Identification of nine genes as novel susceptibility loci for early-onset ischemic stroke, intracerebral hemorrhage, or subarachnoid hemorrhage

YOSHIJI YAMADA[1,2], KIMIHIKO KATO[1,3], MITSUTOSHI OGURI[1,4], HIDEKI HORIBE[5], TETSUO FUJIMAKI[6], YOSHIKI YASUKOCHI[1,2], ICHIRO TAKEUCHI[2,7,8] and JUN SAKUMA[2,8,9]

[1]Department of Human Functional Genomics, Advanced Science Research Promotion Center, Mie University, Tsu, Mie 514-8507; [2]CREST, Japan Science and Technology Agency, Kawaguchi, Saitama 332-0012; [3]Department of Internal Medicine, Meitoh Hospital, Nagoya, Aichi 465-0025; [4]Department of Cardiology, Kasugai Municipal Hospital, Kasugai, Aichi 486-8510; [5]Department of Cardiovascular Medicine, Gifu Prefectural Tajimi Hospital, Tajimi, Gifu 507-8522; [6]Department of Cardiovascular Medicine, Northern Mie Medical Center Inabe General Hospital, Inabe, Mie 511-0428; [7]Department of Computer Science, Nagoya Institute of Technology, Nagoya, Aichi 466-8555; [8]RIKEN Center for Advanced Intelligence Project, Tokyo 103-0027; [9]Computer Science Department, College of Information Science, University of Tsukuba, Tsukuba, Ibaraki 305-8573, Japan

Received April 30, 2018; Accepted May 23, 2018

DOI: 10.3892/br.2018.1104

Abstract. Given that substantial genetic components have been shown in ischemic stroke, intracerebral hemorrhage (ICH), and subarachnoid hemorrhage (SAH), heritability may be higher in early-onset than late-onset individuals with these conditions. Although genome-wide association studies (GWASs) have identified various genes and loci significantly associated with ischemic stroke, ICH, or intracranial aneurysm mainly in European ancestry populations, genetic variants that contribute to susceptibility to these disorders remain to be identified definitively. We performed exome-wide association studies (EWASs) to identify genetic variants that confer susceptibility to ischemic stroke, ICH, or SAH in early-onset subjects with these conditions. A total of 6,649 individuals aged ≤65 years were examined. For the EWAS of ischemic or hemorrhagic stroke, 6,224 individuals (450 subjects with ischemic stroke, 5,774 controls) or 6,179 individuals (261 subjects with ICH, 176 subjects with SAH, 5,742 controls), respectively, were examined. EWASs were performed with the use of Illumina Human Exome-12 v1.2 DNA Analysis BeadChip or Infinium Exome-24 v1.0 BeadChip. To compensate for multiple comparisons of allele frequencies with ischemic stroke, ICH, or SAH, we applied a false discovery rate (FDR) of <0.05 for statistical significance of association. The association of allele frequencies of 31,245 single nucleotide polymorphisms (SNPs) that passed quality control to ischemic stroke was examined with Fisher's exact test, and 31 SNPs were significantly (FDR <0.05) associated with ischemic stroke. The association of allele frequencies of 31,253 or 30,970 SNPs to ICH or SAH, respectively, was examined with Fisher's exact test, and six or two SNPs were significantly associated with ICH or SAH, respectively. Multivariable logistic regression analysis with adjustment for age, sex, and the prevalence of hypertension and diabetes mellitus revealed that 12 SNPs were significantly [P<0.0004 (0.05/124)] related to ischemic stroke. Similar analysis with adjustment for age, sex, and the prevalence of hypertension revealed that six or two SNPs were significantly [P<0.0016 (0.05/32)] related to ICH or SAH, respectively. After examination of linkage disequilibrium of identified SNPs and results of previous GWASs, we identified *HHIPL2*, *CTNNA3*, *LOC643770*, *UTP20*, and *TRIB3* as susceptibility loci for ischemic stroke, *DNTTIP2* and *FAM205A* as susceptibility loci for ICH, and *FAM160A1* and *OR52E4* as such loci for SAH. Therefore, to the best of our knowledge, we have newly identified nine genes that confer susceptibility to early-onset ischemic stroke, ICH, or SAH. Determination of genotypes for the SNPs in these genes may prove informative for assessment of the genetic risk for ischemic stroke, ICH, or SAH in Japanese.

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 study

Introduction

Stroke is the leading cause of severe disability and a life-threatening condition (1). In 2015, there were 6.3 million stroke deaths worldwide (11.8% of total deaths), making stroke the second leading cause of death behind ischemic heart disease (2). Of all strokes, 87% are ischemic stroke, 10% are intracerebral

hemorrhage (ICH), and 3% are subarachnoid hemorrhage (SAH) in the United States (2). The etiology of common forms of stroke is multifactorial and includes both genetic and environmental factors (2,3). Studies with twins, siblings, and families provided substantial evidence for heritability of stroke (2,4). Given that personalized prevention is important to reduce overall burden of stroke, identification of genetic variants for stroke risk is key both for risk prediction and for potential intervention to avert future cerebrovascular events.

The Trial of ORG 10172 in Acute Stroke Treatment (TOAST) study (5) classified ischemic stroke into five subtypes: i) Large-artery atherosclerosis; ii) cardioembolism; iii) small-vessel occlusion; iv) stroke of other determined etiology; and v) stroke of undetermined etiology. A family history study of 1,000 individuals with ischemic stroke and 800 controls showed that a family history of stroke was a risk factor for both large-vessel atherosclerosis and small-vessel occlusion, especially in cases aged <65 years (6). The heritability of ischemic stroke was estimated to be 40.3% for large-vessel disease, 32.6% for cardioembolic stroke, 16.1% for small-vessel disease, and 37.9% for all ischemic stroke (7). These observations suggest that genetic components play important roles in the pathogenesis of ischemic stroke, in addition to conventional risk factors such as hypertension and diabetes mellitus (2,8,9).

Genome-wide association studies (GWASs) in European ancestry populations identified various genes and loci that confer susceptibility to ischemic stroke (10-17). A recent GWAS identified *HDAC9* and chromosome 1p13.2 (near *TSPAN2*) as susceptibility loci for large-vessel disease, *PITX2* and *ZFHX3* as 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 (18). A more recent multiancestry meta-analysis of GWASs identified 32 loci including 22 new loci that confer susceptibility to stroke (19).

ICH accounts for a large proportion of severe or fatal cases of stroke, with its most important risk factors being hypertension and advanced age (20). Familial aggregation of ICH cases was demonstrated in a prospective study in North Carolina in the United States, showing that 10% of affected individuals had a family history of ICH (21). Genetic factors may influence, not only the development of ICH, but also the prevalence of risk factors for this condition, such as hypertension (22).

ICH has a substantial genetic component with heritability of deep or lobar ICH being estimated at 34 and 73%, respectively (23). Previous genetic association studies suggested several genes and loci are involved in the predisposition to ICH (24-27). A 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 a locus for non-lobar ICH (28).

SAH is commonly caused by rupture of an aneurysm in an intracranial artery (29). Given that a family history is an important risk factor for the development of intracranial aneurysm, genetic components may play important roles in the development of SAH (30-32). The heritability of SAH was estimated as 41% (33). GWASs have implicated several loci and genes that confer susceptibility to intracranial aneurysm (34-38). A meta-analysis of GWASs identified 19 genetic variants associated with intracranial aneurysm (39).

Although several single nucleotide polymorphisms (SNPs) have been found to be significantly associated with ischemic stroke (40,41) or intracranial aneurysm (42) in Japanese subjects, genetic variants that confer susceptibility to ischemic stroke, ICH, or SAH in Japanese remain to be identified definitively.

In a family and twin study of ischemic stroke, heritability was higher in early-onset than late-onset individuals with this condition (6,43,44), suggesting that early-onset ischemic stroke has a strong genetic component. Similar to ischemic stroke, early-onset ICH (45) and SAH (46) were shown to have strong genetic components. Given that a genetic contribution may be greater in early-onset forms than in late-onset forms of ischemic stroke, ICH, and SAH, statistical power of the genetic association study may be increased by focusing on early-onset subjects with diseases (6).

We performed exome-wide association studies (EWASs) with the use of human exome array-based genotyping methods to identify genetic variants that confer susceptibility to ischemic stroke, ICH, or SAH in Japanese individuals. To increase the statistical power of EWASs, we examined early-onset subjects with these conditions.

Materials and methods

Study subjects. In our previous study (47), the median age of subjects with ischemic stroke, ICH, or SAH was 74, 71, or 60 years, respectively. We thus defined patients aged ≤65 years as early-onset cases in the present study. A total of 6,649 individuals aged ≤65 years were examined. For the EWAS of ischemic stroke, 6,224 individuals (450 subjects with ischemic stroke, and 5,774 controls) were examined. For the EWAS of hemorrhagic stroke, 6,179 individuals (261 subjects with ICH, 176 subjects with SAH, and 5,742 controls) were examined. Most control individuals were the same for the studies of ischemic and hemorrhagic stroke. The subjects were recruited from individuals who visited outpatient clinics of or were admitted to participating hospitals in Japan (Gifu Prefectural Tajimi Hospital, Tajimi; Gifu Prefectural General Medical Center, Gifu; Japanese Red Cross Nagoya First Hospital, Nagoya; Northern Mie Medical Center Inabe General Hospital, Inabe; Hirosaki University Hospital and Hirosaki Stroke and Rehabilitation Center, Hirosaki) because of various symptoms or for an annual health checkup between October 2002 and March 2014; or who were community-dwelling individuals recruited to a population-based cohort study in Inabe between March 2010 and September 2014 (48).

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 (49). Given that susceptibility loci for ischemic stroke are subtype-specific (18), 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 controls. The absence of stroke history was evaluated with a detailed questionnaire and was confirmed by the absence of a history of neurological deficits.

EWASs. Venous blood 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 Srl, Trieste, Italy) or SMITEST EX-R&D (Medical & Biological Laboratories, Co., Ltd., Nagoya, Japan). EWASs were performed with the use of a Human Exome-12 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 (50). SNPs contained in only one of the exome arrays (~2.6% of all SNPs) were excluded from analysis. We performed quality control (51) 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 checked 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 checked 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 minor allele frequency of <1.0%. vi) SNPs whose genotype distributions deviated significantly (P<0.01) from Hardy-Weinberg equilibrium in control individuals were discarded. vii) Genotype data were examined for population stratification by principal components analysis (52), and population outliers were excluded from the analysis. Totals of 31,245, 31,253, or 30,970 SNPs that passed quality control for the study of ischemic stroke, ICH, or SAH, respectively, were subjected to analysis.

Statistical analysis. For analysis of characteristics of the study subjects, quantitative data were presented as means ± SD, and were compared between subjects with ischemic stroke, ICH, or SAH and controls with the unpaired Student's t-test. Categorical data were compared between two groups with Pearson's Chi-square 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 corresponding controls with Fisher's exact test. To compensate for multiple comparisons of allele frequencies with ischemic stroke, ICH, or SAH, we applied a false discovery rate (FDR) (53) for statistical significance of association. The significance level was set at a FDR of <0.05 for each EWAS. The inflation factor (λ) was 1.06 for ischemic stroke, 1.10 for ICH, and 1.11 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. A 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 association of genotypes of SNPs to intermediate phenotypes was examined with Pearson's Chi-square test and P-values were shown. Bonferroni's correction was applied to other statistical analyses as indicated. Statistical tests were performed with JMP Genomics version 9.0 software (SAS Institute, Cary, NC, USA).

Results

Characteristics of subjects. The characteristics of the 6,224 subjects enrolled in the ischemic stroke study are shown in Table I. Age, the frequency of men, and the prevalence of hypertension, diabetes mellitus, dyslipidemia, and chronic kidney disease as well as systolic and diastolic blood pressure (BP), fasting plasma glucose (FPG) level, blood glycosylated hemoglobin (hemoglobin A_{1c}) content, and serum concentrations of triglycerides were greater, whereas serum concentration of high density lipoprotein (HDL)-cholesterol and estimated glomerular filtration rate (eGFR) were lower, in subjects with ischemic stroke than in controls.

The characteristics of the subjects enrolled in the hemorrhagic stroke study are shown in Table II. Age, the frequency of men, and the prevalence of hypertension, diabetes mellitus, and chronic kidney disease as well as systolic and diastolic BP, FPG level, blood hemoglobin A_{1c} content, and serum concentrations of triglycerides were greater, whereas serum concentrations of HDL-cholesterol and low density lipoprotein (LDL)-cholesterol were lower, in subjects with ICH than in controls. The prevalence of hypertension, diabetes mellitus, and chronic kidney disease as well as systolic and diastolic BP, FPG level, and serum concentrations of triglycerides were greater, whereas the prevalence of dyslipidemia and the serum

Table I. Characteristics of subjects with ischemic stroke and control individuals.

Characteristic	Control	Ischemic stroke	P-value
No. of subjects	5,774	450	
Age (years)	50.6±10.2	56.7±7.1	<0.0001
Sex (men/women, %)	52.1/47.9	67.8/32.2	<0.0001
Smoking (%)	42.5	35.6	0.0093
Obesity (%)	31.0	33.3	0.3484
Body mass index (kg/m^2)	23.2±3.5	23.9±3.8	0.0002
Hypertension (%)	31.7	72.5	<0.0001
Systolic BP (mmHg)	121±18	149±30	<0.0001
Diastolic BP (mmHg)	75±13	86±17	<0.0001
Diabetes mellitus (%)	12.7	47.5	<0.0001
Fasting plasma glucose (mmol/l)	5.66±1.78	7.16±3.00	<0.0001
Blood hemoglobin A_{1c} (%)	5.72±0.96	6.52±1.66	<0.0001
Dyslipidemia (%)	56.9	66.3	0.0001
Serum triglycerides (mmol/l)	1.32±0.98	1.67±1.03	<0.0001
Serum HDL-cholesterol (mmol/l)	1.65±0.45	1.30±0.42	<0.0001
Serum LDL-cholesterol (mmol/l)	3.18±0.83	3.13±0.93	0.5012
Chronic kidney disease (%)	10.3	31.2	<0.0001
Serum creatinine (μmol/l)	69.8±61.0	88.4±120.2	0.0041
eGFR (ml min^{-1} 1.73 m^{-2})	78.7±17.1	71.1±23.7	<0.0001
Hyperuricemia (%)	15.2	19.1	0.0290
Serum uric acid (μmol/l)	321±89	337±96	0.0027

Quantitative data are means ± standard deviations and were compared between subjects with ischemic stroke and controls with the unpaired Student's t-test. Categorical data were compared between two groups with Pearson's Chi-square test. Based on Bonferroni's correction, a P-value of <0.0025 (0.05/20) was considered statistically significant. BP, blood pressure; HDL, high density lipoprotein; LDL, low density lipoprotein; eGFR, estimated glomerular filtration rate.

concentration of HDL-cholesterol were lower, in subjects with SAH than in controls.

EWAS for ischemic stroke, ICH, or SAH. We examined the association of allele frequencies of 31,245 SNPs that passed quality control to ischemic stroke with the use of Fisher's exact test, and detected that 31 SNPs were significantly (FDR <0.05) associated with ischemic stroke (Table III). The relation of allele frequencies of 31,253 or 30,970 SNPs to ICH or SAH, respectively, was examined with Fisher's exact test. Six or two SNPs were significantly associated with ICH or SAH, respectively (Table IV).

Multivariable logistic regression analysis of the association of SNPs to ischemic stroke, ICH, or SAH. The association of the 31 identified SNPs in the EWAS of ischemic stroke was further examined by multivariable logistic regression analysis with adjustment for age, sex, and the prevalence of hypertension and diabetes mellitus (Table V). The 12 SNPs were significantly [P<0.0004 (0.05/124) in at least one genetic model] related to ischemic stroke. The association of the six or two SNPs identified in the EWAS for ICH or SAH, respectively, to these conditions was examined by multivariable logistic regression analysis with adjustment for age, sex, and the prevalence of hypertension (Table VI). The SNPs were significantly [P<0.0016 (0.05/32)] related to ICH or SAH.

Relationship of SNPs associated with ischemic stroke, ICH, or SAH to intermediate phenotypes. We examined the relationship of the 12 SNPs associated with ischemic stroke to intermediate phenotypes of this condition (including hypertension, diabetes mellitus, hypertriglyceridemia, hypo-HDL-cholesterolemia, hyper-LDL-cholesterolemia, chronic kidney disease, obesity, and hyperuricemia) with the use of Pearson's Chi-square test. None of the SNPs was related to intermediate phenotypes (Table VII). The relationship of six or two SNPs associated with ICH or SAH, respectively, to intermediate phenotypes of these conditions was also examined. The rs12229654 at chromosome 12q24.1, rs671 of *ALDH2*, and rs11066015 of *ACAD10* associated with ICH were significantly related to hypertension, hyper-LDL-cholesterolemia, and hyperuricemia, whereas none of SNPs associated with SAH was related to intermediate phenotypes (Table VIII).

Linkage disequilibrium analyses. We examined linkage disequilibrium (LD) among SNPs associated with ischemic stroke or ICH. For the ischemic stroke study, rs3130688 at chromosome 6p21.3 and rs2308557 of *HLA-C* were in complete LD [square of the correlation coefficient (r^2), 1.000], whereas rs3130981 and rs3130984 of *CDSN* were not in LD. For the ICH study, there was significant LD (r^2, 0.650 to 0.995) among rs12229654 at 12q24.1, rs11066015 of *ACAD10*, and rs671 of *ALDH2* (data not shown).

Table II. Characteristics of subjects with ICH or SAH and control individuals.

Characteristic	Control	ICH	P-value	SAH	P-value
No. of subjects	5,742	261		176	
Age (years)	50.5±10.2	55.1±7.6	<0.0001	52.2±9.2	0.0172
Sex (men/women, %)	52.1/47.9	70.9/29.1	<0.0001	42.6/57.4	0.0134
Smoking (%)	42.4	37.6	0.1915	33.0	0.0434
Obesity (%)	30.9	30.9	0.9956	25.4	0.2094
Body mass index (kg/m^2)	23.2±3.5	23.3±3.8	0.5923	23.1±3.2	0.8343
Hypertension (%)	31.6	73.5	<0.0001	59.4	<0.0001
Systolic BP (mmHg)	121±18	150±29	<0.0001	149±27	<0.0001
Diastolic BP (mmHg)	75±13	88±17	<0.0001	85±16	<0.0001
Diabetes mellitus (%)	12.7	33.1	<0.0001	21.5	0.0012
Fasting plasma glucose (mmol/l)	5.66±1.83	6.66±2.44	<0.0001	6.61±2.44	0.0002
Blood hemoglobin A_{1c} (%)	5.72±0.97	6.24±1.33	<0.0001	5.97±1.28	0.1712
Dyslipidemia (%)	56.7	54.5	0.4854	42.5	0.0002
Serum triglycerides (mmol/l)	1.31±0.98	1.63±0.91	<0.0001	1.89±1.63	0.0003
Serum HDL-cholesterol (mmol/l)	1.65±0.45	1.27±0.46	<0.0001	1.35±0.37	<0.0001
Serum LDL-cholesterol (mmol/l)	3.18±0.83	2.90±0.85	<0.0001	2.95±0.93	0.0159
Chronic kidney disease (%)	10.3	18.3	0.0004	25.0	<0.0001
Serum creatinine (μmol/l)	69.8±61.0	70.7±31.8	0.7555	68.1±34.5	0.5044
eGFR (ml min^{-1} 1.73 m^{-2})	78.7±17.1	79.4±24.1	0.6912	79.1±27.5	0.8795
Hyperuricemia (%)	15.1	17.7	0.2616	10.3	0.0817
Serum uric acid (μmol/l)	321±89	335±113	0.0872	303±150	0.2510

Quantitative data are means ± standard deviations and were compared between subjects with ICH or SAH and controls with the unpaired Student's t-test. Categorical data were compared between two groups with Pearson's Chi-square test. Based on Bonferroni's correction, a P-value of <0.0013 (0.05/40) was considered statistically significant. ICH, intracerebral hemorrhage; SAH, subarachnoid hemorrhage; BP, blood pressure; HDL, high density lipoprotein; LDL, low density lipoprotein; eGFR, estimated glomerular filtration rate.

Association of genes, chromosomal loci, and SNPs identified in the present study to phenotypes reported by previously GWASs. In the ischemic stroke study, *CDK18* was shown to be related to type 1 diabetes mellitus (T1DM); *CDSN* to T1DM and serum concentrations of triglycerides; *HLA-C* to T1DM and serum concentrations of triglycerides and LDL-cholesterol; *NYNRIN* to serum concentration of LDL-cholesterol; and *GOSR2* to systolic BP. The remaining five genes (*HHIPL2*, *CTNNA3*, *LOC643770*, *UTP20*, *TRIB3*) were not found to be related to ischemic stroke or other cerebrovascular disease-related phenotypes (Table IX). In the hemorrhagic stroke study, *SVEP1* was shown to be related to coronary artery disease (CAD); chromosome 12q24.1 to serum HDL-cholesterol level; *ACAD10* to CAD, serum LDL-cholesterol level, T1DM, and diastolic BP; and *ALDH2* to CAD, myocardial infarction (MI), serum concentrations of HDL-cholesterol and LDL-cholesterol, T1DM, and systolic and diastolic BP. The remaining four genes (*DNTTIP2*, *FAM205A*, *FAM160A1*, *OR52E4*) were not related to ICH, SAH, or other cerebrovascular disease-related phenotypes (Table X).

Discussion

Given that stroke is a serious condition and is a global public health problem (1,2,20,29), identification of genetic variants that confer susceptibility to ischemic stroke, ICH, and SAH is clinically important to prevent these conditions. In the present study, we performed EWASs for ischemic stroke, ICH, and SAH in early-onset subjects who may have greater genetic components compared with late-onset individuals.

In the study of ischemic stroke, among 10 genes and one chromosomal locus identified, *CDK18*, *CDSN*, *HLA-C*, *NYNRIN* and *GOSR2* were shown to be related to T1DM (55-56), serum concentrations of triglycerides or LDL-cholesterol (57), or systolic BP (58,59) which are risk factors for ischemic stroke. Although rs3130688 at chromosomal 6p21.3 was not related to cerebrovascular phenotype, this region was previously related to CAD (60). We thus identified *HHIPL2*, *CTNNA3*, *LOC643770*, *UTP20* and *TRIB3* as novel susceptibility loci for ischemic stroke. The five genes associated with ischemic stroke were not related to intermediate phenotypes, although the relation of *UTP20* to chronic kidney disease was borderline significance. The underline molecular mechanisms of the association of these genes with ischemic stroke remain unclear.

In the study of hemorrhagic stroke, of the seven genes and one chromosomal locus identified, *SVEP1*, *ACAD10* and *ALDH2* were shown to be related to CAD or MI (60-62). *ACAD10* and *ALDH2* were previously related to systolic or diastolic BP (59,63); and these genes as well as chromosome 12q24.1 were previously related to the serum concentrations

Table III. The 31 SNPs significantly (FDR <0.05) associated with ischemic stroke in the exome-wide association study.

Gene	SNP	Nucleotide substitution[a]	Amino acid substitution	Chromosome	Position	MAF (%)	Allele OR	P-value (allele frequency)	FDR (allele frequency)
PLCB2	rs200787930	C/T	E1106K	15	40289298	1.2	0.07	3.81×10^{-9}	6.48×10^{-6}
VPS33B	rs199921354	C/T	R80Q	15	91013841	1.2	0.07	5.62×10^{-9}	8.83×10^{-6}
CXCL8	rs188378669	G/T	E31*	4	73741568	1.2	0.07	5.71×10^{-9}	8.83×10^{-6}
MARCH1	rs61734696	G/T	Q137K	4	164197303	1.2	0.07	5.89×10^{-9}	8.83×10^{-6}
ADGRL3	rs192210727	G/T	R580I	4	61909615	1.3	0.07	5.85×10^{-9}	8.83×10^{-6}
TMOD4	rs115287176	G/A	R277W	1	151170961	1.2	0.08	8.35×10^{-9}	1.22×10^{-5}
COL6A3	rs146092501	C/T	E1386K	2	237371861	1.2	0.08	1.25×10^{-8}	1.72×10^{-5}
ZNF77	rs146879198	G/A	R340*	19	2934109	1.2	0.08	1.25×10^{-8}	1.72×10^{-5}
NYNRIN	rs149771079	G/A	D467N	14	24409193	1.2	3.33	5.03×10^{-8}	6.23×10^{-5}
GOSR2	rs1052586	T/C		17	46941097	48.7	0.70	2.01×10^{-6}	0.0020
	rs12662501	C/T		6	31223073	7.3	1.68	1.14×10^{-5}	0.0104
	rs17435433	T/C		2	88210097	25.8	1.39	1.24×10^{-5}	0.0111
	rs7453967	T/G		6	31346466	15.3	1.49	1.85×10^{-5}	0.0153
HLA-C	rs2308557	G/A	S101N	6	31271640	8.4	1.62	2.54×10^{-5}	0.0205
	rs3130688	T/C		6	31242439	18.6	1.43	2.77×10^{-5}	0.0215
HHIPL2	rs3748665	C/T	R394Q	1	222540279	7.4	0.51	2.88×10^{-5}	0.0218
MUC22	rs11756038	A/G	T1376A	6	31029557	5.0	1.73	5.41×10^{-5}	0.0392
LOC643770	rs829881	C/A		12	98487450	37.4	1.33	6.09×10^{-5}	0.0435
TRIB3	rs2295490	A/G	Q84R	20	388261	23.0	1.37	6.35×10^{-5}	0.0449
CDSN	rs3130984	C/T	S143N	6	31117187	13.4	1.47	6.49×10^{-5}	0.0449
HLA-DQB1	rs1130375	C/G	A45G	6	32665043	28.3	0.73	6.49×10^{-5}	0.0449
CTNNA3	rs10997469	C/T		10	66986527	26.1	1.35	7.04×10^{-5}	0.0477
C6orf15	rs2233977	T/C	V81A	6	31112117	44.0	0.76	7.44×10^{-5}	0.0492
	rs4713433	C/A		6	31100249	44.0	0.76	7.44×10^{-5}	0.0492
CDSN	rs3130981	C/T	D527N	6	31116036	13.6	1.46	7.48×10^{-5}	0.0492
DDAH1	rs12742253	T/G		1	85505292	44.9	0.76	8.45×10^{-5}	0.0499
HLA-DQB1	rs1130370	A/C	Y69D	6	32664972	18.6	0.69	8.19×10^{-5}	0.0499
	rs3131931	A/T		6	30977488	5.6	1.69	7.72×10^{-5}	0.0499
CDK18	rs77571454	G/A	G466E	1	205531350	5.1	1.72	8.4×10^{-5}	0.0499
CTNNA3	rs1925608	A/C		10	66990654	32.2	1.33	8.35×10^{-5}	0.0499
UTP20	rs117417637	G/A	R1520H	12	101344704	1.8	2.33	8.07×10^{-5}	0.0499

Allele frequencies were analyzed with Fisher's exact test. [a]Major allele/minor allele. SNP, single nucleotide polymorphism; FDR, false discovery rate; MAF, minor allele frequency; OR, odds ratio.

Table IV. The eight SNPs significantly (FDR <0.05) associated with intracerebral hemorrhage or subarachnoid hemorrhage in the exome-wide association study.

Gene	SNP	Nucleotide substitution[a]	Amino acid substitution	Chromosome	Position	MAF (%)	Allele OR	P-value (allele frequency)	FDR (allele frequency)
Intracerebral hemorrhage									
	rs12229654	T/G		12	110976657	22.5	0.59	7.66×10^{-6}	0.0093
FAM205A	rs3739881	A/C	I999S	9	34724244	36.9	0.66	1.53×10^{-5}	0.0178
SVEP1	rs7030192	G/A	A2750V	9	110407351	40.9	1.47	1.71×10^{-5}	0.0192
DNTTIP2	rs3747965	T/G	D309E	1	93877008	42.5	1.47	2.86×10^{-5}	0.0316
ALDH2	rs671	G/A	E504K	12	111803962	27.6	0.65	4.76×10^{-5}	0.0483
ACAD10	rs11066015	G/A		12	111730205	27.5	0.65	4.76×10^{-5}	0.0483
Subarachnoid hemorrhage									
OR52E4	rs11823828	T/G	F227L	11	5884973	36.6	1.82	2.48×10^{-6}	0.0063
FAM160A1	rs2709828	C/T		4	151434116	33.1	0.57	5.96×10^{-6}	0.0145

Allele frequencies were analyzed with Fisher's exact test. [a]Major allele/minor allele. SNP, single nucleotide polymorphism; FDR, false discovery rate; MAF, minor allele frequency; OR, odds ratio.

Table V. Association of SNPs to ischemic stroke as determined by multivariable logistic regression analysis.

			Dominant			Recessive			Additive 1			Additive 2		
Gene	SNP		P-value	OR	95% CI	P-value	OR	95% CI	P-value	OR	95% CI	P-value	OR	95% CI
NYNRIN	rs149771079	G/A	<0.0001	3.21	2.03-5.07				<0.0001	3.21	2.03-5.07			
GOSR2	rs1052586	T/C	0.0077	0.71	0.55-0.91	<0.0001	0.47	0.34-0.64	0.2492			<0.0001	0.42	0.29-0.61
HLA-C	rs2308557	G/A	0.0004	1.61	1.24-2.09	0.4809			0.0002	1.65	1.27-2.15	0.5880		
	rs3130688	T/C	<0.0001	1.59	1.28-1.97	0.5185			<0.0001	1.61	1.28-2.02	0.2174		
HHIPL2	rs3748665	C/T	0.0003	0.50	0.34-0.73	0.4082			0.0005	0.50	0.34-0.74	0.4163		
LOC643770	rs829881	C/A	0.0004	1.51	1.20-1.90	<0.0001	1.71	1.31-2.24	0.0146	1.36	1.06-1.73	<0.0001	2.03	1.50-2.76
TRIB3	rs2295490	A/G	<0.0001	1.53	1.24-1.90	0.0021	1.88	1.26-2.81	0.0015	1.44	1.15-1.81	0.0002	2.18	1.44-3.30
CDSN	rs3130984	C/T	<0.0001	1.62	1.28-2.04	0.9814			<0.0001	1.65	1.30-2.09	0.7537		
CDSN	rs3130981	C/T	<0.0001	1.61	1.28-2.03	0.9850			<0.0001	1.64	1.30-2.08	0.7588		
CDK18	rs77571454	G/A	0.0001	1.86	1.36-2.55	0.2692			<0.0001	1.90	1.38-2.61	0.9969		
CTNNA3	rs1925608	A/C	0.0080	1.34	1.08-1.67	0.0018	1.64	1.20-2.24	0.0720			0.0003	1.83	1.31-2.56
UTP20	rs117417637	G/A	<0.0001	2.72	1.76-4.20	0.5873			<0.0001	2.76	1.79-4.26	0.9965		

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 <0.0004 (0.05/124) was considered statistically significant. SNP, single nucleotide polymorphism; OR, odds ratio; CI, confidence interval.

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

Gene	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
Intracerebral hemorrhage														
	rs12229654	T/G	0.0002	0.58	0.44-0.77	0.1592			0.0005	0.59	0.44-0.79	0.0814		
FAM205A	rs3739881	A/C	0.0008	0.64	0.50-0.83	0.0008	0.42	0.26-0.70	0.0214	0.73	0.56-0.95	0.0001	0.36	0.22-0.61
SVEP1	rs7030192	G/A	0.0003	1.73	1.28-2.34	0.0057	1.55	1.14-2.11	0.0027	1.61	1.18-2.21	0.0001	2.11	1.45-3.08
DNTTIP2	rs3747965	T/G	0.0063	1.52	1.13-2.05	0.0002	1.78	1.31-2.41	0.1320			<0.0001	2.11	1.45-3.06
ALDH2	rs671	G/A	0.0004	0.61	0.47-0.80	0.7561			0.0003	0.58	0.43-0.78	0.3117		
ACAD10	rs11066015	G/A	0.0004	0.61	0.47-0.81	0.7534			0.0003	0.58	0.44-0.78	0.3118		
Subarachnoid hemorrhage														
OR52E4	rs11823828	T/G	0.0349	1.48	1.03-2.14	<0.0001	3.12	2.14-4.54	0.9690			<0.0001	3.11	2.02-4.78
FAM160A1	rs2709828	C/T	0.0002	0.55	0.41-0.75	0.0025	0.31	0.14-0.66	0.0047	0.63	0.46-0.87	0.0005	0.25	0.12-0.54

Multivariable logistic regression analysis was performed with adjustment for age, sex, and the prevalence of hypertension. Based on Bonferroni's correction, a P-value of <0.0016 (0.05/32) was considered statistically significant. SNP, single nucleotide polymorphism; OR, odds ratio; CI, confidence interval.

Table VII. Relation of SNPs associated with ischemic stroke to intermediate phenotypes.

Gene	SNP		Hypertension	DM	Hyper-TG	Hypo-HDL	Hyper-LDL	CKD	Obesity	Hyperuricemia
NYNRIN	rs149771079	G/A	0.0054	0.0686	0.2117	0.6340	0.9385	0.6942	0.9354	0.2993
GOSR2	rs1052586	T/C	0.7587	0.7483	0.2289	0.7387	0.2092	0.3107	0.1309	0.8628
HLA-C	rs2308557	G/A	0.0188	0.1881	0.4974	0.1409	0.7632	0.5904	0.4666	0.0329
	rs3130688	T/C	0.3913	0.4724	0.0746	0.4322	0.0165	0.5385	0.4440	0.7564
HHIPL2	rs3748665	C/T	0.9811	0.7782	0.6295	0.4509	0.3465	0.0781	0.2633	0.4430
LOC643770	rs829881	C/A	0.4804	0.1812	0.2885	0.5240	0.7785	0.1223	0.1292	0.5389
TRIB3	rs2295490	A/G	0.6434	0.6909	0.5038	0.1135	0.1603	0.0762	0.2842	0.2079
CDSN	rs3130984	C/T	0.4946	0.7936	0.4541	0.3164	0.0451	0.8039	0.9544	0.0920
CDSN	rs3130981	C/T	0.4921	0.8212	0.4463	0.3344	0.0459	0.8039	0.9541	0.0923
CDK18	rs77571454	G/A	0.5132	0.7996	0.7437	0.7177	0.9208	0.5647	0.2642	0.3218
CTNNA3	rs1925608	A/C	0.5025	0.8597	0.2367	0.8708	0.1496	0.7296	0.7227	0.4908
UTP20	rs117417637	G/A	0.2032	0.9101	0.5860	0.1956	0.0396	0.0006	0.1973	0.4511

Data are P-values. The relationship of genotypes of each SNP to intermediate phenotypes was examined with Pearson's Chi-square test. SNP, single nucleotide polymorphism; DM, diabetes mellitus; hyper-TG, hypertriglyceridemia; hypo-HDL, hypo-HDL-cholesterolemia; hyper-LDL, hyper-LDL-cholesterolemia; CKD, chronic kidney disease. Based on Bonferroni's correction; a P-value of <0.0005 (0.05/96) was considered statistically significant.

Table VIII. Relationship of SNPs associated with hemorrhagic stroke to intermediate phenotypes.

Gene	SNP		Hypertension	DM	Hyper-TG	Hypo-HDL	Hyper-LDL	CKD	Obesity	Hyperuricemia
Intracerebral hemorrhage										
	rs12229654	T/G	**<0.0001**	0.0430	0.1567	0.4724	**0.0001**	0.5229	0.0456	**<0.0001**
FAM205A	rs3739881	A/C	0.6723	0.1525	0.0535	0.1239	0.0129	0.5238	0.9775	0.5100
SVEP1	rs7030192	G/A	0.0317	0.8434	0.6859	0.6314	0.9587	0.0907	0.6029	0.1534
DNTTIP2	rs3747965	T/G	0.3406	0.0796	0.0316	0.0921	0.7314	0.3182	0.1242	0.4615
ALDH2	rs671	G/A	**<0.0001**	0.0373	0.0677	0.0286	**<0.0001**	0.5559	0.0086	**<0.0001**
ACAD10	rs11066015	G/A	**<0.0001**	0.0708	0.0621	0.0249	**<0.0001**	0.6322	0.0101	**<0.0001**
Subarachnoid hemorrhage										
OR52E4	rs11823828	T/G	0.0325	0.2354	0.1035	0.9486	0.9093	0.2094	0.4177	0.0205
FAM160A1	rs2709828	C/T	0.5113	0.9486	0.8178	0.3166	0.1606	0.1469	0.9885	0.7020

Data are P-values. The relation of genotypes of each SNP to intermediate phenotypes was examined with Pearson's Chi-square test. Based on Bonferroni's correction; a P-value of <0.0008 (0.05/64) was considered statistically significant and is shown in bold. SNP, single nucleotide polymorphism; DM, diabetes mellitus; hyper-TG, hypertriglyceridemia; hypo-HDL, hypo-HDL-cholesterolemia; hyper-LDL, hyper-LDL-cholesterolemia; CKD, chronic kidney disease.

of HDL-cholesterol or LDL-cholesterol (57,64). These phenotypes are related to cerebrovascular disease. We thus identified *DNTTIP2* and *FAM205A* as new susceptibility loci for ICH, and *FAM160A1* and *OR52E4* as loci for SAH. Given that the four genes associated with ICH or SAH were not related to intermediate phenotypes, the functional relevance of the association of these genes with ICH or SAH remains to be elucidated.

We previously showed that four, six, or three SNPs were associated with ischemic stroke (P<0.01), ICH (P<0.05), or SAH (P<0.05), respectively, as determined by multivariable logistic regression analysis with adjustment for covariates after the initial EWAS screening among both early- and late-onset subjects with these conditions (47). The relationship of four SNPs to ischemic stroke was not replicated (P<0.05) in the present study. The relation of one of six SNPs [rs138533962 (P=0.0019)] to ICH was replicated in the present study. The association of one of three SNPs [rs117564807 (P=0.0454)] to SAH was replicated in the present study. The results suggest that genetic variants that confer susceptibility to ischemic stroke, ICH, or SAH may differ, in part, between early-onset and late-onset subjects with these conditions.

There are several limitations to our study: i) Given that the results were not replicated, their validation will be necessary in independent study populations or in other ethnic groups. ii) It is possible that SNPs identified in the present study are in LD with other genetic variants in the same gene or in other nearby genes that are actually responsible for the development of ischemic stroke, ICH, or SAH. iii) The functional relevance of identified SNPs to the pathogenesis of ischemic stroke, ICH, or SAH remains to be elucidated.

In conclusion, we have newly identified five (*HHIPL2*, *CTNNA3*, *LOC643770*, *UTP20*, *TRIB3*), two (*DNTTIP2*, *FAM205A*), or two (*FAM160A1*, *OR52E4*) genes as susceptibility loci for early-onset ischemic stroke, ICH, or SAH, respectively. Determination of genotypes for the SNPs in these genes may prove informative for assessment of the genetic risk for ischemic stroke, ICH, or SAH in Japanese subjects.

Acknowledgements

Not applicable.

Funding

This study was supported by CREST (JPMJCR1302), Japan Science and Technology Agency (to YYam, JS and IT).

Availability of data and materials

All data underlying the findings described in the article are available upon request from the corresponding author.

Authors' contributions

YYam contributed to conception and design of the study; to acquisition, analysis, and interpretation of the data; and to drafting of the manuscript. KK, MO, HH and TF each contrib-

Table IX. Relationship of genes, chromosomal loci, and SNPs associated with ischemic stroke in the present study to previously reported cerebrovascular disease-related phenotypes.

Gene/chr. locus	SNP	Chr.	Position	Previously reported phenotypes
CDK18	rs77571454	1	205531350	Type 1 diabetes (21980299)
HHIPL2	rs3748665	1	222540279	None
CDSN	rs3130981	6	31116036	Type 1 diabetes (17554300),
	rs3130984	6	31117187	triglycerides (20686565)
6p21.3	rs3130688	6	31242439	None
HLA-C	rs2308557	6	31271640	Type 1 diabetes (17632545), total cholesterol (20686565), triglycerides (20686565), LDL-cholesterol (20686565)
CTNNA3	rs1925608	10	66990654	None
LOC643770	rs829881	12	98487450	None
UTP20	rs117417637	12	101344704	None
NYNRIN	rs149771079	14	24409193	LDL-cholesterol (20686565), total cholesterol (20686565)
GOSR2	rs1052586	17	46941097	Systolic blood pressure (21909110, 21909115)
TRIB3	rs2295490	20	388261	None

Data were obtained from genome-wide repository of associations between SNPs and phenotypes (GRASP) search database (https://grasp.nhlbi.nih.gov/Search.aspx) with a P-value of $<1.0x10^{-6}$. Numbers in parentheses are PubMed IDs. SNP, single nucleotide polymorphism; Chr., chromosome; LDL, low density lipoprotein.

Table X. Relationship of genes and SNPs associated with intracerebral hemorrhage or subarachnoid hemorrhage in the present study to previously reported cerebrovascular disease-related phenotypes.

Gene/chr. locus	SNP	Chr.	Position	Previously reported phenotypes
Intracerebral hemorrhage				
DNTTIP2	rs3747965	1	93877008	None
FAM205A	rs3739881	9	34724244	None
SVEP1	rs7030192	9	110407351	Coronary artery disease (23364394)
12q24.1	rs12229654	12	110976657	HDL-cholesterol (21909109)
ACAD10	rs11066015	12	111730205	Coronary artery disease (23364394, 23202125), LDL-cholesterol (20686565), type 1 diabetes (17554300), diastolic blood pressure (21909115)
ALDH2	rs671	12	111803962	HDL-cholesterol (21572416, 21372407), myocardial infarction (21971053), coronary artery disease (21971053, 21572416, 23202125), diastolic blood pressure (21572416, 21909115), systolic blood pressure (21572416), LDL-cholesterol (21572416, 20686565), type 1 diabetes (17554300)
Subarachnoid hemorrhage				
FAM160A1	rs2709828	4	151434116	None
OR52E4	rs11823828	11	5884973	None

Data were obtained from genome-wide repository of associations between SNPs and phenotypes (GRASP) search database (https://grasp.nhlbi.nih.gov/Search.aspx) with a P-value of $<1.0x10^{-6}$. Numbers in parentheses are PubMed IDs. SNP, single nucleotide polymorphism; Chr., chromosome; HDL, high density lipoprotein; LDL, low density lipoprotein.

uted to acquisition of the data and to revision of the manuscript. YYasukochi, IT and JS contributed to analysis and interpretation of the data as well as to revision of the manuscript.

Ethics approval and consent to participate

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, Hirosaki University Graduate School of Medicine, and participating hospitals (Gifu Prefectural Tajimi Hospital, Gifu Prefectural General Medical Center, Japanese Red Cross Nagoya First Hospital, Northern Mie Medical Center Inabe General Hospital, and Hirosaki Stroke and Rehabilitation Center). Written informed consent was obtained from all subjects.

Consent for publication

All authors approved submission of the final version of the article for publication.

Competing interests

The authors declare that they have no competing interests.

References

1. Feigin VL, Forouzanfar MH, Krishnamurthi R, Mensah GA, Connor M, Bennett DA, Moran AE, Sacco RL, Anderson L, Truelsen T, *et al*; Global Burden of Diseases, Injuries, and Risk Factors Study 2010 (GBD 2010) and the GBD Stroke Experts Group: Global and regional burden of stroke during 1990-2010: Findings from the Global Burden of Disease Study 2010. Lancet 383: 245-254, 2014.
2. Benjamin EJ, Virani SS, Callaway CW, Chang AR, Cheng S, Chiuve SE, Cushman M, Delling FN, Deo R, de Ferranti SD *et al*; American Heart Association Council on Epidemiology and Prevention Statistics Committee and Stroke Statistics Subcommittee: Heart disease and stroke statistics-2018 update: A report from the American Heart Association. Circulation 137: e67-e492, 2018.
3. Tournier-Lasserve E: New players in the genetics of stroke. N Engl J Med 347: 1711-1712, 2002.
4. Bak S, Gaist D, Sindrup SH, Skytthe A and Christensen K: Genetic liability in stroke: A long-term follow-up study of Danish twins. Stroke 33: 769-774, 2002.
5. Adams HP Jr, Bendixen BH, Kappelle LJ, Biller J, Love BB, Gordon DL and Marsh EE III: Classification of subtype of acute ischemic stroke. Definitions for use in a multicenter clinical trial. TOAST. Trial of Org 10172 in Acute Stroke Treatment. Stroke 24: 35-41, 1993.
6. Jerrard-Dunne P, Cloud G, Hassan A and Markus HS: Evaluating the genetic component of ischemic stroke subtypes: A family history study. Stroke 34: 1364-1369, 2003.
7. 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.
8. Falcone GJ, Malik R, Dichgans M and Rosand J: Current concepts and clinical applications of stroke genetics. Lancet Neurol 13: 405-418, 2014.
9. Malik R and Dichgans M: Challenges and opportunities in stroke genetics. Cardiovasc Res, 14 Mar, 2018 (Epub ahead of print).
10. 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.
11. 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.
12. 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.
13. 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.
14. 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.
15. 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.
16. 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.
17. Chauhan G, Arnold CR, Chu AY, Fornage M, Reyahi A, Bis JC, Havulinna AS, Sargurupremraj M, Smith AV, Adams HHH, *et al*; Neurology Working Group of the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium, the Stroke Genetics Network (SiGN), and the 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.
18. Pulit SL, McArdle PF, Wong Q, Malik R, Gwinn K, Achterberg S, Algra A, Amouyel P, Anderson CD, Arnett DK, *et al*; NINDS Stroke Genetics Network (SiGN); 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.
19. Malik R, Chauhan G, Traylor M, Sargurupremraj M, Okada Y, Mishra A, Rutten-Jacobs L, Giese AK, van der Laan SW, Gretarsdottir S, *et al*; AFGen Consortium; Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium; International Genomics of Blood Pressure (iGEN-BP) Consortium; INVENT Consortium; STARNET; BioBank Japan Cooperative Hospital Group; COMPASS Consortium; EPIC-CVD Consortium; EPIC-InterAct Consortium; International Stroke Genetics Consortium (ISGC); METASTROKE Consortium; Neurology Working Group of the CHARGE Consortium; NINDS Stroke Genetics Network (SiGN); UK Young Lacunar DNA Study; MEGASTROKE Consortium; MEGASTROKE Consortium: Multiancestry genome-wide association study of 520,000 subjects identifies 32 loci associated with stroke and stroke subtypes. Nat Genet 50: 524-537, 2018.
20. Qureshi AI, Tuhrim S, Broderick JP, Batjer HH, Hondo H and Hanley DF: Spontaneous intracerebral hemorrhage. N Engl J Med 344: 1450-1460, 2001.
21. Alberts MJ, McCarron MO, Hoffmann KL and Graffagnino C: Familial clustering of intracerebral hemorrhage: A prospective study in North Carolina. Neuroepidemiology 21: 18-21, 2002.
22. 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.
23. 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.

24. 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.
25. 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.
26. Jeanne M, Labelle-Dumais C, Jorgensen J, Kauffman WB, Mancini GM, Favor J, Valant V, Greenberg SM, Rosand J and Gould DB: COL4A2 mutations impair COL4A1 and COL4A2 secretion and cause hemorrhagic stroke. Am J Hum Genet 90: 91-101, 2012.
27. Rannikmäe K, Davies G, Thomson PA, Bevan S, Devan WJ, Falcone GJ, Traylor M, Anderson CD, Battey TW, Radmanesh F, *et al*; METASTROKE Consortium; CHARGE WMH Group; ISGC ICH GWAS Study Collaboration; WMH in Ischemic Stroke GWAS Study Collaboration; International Stroke Genetics Consortium: Common variation in COL4A1/COL4A2 is associated with sporadic cerebral small vessel disease. Neurology 84: 918-926, 2015.
28. 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.
29. Macdonald RL and Schweizer TA: Spontaneous subarachnoid haemorrhage. Lancet 389: 655-666, 2017.
30. Okamoto K, Horisawa R, Kawamura T, Asai A, Ogino M, Takagi T and Ohno Y: Family history and risk of subarachnoid hemorrhage: A case-control study in Nagoya, Japan. Stroke 34: 422-426, 2003.
31. Bor AS, Rinkel GJ, Adami J, Koffijberg H, Ekbom A, Buskens E, Blomqvist P and Granath F: Risk of subarachnoid haemorrhage according to number of affected relatives: A population based case-control study. Brain 131: 2662-2665, 2008.
32. Tromp G, Weinsheimer S, Ronkainen A and Kuivaniemi H: Molecular basis and genetic predisposition to intracranial aneurysm. Ann Med 46: 597-606, 2014.
33. Korja M, Silventoinen K, McCarron P, Zdravkovic S, Skytthe A, Haapanen A, de Faire U, Pedersen NL, Christensen K, Koskenvuo M, *et al*; GenomEUtwin Project: Genetic epidemiology of spontaneous subarachnoid hemorrhage: Nordic Twin Study. Stroke 41: 2458-2462, 2010.
34. 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.
35. 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.
36. 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.
37. 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.
38. van 't Hof FN, Ruigrok YM, Lee CH, Ripke S, Anderson G, de Andrade M, Baas AF, Blankensteijn JD, Böttinger EP, Bown MJ, *et al*; Aneurysm Consortium; Vascular Research Consortium of New Zealand: Shared genetic risk factors of intracranial, abdominal, and thoracic aneurysms. J Am Heart Assoc 5: e002603, 2016.
39. 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.
40. 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.
41. 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.
42. 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.
43. Cheng YC, Cole JW, Kittner SJ and Mitchell BD: Genetics of ischemic stroke in young adults. Circ Cardiovasc Genet 7: 383-392, 2014.
44. Brass LM, Isaacsohn JL, Merikangas KR and Robinette CD: A study of twins and stroke. Stroke 23: 221-223, 1992.
45. Woo D, Sekar P, Chakraborty R, Haverbusch MA, Flaherty ML, Kissela BM, Kleindorfer D, Schneider A, Khoury J, Sauerbeck LR, *et al*: Genetic epidemiology of intracerebral hemorrhage. J Stroke Cerebrovasc Dis 14: 239-243, 2005.
46. Sundquist J, Li X, Sundquist K and Hemminki K: Risks of subarachnoid hemorrhage in siblings: A nationwide epidemiological study from Sweden. Neuroepidemiology 29: 178-184, 2007.
47. Yamada Y, Sakuma J, Takeuchi I, Yasukochi Y, Kato K, Oguri M, Fujimaki T, Horibe H, Muramatsu M, Sawabe M, *et al*: Identification of six polymorphisms as novel susceptibility loci for ischemic or hemorrhagic stroke by exome-wide association studies. Int J Mol Med 39: 1477-1491, 2017.
48. Yamada Y, Matsui K, Takeuchi I, Oguri M and Fujimaki T: Association of genetic variants with hypertension in a longitudinal population-based genetic epidemiological study. Int J Mol Med 35: 1189-1198, 2015.
49. 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.
50. 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.
51. 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.
52. 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.
53. Benjamini Y and Hochberg Y: Controlling the false discovery rate: A practical and powerful approach to multiple testing. J R Stat Soc B 57: 289-300, 1995.
54. 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.
55. Hakonarson H, Grant SF, Bradfield JP, Marchand L, Kim CE, Glessner JT, Grabs R, Casalunovo T, Taback SP, Frackelton EC, *et al*: A genome-wide association study identifies KIAA0350 as a type 1 diabetes gene. Nature 448: 591-594, 2007.
56. Bradfield JP, Qu HQ, Wang K, Zhang H, Sleiman PM, Kim CE, Mentch FD, Qiu H, Glessner JT, Thomas KA, *et al*: A genome-wide meta-analysis of six type 1 diabetes cohorts identifies multiple associated loci. PLoS Genet 7: e1002293, 2011.
57. Teslovich TM, Musunuru K, Smith AV, Edmondson AC, Stylianou IM, Koseki M, Pirruccello JP, Ripatti S, Chasman DI, Willer CJ, *et al*: Biological, clinical and population relevance of 95 loci for blood lipids. Nature 466: 707-713, 2010.
58. Wain LV, Verwoert GC, O'Reilly PF, Shi G, Johnson T, Johnson AD, Bochud M, Rice KM, Henneman P, Smith AV, *et al*; LifeLines Cohort Study; EchoGen consortium; AortaGen Consortium; CHARGE Consortium Heart Failure Working Group; KidneyGen consortium; CKDGen consortium; Cardiogenics consortium; CardioGram: Genome-wide association study identifies six new loci influencing pulse pressure and mean arterial pressure. Nat Genet 43: 1005-1011, 2011.

59. Ehret GB, Munroe PB, Rice KM, Bochud M, Johnson AD, Chasman DI, Smith AV, Tobin MD, Verwoert GC, Hwang SJ, *et al*; CHARGE-HF consortium: Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. Nature 478: 103-109, 2011.
60. Takeuchi F, Yokota M, Yamamoto K, Nakashima E, Katsuya T, Asano H, Isono M, Nabika T, Sugiyama T, Fujioka A, *et al*: Genome-wide association study of coronary artery disease in the Japanese. Eur J Hum Genet 20: 333-340, 2012.
61. Lee JY, Lee BS, Shin DJ, Woo Park K, Shin YA, Joong Kim K, Heo L, Young Lee J, Kyoung Kim Y, Jin Kim Y, *et al*: A genome-wide association study of a coronary artery disease risk variant. J Hum Genet 58: 120-126, 2013.
62. CARDIoGRAMplusC4D Consortium, Deloukas P, Kanoni S, Willenborg C, Farrall M, Assimes TL, Thompson JR, Ingelsson E, Saleheen D, Erdmann J, *et al*: Large-scale association analysis identifies new risk loci for coronary artery disease. Nat Genet 45: 25-33, 2013.
63. Ehret GB, Munroe PB, Rice KM, Bochud M, Johnson AD, Chasman DI, Smith AV, Tobin MD, Verwoert GC, Hwang SJ, *et al*; CHARGE-HF consortium: Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. Nature 478: 103-109, 2011.
64. Kato N, Takeuchi F, Tabara Y, Kelly TN, Go MJ, Sim X, Tay WT, Chen CH, Zhang Y, Yamamoto K, *et al*: Meta-analysis of genome-wide association studies identifies common variants associated with blood pressure variation in east Asians. Nat Genet 43: 531-538, 2011.
65. Kim YJ, Go MJ, Hu C, Hong CB, Kim YK, Lee JY, Hwang JY, Oh JH, Kim DJ, Kim NH, *et al*; MAGIC consortium: Large-scale genome-wide association studies in East Asians identify new genetic loci influencing metabolic traits. Nat Genet 43: 990-995, 2011.