Functional SNP of ARHGEF10 confers risk of atherothrombotic stroke

Tomonaga Matsushita1,2, Kyota Ashikawa1, Koji Yonemoto3, Yoichiro Hirakawa3, Jun Hata3, Hanae Amitani1, Yasufumi Doi3, Toshiharu Ninomiya3, Takanari Kitazono2, Setsuro Ibayashi2, Mitsuo Iida2, Yusuke Nakamura4, Yutaka Kiyohara3 and Michiaki Kubo1,2,3,*

1Laboratory for Genotyping Development, Center for Genomic Medicine, RIKEN, Yokohama, Kanagawa, Japan, 2Department of Medicine and Clinical Science and 3Department of Environmental Medicine, Graduate School of Medical Sciences, Kyushu University, Fukuoka, Japan and 4Laboratory for Molecular Medicine, Human Genome Center, The Institute of Medical Science, University of Tokyo, Tokyo, Japan

Received September 5, 2009; Revised and Accepted December 23, 2009

Although stroke is a common cause of death and a major cause of disability all over the world, genetic components of common forms of ischemic stroke are largely unknown. To identify susceptibility genes of atherothrombotic stroke, we performed a large case–control association study and a replication study in a total of 2775 cases with atherothrombotic stroke and 2839 controls. Through the analysis in 860 cases and 860 age- and sex-matched controls, we found that a single-nucleotide polymorphism (SNP), rs2280887, in the ARHGEF10 gene was significantly associated with atherothrombotic stroke even after the adjustment of multiple testing by a permutation test [unadjusted \( P = 1.2 \times 10^{-6}, \ odds\ ratio = 1.80, 95\% confidence interval (CI) = 1.42–2.28\]. This association was replicated in independent 1915 cases and 1979 controls. Subsequent fine mapping found another three SNPs which showed similar association due to strong linkage disequilibrium to rs2280887 \( (r^2 > 0.95)\). In the functional analyses of these four highly associated SNPs, using luciferase assay and electrophoretic mobility shift assay we found that rs4376531 affected ARHGEF10 transcriptional activity due to the different Sp1-binding affinity. In small GTPase activity assay, we found that a gene product of ARHGEF10 specifically activated RhoA. A population-based cohort study revealed the subjects with rs4376531 CC or CG to increase the incidence of ischemic stroke \( (P = 0.033, \ hazard\ ratio = 1.79, 95\% CI = 1.05–3.04)\). Our data suggest that the functional SNP of ARHGEF10 confers the susceptibility to atherothrombotic stroke.

INTRODUCTION

Stroke is a common cause of death and a major cause of disability all over the world (1). Particularly, in the countries having a larger proportion of older people, the burden has been increasing more significantly. Twin and family studies have indicated that the risk for ischemic stroke is related to multiple genetic and environmental factors (2). Identification of susceptibility genes for ischemic stroke is expected to elucidate new pathophysiological mechanisms of the disease and to lead to the development of novel preventive measures. Although previous genome-wide association studies (GWASs) reported several susceptibility genes (3,4), genetic components of common forms of ischemic stroke are still largely undetermined.

Ischemic stroke is usually classified into several subtypes. From the aspects of pathophysiological mechanisms and preventive measures, ischemic stroke can be classified into atherothrombotic stroke and cardioembolic stroke (3,5,6). Atherothrombotic stroke is mainly caused by atherosclerosis in arteries of various sizes, and the main preventive measure is the control of cardiovascular risk factors such as hypertension, diabetes, dyslipidemia and smoking. In contrast, cardioembolic stroke is mainly caused by cardiac diseases such as atrial fibrillation and valvular heart disease.
disease, and the main preventive measure is the use of anticoagulants.

As for the genetics of these two subtypes, Jerrard-Dunne et al. showed that the genetic component of atherothrombotic stroke is stronger than that of cardioembolic subtype, and suggested that the genetic study might be more efficient by focusing on atherothrombotic stroke (7). Therefore, we performed a large-scale case–control association study in a Japanese population (8,9) by focusing on atherothrombotic stroke and identified a gene encoding guanine nucleotide exchange factor 10 (ARHGEF10) on chromosome 8p23 as a new susceptibility gene for atherothrombotic stroke. We found that a functional single-nucleotide polymorphism (SNP) of this gene affects its transcriptional activity by altering Sp1-binding affinity. Furthermore, small GTPase activity assay showed that a gene product of ARHGEF10 specifically activates RhoA. Since RhoA–Rho kinase pathway has an important role for the pathogenesis of cardiovascular disease and atherosclerosis, the functional SNP of ARHGEF10 may be involved in the susceptibility to the development of ischemic stroke.

RESULTS

Case–control association study

We previously performed a two-stage association analysis using 1112 cases with ischemic stroke and 1112 age- and sex-matched controls by examining 52,608 gene-based tag-SNPs selected from the JSNP database (8,9). To identify SNPs possibly associated with atherothrombotic stroke, we further analyzed candidate SNPs by focusing on 860 cases with atherothrombotic stroke and 860 age- and sex-matched controls (Set 1). We found that an SNP rs2280887 in intron 17 of ARHGEF10 on chromosome 8p23 was found to be strongly associated with atherothrombotic stroke (P = 1.2 × 10^-6 for dominant model; Table 1). This association remained significant after a permutation test for the adjustment of multiple testing (P = 0.0006). Although this SNP revealed a weak association even when we analyzed all ischemic stroke cases, no association was observed in the case of cardioembolic stroke (Supplementary Material, Table S1).

We subsequently selected and genotyped 93 tag-SNPs across the ARHGEF10 gene selected from phase II of HapMap JPT data. We found an SNP rs4480162 in intron 17 of ARHGEF10 to be in absolute linkage disequilibrium (LD) with rs2280887 (D' = 1.0 and r² = 1.0) and be significantly associated with atherothrombotic stroke (P = 6.9 × 10^-7 for dominant model; Table 1). LD analysis showed that these two SNPs, rs4480162 and rs2280887, were located in a small LD block (block A) spanning 15.7 kb region, which corresponds to a region from intron 15 to intron 18 of ARHGEF10 (Supplementary Material, Fig. S1). None of the remaining 92 tag-SNPs showed significant association with the disease. We subsequently searched for variants in this 15.7 kb region by direct sequencing using 48 affected individuals. This resequencing identified a total of 81 variants, of which 50 variants were already registered in dbSNP database, and 31 variants were new. After the exclusion of the variants genotyped or with minor allele frequency (MAF) of <0.05, we genotyped 43 additional SNPs. Figure 1 shows the result of fine mapping around the candidate region of ARHGEF10. The association of the SNPs with the disease was limited to the block A region of ARHGEF10. In block A, additional two SNPs, rs35234164 and rs4376531, were found to have significant associations similar to rs2280887 (Fig. 1C and Table 1). rs4376531 was only two bases apart from rs4480162, and these SNPs were absolutely linked with rs2280887. rs35234164 was a one-base insertion/deletion (T/del) polymorphism located at intron 16 and strongly linked with other three SNPs (each pairwise D' = 1.0 and r² = 0.95). No other SNP in block A was associated with atherothrombotic stroke. These four SNPs were found to be associated with atherothrombotic stroke in another case–control set of 1915 cases and 1979 controls (Set 2, P = 0.010 for dominant model; Table 1). The results of association analyses of those SNPs under allele and recessive models are shown in Supplementary Material, Table S2.

Susceptible allele of rs4376531 affects ARHGEF10 transcriptional activity through the difference in Sp1-binding affinity

The four SNPs were located at intron 16 or 17 and mapped ~80 kb apart from 5'-untranslated region (UTR) and 50 kb apart from 3'-UTR. None of the four SNPs was located in splice donor, acceptor or branch sites of intron 16 or 17. Furthermore, the UCSC Genome Browser database indicated no additional annotated gene or non-coding RNA in the block A region. We hypothesized that some of these SNPs might exert some effect on transcription and prepared 5'-end biotin-labeled oligonucleotide probes that were derived from the genomic sequences corresponding to these SNPs. Although rs4480162 and rs4376531 were absolutely linked and only haplotypes of C-G and G-C were existed in our population, we also synthesized rs4480162_C/rs4376531_G and rs4480162_G/rs4376531_C probes to elucidate the function of each SNP. Electrophoretic mobility shift assay (EMSA) experiments using these oligonucleotides with nuclear extract of LoVo cells, in which the expression of ARHGEF10 gene transcript is high, found a shifted band of a DNA–protein complex with a strong intensity in lanes corresponding to the susceptible allele of rs4376531 (C-G and G-G, Fig. 2A). This shifted band was weak in lanes corresponding to the non-susceptible allele (C-C and G-C). Although we also observed shifted bands for other oligonucleotides, no difference in the intensity between susceptible and non-susceptible alleles was observed. The competition assay with the unlabeled oligonucleotides demonstrated that the oligonucleotides containing the susceptibility allele of rs4376531 (C-G and G-G) inhibited the formation of DNA–protein complex in a dose-dependent manner but the other oligonucleotides (C-C and G-C) did not (Fig. 2B), suggesting that some nuclear proteins specifically bound to the DNA fragment corresponding to the susceptible allele of rs4376531. To identify which transcriptional factor binds to this susceptible allele, we added excess amount of unlabeled oligonucleotides corresponding to consensus sequences of various transcriptional factors as competitor and found that the unlabeled Sp1-binding consensus oligonucleotide effectively inhibited the formation of the DNA–protein complex (Fig. 2C).
Table 1. Association results among the four SNPs in ARHGEF10 for atherothrombotic stroke

<table>
<thead>
<tr>
<th>SNP (allele)</th>
<th>Set</th>
<th>Case</th>
<th>Control</th>
<th>Dominant model</th>
<th>Additive model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>11</td>
<td>12</td>
<td>22  Total P-value</td>
<td>OR 95% CI P-value</td>
</tr>
<tr>
<td>rs2280887 (G/C)</td>
<td>Set 1</td>
<td>9</td>
<td>211</td>
<td>638 858 15 123 719 857 1.2 × 10⁻⁶</td>
<td>1.80 1.42–2.28 5.2 × 10⁻⁵</td>
</tr>
<tr>
<td></td>
<td></td>
<td>22</td>
<td>24</td>
<td>389 1501 1914 23 341 1615 1979 2.1 × 10⁻⁶</td>
<td>1.38 1.21–1.57 2.8 × 10⁻⁵</td>
</tr>
<tr>
<td>rs35234164 (-/T)</td>
<td>Set 1</td>
<td>8</td>
<td>206</td>
<td>642 856 14 116 729 859 3.4 × 10⁻⁷</td>
<td>1.87 1.47–2.38 1.9 × 10⁻⁵</td>
</tr>
<tr>
<td></td>
<td></td>
<td>22</td>
<td>277</td>
<td>1507 1906 23 326 1627 1976 7.0 × 10⁻⁷</td>
<td>1.40 1.23–1.60 1.5 × 10⁻⁵</td>
</tr>
<tr>
<td>rs4480162 (C/G)</td>
<td>Set 1</td>
<td>9</td>
<td>212</td>
<td>636 857 15 123 722 860 6.9 × 10⁻⁷</td>
<td>1.82 1.43–2.31 3.4 × 10⁻⁵</td>
</tr>
<tr>
<td></td>
<td></td>
<td>22</td>
<td>23</td>
<td>390 1502 1915 23 342 1613 1978 0.015</td>
<td>1.22 1.04–1.42 0.024</td>
</tr>
<tr>
<td>rs4376531 (G/C)</td>
<td>Set 1</td>
<td>9</td>
<td>212</td>
<td>636 857 15 123 722 860 6.9 × 10⁻⁷</td>
<td>1.82 1.43–2.31 3.4 × 10⁻⁵</td>
</tr>
<tr>
<td></td>
<td></td>
<td>22</td>
<td>23</td>
<td>390 1499 1912 23 344 1610 1977 0.018</td>
<td>1.21 1.03–1.41 0.028</td>
</tr>
<tr>
<td></td>
<td>Combined</td>
<td></td>
<td></td>
<td>2.6 × 10⁻⁶</td>
<td>1.37 1.20–1.56 3.8 × 10⁻⁵</td>
</tr>
</tbody>
</table>

For additive genetic model, we used logistic regression analysis coding genotypes 11, 12 and 22 as 2, 1 and 0, respectively. OR, odds ratio; CI, confidence interval.

Moreover, when we added anti-Sp1 antibody to the mixture, the band was further shifted to a higher molecular position, indicating the specific binding of the Sp1 protein to the susceptible allele of rs4376531. Similar shifted band was observed when we used rs4480162_G/rs4376531_G probe (data not shown).

To test whether rs4376531 affects the ARHGEF10 transcriptional activity, we performed a luciferase assay using LoVo cells. We subcloned the sequences corresponding C-G and G-C EMSA probes into the pGL3-promoter vector. Luciferase activity was enhanced in the cells transfected with the reporter vector containing the susceptible allele of rs4376531 but the enhancement was low in the cells transfected with the vector containing the non-susceptible allele (Fig. 2D). These findings indicated that rs4376531 might affect ARHGEF10 transcriptional activity through the difference in the binding affinity of Sp1 transcriptional factor.

Although the function of ARHGEF10 is not well understood, ARHGEF10 is a member of the family of guanine nucleotide exchange factors (GEFs), which regulate the activity of small Rho GTPases by catalyzing the exchange of bound GDP by GTP. To elucidate the role of ARHGEF10 in the pathogenesis of atherothrombotic stroke, we examined the effect of ARHGEF10 on the activation of RhoA, Rac1 and Cdc42 by small GTPase activity assay. As shown in Figure 3, overexpression of ARHGEF10 led to an increase in the GTP-bounded RhoA, indicating that ARHGEF10 might activate RhoA. In contrast, overexpression of ARHGEF10 had no effect on the GTP-bounded Rac1 or Cdc42. Since Sp1 is abundantly expressed in multiple tissues, the subjects with the disease-susceptible allele of rs4376531 are expected to have higher expression of ARHGEF10 transcripts and might result in the higher activity of RhoA–Rho kinase pathway.

rs4376531 increases the incidence of ischemic stroke

Finally, we examined the effect of rs4376531 on the incidence of ischemic stroke using a population-based cohort study. During a 14-year follow-up of the cohort, 67 events of first-ever ischemic stroke were observed among 1656 subjects without a history of stroke at baseline examination. Figure 4 shows Kaplan–Meier estimates of the incidence of ischemic stroke by rs4376531. The cumulative incidence was 6.1% in the subjects who had at least one susceptible allele and 3.6% in the subjects with the homozygous of non-susceptible allele (P = 0.042 for log-rank test). Age- and sex-adjusted risk of atherothrombotic stroke was significantly higher in the subjects with susceptible allele of rs4376531 (adjusted P = 0.033, hazard ratio = 1.79, 95% confidence interval = 1.05–3.04).

**DISCUSSION**

In this study, we analyzed the data of a large-scale case–control association study by focusing on atherothrombotic stroke. We found a new candidate locus, rs2280887, located in intron 17 of ARHGEF10. This SNP was significantly associated with atherothrombotic stroke even after the adjustment of multiple testing, and the association was replicated in other case–control samples. Fine mapping of the ARHGEF10 gene identified four highly linked SNPs (rs2280887, rs35234164, rs4480162 and rs4376531) as candidates with functional significance. Functional analysis of these four SNPs demonstrated that an SNP, rs4376531, altered the binding affinity of the Sp1 transcriptional factor and might enhance the ARHGEF10 transcriptional activity in individuals with the susceptible allele. We also found that ARHGEF10 specifically activated RhoA, which has an important role in various process of atherosclerosis. From these findings, we suggest that the subjects with the susceptible allele of rs4376531 in ARHGEF10 will have higher expression of transcript and might have higher RhoA activity. Since RhoA–Rho kinase pathway is involved in the pathogenesis of atherosclerosis, the functional SNP of ARHGEF10 might confer the development of atherothrombotic stroke. A population-based cohort study supported this hypothesis.

Although several GWASs for ischemic stroke have been reported (3,4), they did not detect the association of SNPs in ARHGEF10 with ischemic stroke. Among the four highly associated SNPs in this study, rs4480162 and rs4376531 were registered in the HapMap database. From the database,
MAF of rs4480162 was 0.167 for CEU, 0.475 for YRI, 0.078 for CHB and 0.091 for JPT. Our sequence data showed that rs4376531 is absolutely linked with rs4480162 in the Japanese population; however, MAF of rs4376531 was 0 for CEU, 0.021 for YRI, 0 for CHB and 0 for JPT. These data speculate that the LD between causative variant (rs4376531) and other three SNPs might be different among different populations if the genotype data of CEU and YRI are correct. Moreover, our large-scale association study included rs2280887 as one of the 52,608 gene-based tag-SNPs selected from JSNP database. However, the current GWAS platforms do not contain all of the four SNPs even in Affymetrix Genome-Wide Human SNP Array 6.0 or Illumina Human1M-Duo BeadChip. Therefore, current GWASs using commercial chips cannot detect the association of SNPs in ARHGEF10 and ischemic stroke.

GEFs activate small GTPases in response to diverse extracellular stimuli and ultimately regulate numerous cellular responses (10). Small GTPases, which were identified as the master regulators of the actin cytoskeleton, control a...
Figure 2. rs4376531 alters the binding affinity of Sp1 and affects ARHGEF10 transcriptional activity. (A) EMSA using 5’ end-labeled 50 bp probes around each allele of SNPs in ARHGEF10. A black arrow indicates the shifted band that shows tighter binding of a nuclear protein to the susceptible allele of rs4376531 (C-G and G-G) than the non-susceptible allele (G-C and C-C). (B) Competition assay with unlabeled self- or non-self-oligonucleotides. DNA–protein complex (black arrow) was more effectively competed by unlabeled oligonucleotides with G-allele of rs4376531 than those with C-allele. (C) Competition assay with unlabeled Sp1-binding consensus oligonucleotide and supershift assay using anti-Sp1 antibody. We could observe additional shifted band more clearly (white arrow) by a longer electrophoresis run time. (D) Luciferase assay. Fifty base pair fragments around each allele of rs4376531 were inserted into pGL3-promoter vector. Each sample was studied in triplicate and data were shown as mean ± SD. Asterisk indicates $P < 0.05$ by Student’s $t$-test.
remarkable diversity of cell functions including contraction, motility, proliferation and apoptosis. Small GTPases act as molecular switches, which cycle between an inactive GDP-bound form and an active GTP-bound form (11). The Rho GEFs mediate the activation of small GTPases by promoting the release of GDP in exchange for GTP (12). ARHGEF10, a member of Rho GEFs, was identified by the sequencing of cDNA clones from the human brain (13) and found to be expressed not only in the brain but also in various tissues including the heart (14). A point mutation (T109I) of ARHGEF10 was reported to co-segregate in the family with slowed motor and sensory nerve conduction velocities of peripheral nerves with autosomal dominant inheritance (15). However, the function of ARHGEF10 is largely unknown. We found that ARHGEF10 specifically activated RhoA and might contribute to the development of atherothrombotic stroke through the regulation of RhoA–Rho kinase activity. Recent linkage studies have demonstrated that several Rho GEFs might be involved in the pathogenesis of atherosclerosis. A linkage scan for type 2 diabetes has identified that non-synonymous SNPs in LARG and PDZ-Rho GEF are associated with insulin sensitivity or insulin resistance (16,17). Another linkage study has found kalirin gene as a candidate gene for early-onset coronary artery disease (18). These results suggest that polymorphisms of Rho GEFs affect small GTPase signaling pathway and result in the pathogenesis of human atherosclerosis. RhoA, which is the most characterized small GTPases (19), and one of its effectors, Rho-kinase, were reported to play an important role in the various processes of atherosclerosis including endothelial dysfunction, inflammation and vascular smooth muscle cell proliferation (20–24). From these aspects, drugs that inhibit RhoA–Rho kinase pathway such as Rho-kinase inhibitor (20) or statin (24) are already available in clinical setting. If our findings are confirmed in further studies, we could expect to use RhoA–Rho kinase pathway inhibitors in the subjects with susceptible allele of ARHGEF10 SNP for the more effective prevention of atherothrombotic stroke.

In conclusion, rs4376531 located in intron 17 of ARHGEF10 was significantly associated with atherothrombotic stroke. Individuals with the susceptible allele of the SNP might have a higher level of ARHGEF10 transcript due to the higher binding-affinity of Sp1 and might have higher RhoA–Rho kinase activity. This higher activity will finally result in the increased incidence of atherothrombotic stroke.

Figure 3. ARHGEF10 activates RhoA specifically. (A) RhoA activity assay. 293FT cells were co-transfected with the plasmids as indicated at the top. Cell lysates were pulled down by GST-Rhotekin and subjected to immunoblot with anti-Myc antibody. Whole-cell lysates were analyzed by immunoblot with anti-Myc and anti-FLAG antibody to detect total RhoA and ARHGEF10, respectively. GTP-bound RhoA and total RhoA were quantified by the intensity of the bands using Multi Gauge software of an LAS-3000 system (bottom). RhoA activity was calculated as the relative ratio of the intensity of GTP-bound RhoA against that of total RhoA. Relative intensity in the control (transfected empty vector) was expressed as 1 arbitrary unit. (B) Rac1 activity assay. (C) Cdc42 activity assay. (B) and (C) were analyzed as in (A) by pull-down assay using GST-PAK-1 instead of GST-Rhotekin. The experiments were repeated at least three times.
in the general population. Our findings might shed light on the elucidation of new atherosclerotic pathogenesis and to the development of preventive therapy for ischemic stroke.

MATERIALS AND METHODS

The flow chart of this study is shown in Supplementary Material, Figure S2.

Study populations

For the large-scale case–control association study, cases with ischemic stroke were registered from seven medical centers in and around Fukuoka City, Japan, in 2004. Details of registration were described previously (8). Briefly, all case subjects were diagnosed by stroke neurologists on the basis of detailed clinical features and ancillary laboratory examinations [such as brain imaging including computed tomography (CT) and magnetic resonance imaging (MRI), cerebral angiography, echocardiography and carotid duplex imaging]. Ischemic stroke was defined as a sudden onset of non-convulsive and focal neurological deficit persisting for >24 h without evidence of hemorrhagic stroke on brain imaging (CT or MRI). Ischemic stroke was further subdivided into atherothrombotic stroke, cardioembolic stroke and undetermined subtype. Subtypes of ischemic stroke were determined on the basis of the Classification of Cerebrovascular Disease III proposed by the National Institute of Neurological Disorders and Stroke (25), as well as on the basis of the diagnostic criteria of the Trial of Org 10172 in Acute Stroke Treatment (TOAST) study (26) and Cerebral Embolism Task Force (27). Details of the diagnostic criteria of ischemic stroke subtypes have been described previously (8). Briefly, small-artery occlusion (lacunar stroke) was diagnosed as the presence of a relevant brain stem or subcortical hemispheric lesion with a diameter of <1.5 cm demonstrated on brain imaging and no evidence of cerebral cortical or cerebellar impairment. Large artery atherosclerotic stroke was diagnosed when the subjects had significant stenosis (>50%) or occlusion of a major cerebral artery with infarct size ≥1.5 cm on brain imaging. The diagnosis of cardioembolic stroke was made on the basis of primary and secondary clinical features suggestive of cardioembolic stroke as reported by the Cerebral Embolism Task Force (27). The category of undetermined subtype included all ischemic stroke cases for which the subtype could not be determined because of insufficient clinical or morphological information. Both small-artery occlusion and large artery atherosclerotic stroke were included in the phenotype of atherothrombotic stroke. Subtypes of ischemic stroke were 860 in atherothrombotic, 136 in cardioembolic and 116 in undetermined subtype. Age- (within 5 years) and sex-matched control subjects were selected from the 3328 participants of the Hisayama screening survey between 2002 and 2003.

For the replication study, case samples were selected from the BioBank Japan project (28). Among the subjects with ischemic stroke in the BioBank Japan, we selected 1915 cases that were diagnosed as atherothrombotic stroke by brain imaging, same as the initial study. The remaining 1979 Hisayama participants who were not enrolled in the initial study were used as controls. Clinical characteristics of the study population in the two case–control sets were shown in Supplementary Material, Table S3.

For the prospective cohort study, we used a cohort population of the Hisayama study established in 1988 (8). In this cohort, 2637 Hisayama residents aged ≥40 years without a history of stroke or coronary heart disease were enrolled in 1988 and continuously followed up for 14 years until the occurrence of cardiovascular diseases or death. Among them, 1656 subjects participated in the examination between 2002 and 2003 were used in the present study.

Genomic DNA was extracted from peripheral blood leukocytes by a standard method in both populations. Written informed consent was obtained from all study subjects in both populations, and this study was approved by the ethics committees of the Graduate School of Medical Sciences, Kyushu University, and Yokohama Institute, RIKEN.

SNP selection and genotyping

In the previous large-scale case–control association study, we used a two-stage approach to identify susceptibility genes of ischemic stroke. We first genotyped 52 608 gene-based tag-SNPs selected from JSNP database using 188 cases and 188 age- and sex-matched controls. In the second stage, 1098 SNPs that showed $P < 0.01$ in the first stage were genotyped in the remaining samples. Details of this large-scale association study were described previously (8). For this study, we combined the data of the first and the second stage and re-analyzed by focusing on atherothrombotic stroke using the matched case–control samples. All SNPs in the large-scale association study were genotyped using the multiplex PCR-based Invader assay (Third Wave Technologies) described previously (29). All genotypes were called by visual inspection, and we determined genotyped success as less than 10 undetermined samples in a 384-well-plate.

For fine mapping across ARHGEF10, we selected tag-SNPs from phase II of the HapMap JPT data by pairwise tagging method with the following criteria: $r^2 > 0.8$, MAF > 5% and call rate > 75%. We genotyped SNPs using the multiplex PCR-based Invader assay or by direct sequencing of PCR.
products using ABI3700 capillary sequencers (Applied Biosystems) according to standard protocols.

Cell culture
Human colon cancer LoVo cells were grown in F12-HAM (Invitrogen) with 10% fetal bovine serum (FBS). Human embryonic kidney fibroblasts 293FT cells were grown in Dulbecco's modified Eagle's medium (Invitrogen) with 10% FBS. These cells were incubated in a humidified atmosphere with 5% CO₂ at 37°C.

Electrophoretic mobility shift assay
5'-biotin-labeled single-strand oligonucleotides were obtained from Invitrogen and annealed. The sequence of EMSA probe rs35234164 del is 5'-CAGTGAAGTAAACATGCCTACCTTAAGATGTCATTATAA-3'; rs35234164 T: 5'-CAGTGAAGTAAACATGCCTACCTTAAGATGTCATTATAA-3'; rs4480162 C/rS376531 G: 5'-AGTCCGACTCCTTAGCTCTAAGATTTAGTCCGACTCCTTAGTGTGAACTCCAGATCCACCTTCTCTGAACCTTGACTCTTGGGCAGTTTTAAGTACGTTAATAGGTTTAAAATTCTCCCGCTGCCAGA-3'; rs4480162 G/rs4376531 C: 5'-AGTGCGAAGATTTAGTCCGACTCCTTAGTGTGAACTCCAGATCCACCTTCTCTGAACCTTGACTCTTGGGCAGTTTTAAGTACGTTAATAGGTTTAAAATTCTCCCGCTGCCAGA-3'; rs4480162 C/rs4376531 C: 5'-AGTGCGAAGATTTAGTCCGACTCCTTAGTGTGAACTCCAGATCCACCTTCTCTGAACCTTGACTCTTGGGCAGTTTTAAGTACGTTAATAGGTTTAAAATTCTCCCGCTGCCAGA-3'; rs35234164 T: 5'-CGTGAAGTAAACATGCCTACCTTAAGATGTCATTATAA-3'; rs4480162 C/rS376531 G: 5'-AGTCCGACTCCTTAGCTCTAAGATTTAGTCCGACTCCTTAGTGTGAACTCCAGATCCACCTTCTCTGAACCTTGACTCTTGGGCAGTTTTAAGTACGTTAATAGGTTTAAAATTCTCCCGCTGCCAGA-3'; rs4480162 G/rs4376531 C: 5'-AGTGCGAAGATTTAGTCCGACTCCTTAGTGTGAACTCCAGATCCACCTTCTCTGAACCTTGACTCTTGGGCAGTTTTAAGTACGTTAATAGGTTTAAAATTCTCCCGCTGCCAGA-3'; rs4480162 C/rs4376531 C: 5'-AGTGCGAAGATTTAGTCCGACTCCTTAGTGTGAACTCCAGATCCACCTTCTCTGAACCTTGACTCTTGGGCAGTTTTAAGTACGTTAATAGGTTTAAAATTCTCCCGCTGCCAGA-3'; rs35234164 T: 5'-CGTGAAGTAAACATGCCTACCTTAAGATGTCATTATAA-3'; rs4480162 C/rS376531 G: 5'-AGTCCGACTCCTTAGCTCTAAGATTTAGTCCGACTCCTTAGTGTGAACTCCAGATCCACCTTCTCTGAACCTTGACTCTTGGGCAGTTTTAAGTACGTTAATAGGTTTAAAATTCTCCCGCTGCCAGA-3'; rs4480162 G/rs4376531 C: 5'-AGTGCGAAGATTTAGTCCGACTCCTTAGTGTGAACTCCAGATCCACCTTCTCTGAACCTTGACTCTTGGGCAGTTTTAAGTACGTTAATAGGTTTAAAATTCTCCCGCTGCCAGA-3'; rs4480162 C/rs4376531 C: 5'-AGTGCGAAGATTTAGTCCGACTCCTTAGTGTGAACTCCAGATCCACCTTCTCTGAACCTTGACTCTTGGGCAGTTTTAAGTACGTTAATAGGTTTAAAATTCTCCCGCTGCCAGA-3';

Luciferase reporter assay
The same DNA sequences around rs4480162 and rs4376531 as the EMSA probes were subcloned into pGL3-promoter luciferase vector (Promega). We transfected LoVo cells with 500 ng of each reporter construct and 50 ng of pRL-CMV vector (Promega) using FuGENE 6 Transfection Reagent (Roche). After 48 h, we collected the cells and measured luciferase activities using Dual Luciferase Assay System (Toyo B-Net).

Small GTPase activity assay
A plasmid designed to express full-length ARHGEF10 was obtained by cloning full-length human ARHGEF10 cDNA into p3XFLAG-CMV-10 expression vector (SIGMA). We constructed three small GTPases overexpression plasmids by cloning full-length human RhoA, Rac1 or Cdc42 cDNA into pcDNA3.1/myc-His expression vector (Invitrogen). The cellular levels of GTP-loaded RhoA, Rac1 and Cdc42 were determined using GST fusion proteins containing the RhoGTPase-binding domain of Rhotekin (GST-RBD) (14-383, Upstate) or PAK-1 (GST-PBD) (14-325, Upstate) as described previously (30,31). In brief, pcDNA3.1-RhoA-Myc or pcDNA3.1-Rac1-Myc or pcDNA3.1-Cdc42-Myc was co-transfected into 293FT cells seeded in 10 cm dishes with p3XFLAG-CMV-10-ARHGEF10 or the corresponding empty vector using FuGENE 6 Transfection Reagent (Roche). After being cultured for 48 h, the cells were lysed in a buffer containing 25 mM HEPES, pH 7.5, 150 mM NaCl, 1% Igepal CA-630, 10 mM MgCl₂, 1 mM EDTA, 10% glycerol and protease inhibitors, and the particular fraction was pelleted by centrifugation. The GTPase-containing supernatant was then incubated for 45–60 min at 4°C with GST fusion proteins bound to glutathione-Sepharose beads. After three times washing of the beads, bound proteins were eluted with sample buffer and separated by SDS–PAGE. ARHGEF10 and the small GTPases were then detected by immunoblotting with commercially available specific anti-FLAG antibody (F3165, SIGMA, 1 μg/ml) and anti-Myc antibody (562, MBL, 1 μl/ml), respectively. Proteins reacting with primary antibodies were visualized by an enhanced chemiluminescence system (GE Healthcare UK Ltd, Amersham) for detecting species-matched secondary antibodies and analyzed with an LAS-3000 system. Quantitative analyses of immunoblots were performed using Multi Gauge version 2.02 software included in an LAS-3000 system.

Statistical analysis
We assessed case–control association analysis by χ² test and Fisher’s exact test, as appropriate. The shift of Hardy–Weinberg equilibrium was also tested by χ² test or Fisher’s exact test. In the association analyses, we used allele, dominant and recessive models. All statistical analyses were performed without the adjustment of age and sex. Meta-analyses of the two case–control sample sets were performed using Mantel–Haenszel method (fixed effect analysis). Heterogeneities across the population were assessed using Cochran’s Q test. For the adjustment of multiple testing in the discovery phase, we performed a random permutation test with 10 000 replications using MULTTEST procedure of SAS software version 9.12 (SAS Institute). LD were calculated as D’ or r², and haplotype blocks were defined by Gabriel’s criteria (32) using Haploview version 4.0 (Broad Institute). Luciferase assay data and small GTPase activity assay data were analyzed by Student’s t-test.

SUPPLEMENTARY MATERIAL
Supplementary Material is available at HMG online.
ACKNOWLEDGEMENTS

We thank the residents of Hisayama town and the patients with ischemic stroke for their participation; T. Omae and the staff of the Division of Health and Welfare of Hisayama for their cooperation; many members of the Hisayama study for assistance; T. Ago, H. Ooboshi, M. Kamouchi, H. Sugimori, J. Kuroda, Y. Kumai, N. Hagiwara, S. Yoshimura (Kyushu University Hospital), K. Tamaki, Y. Wakugawa (Hakujuyuji Hospital), K. Fuji (Fukuoka Red Cross Hospital), Y. Okada, K. Toyoda (National Hospital Organization, Kyushu Medical Center), T. Nagao (Imazu Red Cross Hospital), H. Nakane (National Hospital Organization, Fukuoka Higashi Medical Center), Y. Yamashita, K. Kusuda (Seiai Rehabilitation Hospital) for sample collection. We thank all the patients who participated in the BioBank Japan project. We also thank all members of the BioBank Japan, Institute of Medical Science, The University of Tokyo, and of the Center for Genomic Medicine, RIKEN, for their contribution to the completion of our study.

Conflict of Interest statement. None declared.

FUNDING

This work was supported in part by the Ministry of Education, Culture, Sports, Science and Technology, Japan.

REFERENCES


