
7. Holliday EG, Maguire JM, Evans TJ, Koblar SA, Jannes J, Sturm JW, Hankey GJ, Baker R, Golledge J, Parsons MW, *et al*; Australian Stroke Genetics Collaborative; International Stroke Genetics Consortium; Wellcome Trust Case Control Consortium 2: Common variants at 6p21.1 are associated with large artery atherosclerotic stroke. Nat Genet 44: 1147-1151, 2012.
8. Bellenguez C, Bevan S, Gschwendtner A, Spencer CC, Burgess AI, Pirinen M, Jackson CA, Traylor M, Strange A, Su Z, *et al*; International Stroke Genetics Consortium (ISGC); Wellcome Trust Case Control Consortium 2 (WTCCC2): Genome-wide association study identifies a variant in HDAC9 associated with large vessel ischemic stroke. Nat Genet 44: 328-333, 2012.
9. Traylor M, Farrall M, Holliday EG, Sudlow C, Hopewell JC, Cheng YC, Fornage M, Ikram MA, Malik R, Bevan S, *et al*; Australian Stroke Genetics Collaborative, Wellcome Trust Case Control Consortium 2 (WTCCC2); International Stroke Genetics Consortium: Genetic risk factors for ischaemic stroke and its subtypes (the METASTROKE collaboration): A meta-analysis of genome-wide association studies. Lancet Neurol 11: 951-962, 2012.
10. Kilarski LL, Achterberg S, Devan WJ, Traylor M, Malik R, Lindgren A, Pare G, Sharma P, Slowik A, Thijs V, *et al*; GARNET Collaborative Research Group, Wellcome Trust Case Control Consortium 2, Australian Stroke Genetic Collaborative, the METASTROKE Consortium, and the International Stroke Genetics Consortium: Meta-analysis in more than 17,900 cases of ischemic stroke reveals a novel association at 12q24.12. Neurology 83: 678-685, 2014.
11. Neurology Working Group of the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium; Stroke Genetics Network (SiGN); International Stroke Genetics Consortium (ISGC): Identification of additional risk loci for stroke and small vessel disease: a meta-analysis of genome-wide association studies. Lancet Neurol 15: 695-707, 2016.
12. NINDS Stroke Genetics Network: International Stroke Genetics Consortium (ISGC): Loci associated with ischaemic stroke and its subtypes (SiGN): A genome-wide association study. Lancet Neurol 15: 174-184, 2016.
13. Qureshi AI, Tuhrim S, Broderick JP, Batjer HH, Hondo H and Hanley DF: Spontaneous intracerebral hemorrhage. N Engl J Med 344: 1450-1460, 2001.
14. Sacco RL: Lobar intracerebral hemorrhage. N Engl J Med 342: 276-279, 2000.
15. O'Donnell HC, Rosand J, Knudsen KA, Furie KL, Segal AZ, Chiu RI, Ikeda D and Greenberg SM: Apolipoprotein E genotype and the risk of recurrent lobar intracerebral hemorrhage. N Engl J Med 342: 240-245, 2000.
16. Woo D, Kaushal R, Chakraborty R, Woo J, Haverbusch M, Sekar P, Kissela B, Pancioli A, Jauch E, Kleindorfer D, *et al*: Association of apolipoprotein E4 and haplotypes of the apolipoprotein E gene with lobar intracerebral hemorrhage. Stroke 36: 1874-1879, 2005.
17. Greenberg SM, Vonsattel JP, Segal AZ, Chiu RI, Clatworthy AE, Liao A, Hyman BT and Rebeck GW: Association of apolipoprotein E epsilon2 and vasculopathy in cerebral amyloid angiopathy. Neurology 50: 961-965, 1998.
18. Devan WJ, Falcone GJ, Anderson CD, Jagiella JM, Schmidt H, Hansen BM, Jimenez-Conde J, Giralt-Steinhauer E, Cuadrado-Godia E, Soriano C, *et al*; International Stroke Genetics Consortium: Heritability estimates identify a substantial genetic contribution to risk and outcome of intracerebral hemorrhage. Stroke 44: 1578-1583, 2013.
19. Falcone GJ, Biffi A, Devan WJ, Jagiella JM, Schmidt H, Kissela B, Hansen BM, Jimenez-Conde J, Giralt-Steinhauer E, Elosua R, *et al*; International Stroke Genetics Consortium: Burden of risk alleles for hypertension increases risk of intracerebral hemorrhage. Stroke 43: 2877-2883, 2012.
20. Woo D, Falcone GJ, Devan WJ, Brown WM, Biffi A, Howard TD, Anderson CD, Brouwers HB, Valant V, Battey TW, *et al*; International Stroke Genetics Consortium: Meta-analysis of genome-wide association studies identifies 1q22 as a susceptibility locus for intracerebral hemorrhage. Am J Hum Genet 94: 511-521, 2014.
21. Macdonald RL and Schweizer TA: Spontaneous subarachnoid haemorrhage. Lancet 389: 655-666, 2017.
22. van Gijn J, Kerr RS and Rinkel GJ: Subarachnoid haemorrhage. Lancet 369: 306-318, 2007.
23. Linn FHH, Rinkel GJE, Algra A and van Gijn J: Incidence of subarachnoid hemorrhage: role of region, year, and rate of computed tomography: A meta-analysis. Stroke 27: 625-629, 1996.
24. Johnston SC, Selvin S and Gress DR: The burden, trends, and demographics of mortality from subarachnoid hemorrhage. Neurology 50: 1413-1418, 1998.
25. Tromp G, Weinsheimer S, Ronkainen A and Kuivaniemi H: Molecular basis and genetic predisposition to intracranial aneurysm. Ann Med 46: 597-606, 2014.
26. Bilguvar K, Yasuno K, Niemelä M, Ruigrok YM, von Und Zu Fraunberg M, van Duijn CM, van den Berg LH, Mane S, Mason CE, Choi M, *et al*: Susceptibility loci for intracranial aneurysm in European and Japanese populations. Nat Genet 40: 1472-1477, 2008.
27. Yasuno K, Bilguvar K, Bijlenga P, Low SK, Krischek B, Auburger G, Simon M, Krex D, Arlier Z, Nayak N, *et al*: Genome-wide association study of intracranial aneurysm identifies three new risk loci. Nat Genet 42: 420-425, 2010.
28. Yasuno K, Bakırcıoğlu M, Low SK, Bilgüvar K, Gaál E, Ruigrok YM, Niemelä M, Hata A, Bijlenga P, Kasuya H, *et al*: Common variant near the endothelin receptor type A (EDNRA) gene is associated with intracranial aneurysm risk. Proc Natl Acad Sci USA 108: 19707-19712, 2011.
29. Foroud T, Koller DL, Lai D, Sauerbeck L, Anderson C, Ko N, Deka R, Mosley TH, Fornage M, Woo D, *et al*; FIA Study Investigators: Genome-wide association study of intracranial aneurysms confirms role of Anril and SOX17 in disease risk. Stroke 43: 2846-2852, 2012.
30. Low SK, Takahashi A, Cha PC, Zembutsu H, Kamatani N, Kubo M and Nakamura Y: Genome-wide association study for intracranial aneurysm in the Japanese population identifies three candidate susceptible loci and a functional genetic variant at EDNRA. Hum Mol Genet 21: 2102-2110, 2012.
31. Foroud T, Lai D, Koller D, Van't Hof F, Kurki MI, Anderson CS, Brown RD Jr, Connolly ES, Eriksson JG, Flaherty M, *et al*; Familial Intracranial Aneurysm Study Investigators: Genome-wide association study of intracranial aneurysm identifies a new association on chromosome 7. Stroke 45: 3194-3199, 2014.
32. Alg VS, Sofat R, Houlden H and Werring DJ: Genetic risk factors for intracranial aneurysms: A meta-analysis in more than 116,000 individuals. Neurology 80: 2154-2165, 2013.
33. Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, McCarthy MI, Ramos EM, Cardon LR, Chakravarti A, *et al*: Finding the missing heritability of complex diseases. Nature 461: 747-753, 2009.
34. Kubo M, Hata J, Ninomiya T, Matsuda K, Yonemoto K, Nakano T, Matsushita T, Yamazaki K, Ohnishi Y, Saito S, *et al*: A nonsynonymous SNP in PRKCH (protein kinase C eta) increases the risk of cerebral infarction. Nat Genet 39: 212-217, 2007.
35. Hata J, Matsuda K, Ninomiya T, Yonemoto K, Matsushita T, Ohnishi Y, Saito S, Kitazono T, Ibayashi S, Iida M, *et al*: Functional SNP in an Sp1-binding site of AGTRL1 gene is associated with susceptibility to brain infarction. Hum Mol Genet 16: 630-639, 2007.
36. Matsushita T, Ashikawa K, Yonemoto K, Hirakawa Y, Hata J, Amitani H, Doi Y, Ninomiya T, Kitazono T, Ibayashi S, *et al*: Functional SNP of ARHGEF10 confers risk of atherothrombotic stroke. Hum Mol Genet 19: 1137-1146, 2010.
37. Sacco RL, Kasner SE, Broderick JP, Caplan LR, Connors JJ, Culebras A, Elkind MS, George MG, Hamdan AD, Higashida RT, *et al*; American Heart Association Stroke Council, Council on Cardiovascular Surgery and Anesthesia; Council on Cardiovascular Radiology and Intervention; Council on Cardiovascular and Stroke Nursing; Council on Epidemiology and Prevention; Council on Peripheral Vascular Disease; Council on Nutrition, Physical Activity and Metabolism: An updated definition of stroke for the 21st century: A statement for healthcare professionals from the American Heart Association/American Stroke Association. Stroke 44: 2064-2089, 2013.
38. Grove ML, Yu B, Cochran BJ, Haritunians T, Bis JC, Taylor KD, Hansen M, Borecki IB, Cupples LA, Fornage M, *et al*: Best practices and joint calling of the HumanExome BeadChip: The CHARGE Consortium. PLoS One 8: e68095, 2013.
39. Anderson CA, Pettersson FH, Clarke GM, Cardon LR, Morris AP and Zondervan KT: Data quality control in genetic case-control association studies. Nat Protoc 5: 1564-1573, 2010.

40. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA and Reich D: Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet 38: 904-909, 2006.
41. Kang JQ and Barnes G: A common susceptibility factor of both autism and epilepsy: Functional deficiency of GABA A receptors. J Autism Dev Disord 43: 68-79, 2013.
42. Sinnett D, Wagstaff J, Glatt K, Woolf E, Kirkness EJ and Lalande M: High-resolution mapping of the gamma-aminobutyric acid receptor subunit beta 3 and alpha 5 gene cluster on chromosome 15q11-q13, and localization of breakpoints in two Angelman syndrome patients. Am J Hum Genet 52: 1216-1229, 1993.
43. Glatt K, Sinnett D and Lalande M: The human gamma-aminobutyric acid receptor subunit beta 3 and alpha 5 gene cluster in chromosome 15q11-q13 is rich in highly polymorphic (CA)n repeats. Genomics 19: 157-160, 1994.
44. Scapoli L, Martinelli M, Pezzetti F, Carinci F, Bodo M, Tognon M and Carinci P: Linkage disequilibrium between GABRB3 gene and nonsyndromic familial cleft lip with or without cleft palate. Hum Genet 110: 15-20, 2002.
45. Macdonald RL, Kang JQ and Gallagher MJ: Mutations in GABAA receptor subunits associated with genetic epilepsies. J Physiol 588: 1861-1869, 2010.
46. Delahanty RJ, Kang JQ, Brune CW, Kistner EO, Courchesne E, Cox NJ, Cook EH Jr, Macdonald RL and Sutcliffe JS: Maternal transmission of a rare GABRB3 signal peptide variant is associated with autism. Mol Psychiatry 16: 86-96, 2011.
47. Hooper JD, Clements JA, Quigley JP and Antalis TM: Type II transmembrane serine proteases. Insights into an emerging class of cell surface proteolytic enzymes. J Biol Chem 276: 857-860, 2001.
48. Bugge TH, Antalis TM and Wu Q: Type II transmembrane serine proteases. J Biol Chem 284: 23177-23181, 2009.
49. Antalis TM, Buzza MS, Hodge KM, Hooper JD and Netzel-Arnett S: The cutting edge: Membrane-anchored serine protease activities in the pericellular microenvironment. Biochem J 428: 325-346, 2010.
50. Luostari K, Hartikainen JM, Tengström M, Palvimo JJ, Kataja V, Mannermaa A and Kosma VM: Type II transmembrane serine protease gene variants associate with breast cancer. PLoS One 9: e102519, 2014.
51. Hayano T and Kikuchi M: Molecular cloning of the cDNA encoding a novel protein disulfide isomerase-related protein (PDIR). FEBS Lett 372: 210-214, 1995.
52. Vinaik R, Kozlov G and Gehring K: Structure of the non-catalytic domain of the protein disulfide isomerase-related protein (PDIR) reveals function in protein binding. PLoS One 8: e62021, 2013.
53. de Vries PS, Chasman DI, Sabater-Lleal M, Chen MH, Huffman JE, Steri M, Tang W, Teumer A, Marioni RE, Grossmann V, *et al*: A meta-analysis of 120 246 individuals identifies 18 new loci for fibrinogen concentration. Hum Mol Genet 25: 358-370, 2016.
54. Gieger C, Radhakrishnan A, Cvejic A, Tang W, Porcu E, Pistis G, Serbanovic-Canic J, Elling U, Goodall AH, Labrune Y, *et al*: New gene functions in megakaryopoiesis and platelet formation. Nature 480: 201-208, 2011.
55. Bylund J, Bylund M and Oliw EH: cDna cloning and expression of CYP4F12, a novel human cytochrome P450. Biochem Biophys Res Commun 280: 892-897, 2001.
56. Hashizume T, Imaoka S, Hiroi T, Terauchi Y, Fujii T, Miyazaki H, Kamataki T and Funae Y: cDNA cloning and expression of a novel cytochrome p450 (cyp4f12) from human small intestine. Biochem Biophys Res Commun 280: 1135-1141, 2001.
57. Kikuta Y, Kusunose E and Kusunose M: Prostaglandin and leukotriene omega-hydroxylases. Prostaglandins Other Lipid Mediat 68-69: 345-362, 2002.
58. Cauffiez C, Klinzig F, Rat E, Tournel G, Allorge D, Chevalier D, Pottier N, Lovecchio T, Colombel JF, Lhermitte M, *et al*: Human CYP4F12 genetic polymorphism: Identification and functional characterization of seven variant allozymes. Biochem Pharmacol 68: 2417-2425, 2004.
59. Ye X, Ji C, Huang Q, Cheng C, Tang R, Xu J, Zeng L, Dai J, Wu Q, Gu S, *et al*: Isolation and characterization of a human putative receptor protein kinase cDNA STYK1. Mol Biol Rep 30: 91-96, 2003.
60. Liu L, Yu XZ, Li TS, Song LX, Chen PL, Suo TL, Li YH, Wang SD, Chen Y, Ren YM, *et al*: A novel protein tyrosine kinase NOK that shares homology with platelet- derived growth factor/fibroblast growth factor receptors induces tumorigenesis and metastasis in nude mice. Cancer Res 64: 3491-3499, 2004.
61. Moriai R, Kobayashi D, Amachika T, Tsuji N and Watanabe N: Diagnostic relevance of overexpressed NOK mRNA in breast cancer. Anticancer Res 26: 4969-4973, 2006.
62. Amachika T, Kobayashi D, Moriai R, Tsuji N and Watanabe N: Diagnostic relevance of overexpressed mRNA of novel oncogene with kinase-domain (NOK) in lung cancers. Lung Cancer 56: 337-340, 2007.
63. Jackson KA, Oprea G, Handy J and Kimbro KS: Aberrant STYK1 expression in ovarian cancer tissues and cell lines. J Ovarian Res 2: 15, 2009.
64. Chung S, Tamura K, Furihata M, Uemura M, Daigo Y, Nasu Y, Miki T, Shuin T, Fujioka T, Nakamura Y, *et al*: Overexpression of the potential kinase serine/threonine/tyrosine kinase 1 (STYK1) in castration-resistant prostate cancer. Cancer Sci 100: 2109-2114, 2009.
65. Hu L, Chen HY, Cai J, Zhang Y, Qi CY, Gong H, Zhai YX, Fu H, Yang GZ and Gao CF: Serine threonine tyrosine kinase 1 is a potential prognostic marker in colorectal cancer. BMC Cancer 15: 246, 2015.
66. Wang Z, Qu L, Deng B, Sun X, Wu S, Liao J, Fan J and Peng Z: STYK1 promotes epithelial-mesenchymal transition and tumor metastasis in human hepatocellular carcinoma through MEK/ERK and PI3K/AKT signaling. Sci Rep 6: 33205, 2016.
67. Li J, Wu F, Sheng F, Li YJ, Jin D, Dingxand Zhang S: NOK/STYK1 interacts with GSK-3β and mediates Ser9 phosphorylation through activated Akt. FEBS Lett 586: 3787-3792, 2012.
68. Liu Y, Li T, Hu D and Zhang S: NOK/STYK1 promotes the genesis and remodeling of blood and lymphatic vessels during tumor progression. Biochem Biophys Res Commun 478: 254-259, 2016.
69. Giudice GJ, Emery DJ and Diaz LA: Cloning and primary structural analysis of the bullous pemphigoid autoantigen BP180. J Invest Dermatol 99: 243-250, 1992.
70. Pulkkinen L, Uitto J and Uitto J: Mutation analysis and molecular genetics of epidermolysis bullosa. Matrix Biol 18: 29-42, 1999.
71. Bauer JW and Lanschuetzer C: Type XVII collagen gene mutations in junctional epidermolysis bullosa and prospects for gene therapy. Clin Exp Dermatol 28: 53-60, 2003.
72. Kasperkiewicz M, Zillikens D and Schmidt E: Pemphigoid diseases: Pathogenesis, diagnosis, and treatment. Autoimmunity 45: 55-70, 2012.
73. Chen YJ, Wu CY, Lin MW, Chen TJ, Liao KK, Chen YC, Hwang CY, Chu SY, Chen CC, Lee DD, *et al*: Comorbidity profiles among patients with bullous pemphigoid: A nationwide population-based study. Br J Dermatol 165: 593-599, 2011.
74. Ameglio F, D'Auria L, Cordiali-Fei P, Mussi A, Valenzano L, D'Agosto G, Ferraro C, Bonifati C and Giacalone B: Bullous pemphigoid and pemphigus vulgaris: Correlated behaviour of serum VEGF, sE-selectin and TNF-alpha levels. J Biol Regul Homeost Agents 11: 148-153, 1997.
75. Seppänen A: Collagen XVII: A shared antigen in neurodermatological interactions? Clin Dev Immunol 2013: 240570, 2013.
76. Aoki T and Nishimura M: Molecular mechanism of cerebral aneurysm formation focusing on NF-κB as a key mediator of inflammation. J Biorheol 24: 16-21, 2010.
77. Chyatte D, Bruno G, Desai S and Todor DR: Inflammation and intracranial aneurysms. Neurosurgery 45: 1137-1147, 1999.
78. Kurki MI, Gaál EI, Kettunen J, Lappalainen T, Menelaou A, Anttila V, van't Hof FN, von Und Zu Fraunberg M, Helisalmi S, Hiltunen M, *et al*: High risk population isolate reveals low frequency variants predisposing to intracranial aneurysms. PLoS Genet 10: e1004134, 2014.