Am J Transl Res 2021;13(5):4521-4534
www.ajtr.org /ISSN:1943-8141/AJTR0127668

Original Article

GWAS-linked hot loci predict short-term functional outcome and recurrence of ischemic stroke in Chinese population

Ruixia Zhu, Yating Zhao, Dandan Tian, Na Guo, Chenguang Zhang, Xu Liu

Department of Neurology, First Affiliated Hospital of China Medical University, Shenyang, Liaoning, China

Received December 5, 2020; Accepted March 2, 2021; Epub May 15, 2021; Published May 30, 2021

Abstract: In the past decade, an increasing number of genome-wide association studies (GWASs) have been applied to ischemic stroke (IS) susceptibility and recovery. In our study, six GWAS-linked hot loci (*ALDH2* rs10744777, *HDAC9* rs2107595, *ABO* rs532436, *PATJ* rs76221407, *LOC105372028* rs1842681 and *PTCH1* rs2236406) were selected, genotyped and analyzed in 982 IS patients from northern Chinese population, in order to explore their roles in stroke functional outcome and recurrence risk. We found that *PTCH1* rs2236406 was significantly associated with functional outcome after stroke. Further logistic regression analysis revealed the variant genotype TC/CC of rs2236406 as an independent prognostic factor for poor stroke recovery in Chinese population. Meanwhile, we observed that GA/AA genotype of *ABO* rs532436 was statistically correlated with the increased risk of stroke recurrence, especially for patients with large-artery atherosclerosis. Moreover, multivariate Cox analysis identified *ABO* rs12342 as an independent predictor for IS recurrence. Further functional annotation analysis demonstrated that rs2236406 and rs2043211 were located in the transcriptionally active region, and could change the regulatory motif, transcription factor binding capacity and expression level of RP11-43505.5 (antisense to PTCH1) and ABO, respectively. In summary, our results suggested that *PTCH1* rs2236406 and *ABO* rs532436 may be novel genetic markers and potential therapeutic targets for stroke prognosis. More studies are required to confirm our findings and clarify the underlying molecular mechanisms.

Keywords: Ischemic stroke, GWAS, stroke recovery, stroke recurrence, genetic polymorphism

Introduction

Ischemic stroke (IS) is the leading cause of adult disability and one of the most common causes of death worldwide [1]. It brings great distress to patients and families and imposes a tremendous financial cost on society [2]. In the United States alone, between 2012 and 2030, the total annual medical expenses related to IS are expected to increase from $105.2 billion to $240.7 billion [3]. Due to advances in stroke treatment, the global mortality rate for IS has decreased in the past decade, and most patients could survive the first stroke [4]. However, IS survivors have to face the risk of permanent disability and subsequent stroke recurrence. Therefore, an urgent need exists to explore novel genetic biomarkers and potential therapeutic targets for stroke recovery and recurrence.

Single nucleotide polymorphism (SNP) is believed to play a vital role in many complex diseases, including ischemic stroke [5, 6]. However, traditional candidate gene studies have produced few replicable associations, and another drawback of this hypothesis-driven approach is that genes involved in disease pathogenesis through unknown pathway are ignored [7]. In contrast, genome-wide association studies (GWASs) are designed to provide broad coverage of the entire genome and could identify common genetic polymorphisms that affect disease risk and prognosis. Moreover, the hypothesis-free approach of GWAS has proven successful in exploring novel pathophysiological mechanisms underlying disease [8]. Thus, in the past decade, an increasing number of GWASs have been applied to IS susceptibility and recovery and have transformed the genetics of this complex human disease.

In 2014, Kilarski et al. first found that a novel top SNP rs10744777 on 12q24 locus near *ALDH2* was associated with IS susceptibility in an overall analysis of 17,970 cases and 70,764 controls, which was further replicated in small artery stroke by Stroke Genetics Network (SiGN) GWAS [9, 10]. Moreover, Malik et al. identified two genome-wide significant SNPs, *HDAC9* rs2107595 and *ABO* rs532436, related with IS risk through a GWAS and subsequent transethnic meta-analysis [11]. In 2018, the largest GWAS further confirmed this relationship based on multiancestry population comprising 520,000 subjects [12]. Regarding GWAS focusing on functional outcome of IS patients, a novel variant, rs76221407, in *PATJ* gene was associated with worse outcome at 3 months after stroke in European ancestry [13]. Besides, Söderholm et al. observed the relationship of four genetic polymorphisms (*LOC105372028* rs1842681, *PTCH1* rs2236406, *PLAA* rs13299556 and *NTN4* rs78734480) with functional outcome among 6165 European stroke patients [14].

However, at present, the role of GWAS-related hot SNPs in stroke recurrence remains unclear. Moreover, we do not know the impact of these GWAS-related SNPs on functional recovery after stroke in Chinese population. Thus, we designed a prospective cohort study to explore the relationship between six GWAS-related hot SNPs (*ALDH2* rs10744777, *HDAC9* rs2107595, *ABO* rs532436, *PATJ* rs76221407, *LOC105372028* rs1842681 and *PTCH1* rs2236406) and the short-term prognosis (stroke recurrence and functional outcome) in Chinese population.

Materials and methods

Study subjects

We enrolled 982 first-ever IS patients prospectively between November 2016 and December 2019 from the First Hospital of China Medical University. Among these patients, IS was diagnosed according to the focal neurological deficits lasting more than 24 hours, and future confirmed by magnetic resonance imaging and/or computed tomography. Based on TOAST classification, IS can be classified into five subtypes: small-vessel occlusion (SVO), large-artery atherosclerosis (LAA), cardioembolism, stroke with other determined etiology, as well as stroke with undetermined etiology [15]. In our study, only IS patients with SVO and LAA subtypes were included, while the stroke patients with other subtypes were excluded.

In accordance with the Declaration of Helsinki, the study was conducted with the approval of the ethics committee of the First Hospital of China Medical university. All IS patients were given written informed consent.

Follow-up and outcome

Three months after the index stroke, enrolled patients were prospectively followed up through a clinical visit or telephone interview. Two clinical investigators, who were blinded to the baseline data, evaluated the stroke functional outcome by using the modified Rankin scale (mRS). Then, they divided the IS patients into two groups based on the degree of neurologic functional disability. The mRS score 0-2 has been used to define “favorable outcome” group, while stroke patients with mRS 3-6 were grouped as having poor outcome.

Subsequent follow-ups were performed every three months by using a standard questionnaire until the last follow-up or stroke recurrence. In the end, among 982 IS patients, 940 completed the follow-up with an average time of 14 months, and 42 (4.28%) patients were lost during the follow-up period.

Screening GWAS-linked hot SNPs

PubMed database was searched for all GWASs-focusing on ischemic stroke during the last 10 years. We found 11 studies on IS susceptibility and 2 studies on functional outcome [9-14, 16-22]. Among 11 GWASs involving IS risk, only SNPs that met the following criteria were selected: 1) The *P* value of SNPs exceeding the 5×10^{-8} level of significance typically considered the threshold for GWAS. 2) The SNPs validated in two or more GWASs. 3) Minor allele frequency (MAF) > 0.05 for Chinese Han in Beijing (CHB) population. Finally, three SNPs (*ABO* rs532436, *HDAC9* rs2107595, *ALDH2* rs10744777) were chosen for further genotyping.

In addition, two GWASs on functional outcome after IS were published in 2019 and five SNPs (*PATJ* rs76221407, *LOC105372028* rs1842681, *PTCH1* rs2236406, *PLAA* rs13299556 and *NTN4* rs78734480) were reported to affect

Table 1. Baseline characteristics of ischemic stroke patients grouped by short-term functional outcome

Variables	mRS (0-2) N = 615	mRS (3-6) N = 263	*P*-value
Age ≥60	350	186	<0.01
Male	424	170	0.21
Hypertension	474	208	0.51
Diabetes	221	126	<0.01
Hyperlipidemia	247	93	0.18
Smoking	253	115	0.48
Drinking	134	68	0.19
NIHSS	3.86±1.78	10.34±3.28	<0.01
TOAST	293	190	<0.01

the recovery after stroke [13, 14]. Notably, the MAF of SNPs, rs13299556 and rs78734480, was 0 in CHB population. Thus, we selected the other three SNPs (rs76221407, rs1842681 and rs2236406) for genotyping.

DNA isolation and genotyping

Genomic DNA was extracted from peripheral blood by standard procedures and subsequently stored at temperatures below -20°C. SNPscan technology was used for the genotyping of rs532436, rs2107595, rs10744777, rs76221407, rs1842681 and rs2236406. For quality control, the reproducibility of genotyping was conducted in 5% randomly selected samples and 100% consistency was achieved.

Statistical analysis

Categorical variables were presented as counts and percentages, and compared using Chi-square test. Continuous variables with normal distributions were summarized as mean ± standard deviation (SD), and compared by Student's t test. Binary logistic regression analysis was used to assess the association of six GWAS-linked hot SNPs with short-term functional outcome after IS. Each SNP was analyzed by calculating odds ratio (OR) and 95% confidence interval (CI) under additive, recessive and dominant models. Above statistical analyses as well as interaction analysis with environmental factors were carried out by using SNPStats software (https://www.snpstats.net /start.htm) [23].

Then, the recurrence time was calculated from the diagnosis date of IS to the last time of follow-up or the date of IS recurrence. The recurrence curve was performed using the Kaplan-Meier method, and the difference of recurrence time was assessed by log-rank test. Moreover, Cox proportional hazards regression model was adopted to calculated the hazard ratio (HR) and 95% CI. The SPSS 17.0 and GraphPad Prism version 8 were used to analyze the data. A *P* value of <0.05 was considered statistically significant.

Functional annotation

For GWAS-linked hot SNPs showing statistically significance in predicting the functional outcome and recurrence of IS, we retrieved for their potential functions in HaploReg tool V4.1 (http://archive.broadinstitute.org/mammals/haploreg/haploreg.php) and Genotype-Tissue Expression (GTEx) project (https://www.gtexportal.org) [24-26].

Results

GWAS-linked hot SNPs and functional outcome after stroke

Three months after the index stroke, the stroke functional outcome of 878 eligible patients was evaluated using the mRS scale. Among them, 615 patients had a good outcome, while 263 had a poor outcome. As shown in **Table 1**, compared with the poor outcome group, the patients with good functional outcome were younger and had a lower prevalence of diabetes. Moreover, National Institutes of Health Stroke Scale (NIHSS) score on admission and TOAST classification were also significantly different between the poor and good outcome groups (*P*<0.01).

Table 2 displays the effect of GWAS-linked hot SNPs on the functional outcome after stroke. As a result, only *PTCH1* rs2236406 was observed to be significantly associated with IS outcome in additive, recessive and dominant models. Compared with TT genotype, CC and TC/CC genotype were represented more frequently in the poor outcome group (OR = 1.94, 95% CI = 1.25-3.01; OR = 1.35, 95% CI = 1.00-1.83). Meanwhile, we observed a higher CC genotype frequency for rs2236406 in all participants with poor functional outcome (OR = 1.74, 95% CI = 1.17-2.59). Furthermore, binary logistic regression model was used to determine the relationship of *PTCH1* rs2236406 with functional outcome after stroke, and the

Table 2. Association of GWAS-linked hot SNPs with short-term functional outcome after stroke

Variable	mRS (0-2)	mRS (3-6)	OR (95% CI)	*P*-value	OR (95% CI)[a]	*P*-value[a]
rs10744777						
CC	576 (93.7%)	246 (93.5%)	1.00	0.69	1.00	0.95
TC	38 (6.2%)	17 (6.5%)	1.05 (0.58-1.89)		1.20 (0.39-3.70)	
TT	1 (0.2%)	0 (0%)	0.00 (0.00-NA)		0.00 (0.00-NA)	
Dominant model						
CC	576 (93.7%)	246 (93.5%)	1.00	0.95	1.00	0.76
TC+TT	39 (6.3%)	17 (6.5%)	1.02 (0.57-1.84)		1.20 (0.39-3.70)	
Recessive model						
CC+TC	614 (99.8%)	263 (100%)	1.00	0.4	1.00	0.96
TT	1 (0.2%)	0 (0%)	0.00 (0.00-NA)		0.00 (0.00-NA)	
rs1842681						
GG	556 (90.4%)	229 (87.1%)	1.00	0.087	1.00	0.30
AG	56 (9.1%)	34 (12.9%)	1.47 (0.94-2.32)		1.90 (0.84-4.31)	
AA	3 (0.5%)	0 (0%)	0.00 (0.00-NA)		0.00 (0.00-NA)	
Dominant model						
GG	556 (90.4%)	229 (87.1%)	1.00	0.15	1.00	0.13
AG+AA	59 (9.6%)	34 (12.9%)	1.40 (0.89-2.19)		1.89 (0.83-4.27)	
Recessive model						
GG+AG	612 (99.5%)	263 (100%)	1.00	0.14	1.00	0.80
AA	3 (0.5%)	0 (0%)	0.00 (0.00-NA)		0.00 (0.00-NA)	
rs2107595						
GG	277 (45%)	119 (45.2%)	1.00	0.59	1.00	0.08
GA	268 (43.6%)	108 (41.1%)	0.94 (0.69-1.28)		0.58 (0.32-1.05)	
AA	70 (11.4%)	36 (13.7%)	1.20 (0.76-1.89)		1.36 (0.56-3.29)	
Dominant model						
GG	277 (45%)	119 (45.2%)	1.00	0.96	1.00	0.19
GA+AA	338 (55%)	144 (54.8%)	0.99 (0.74-1.33)		0.69 (0.40-1.21)	
Recessive model						
GG+GA	545 (88.6%)	227 (86.3%)	1.00	0.34	1.00	0.18
AA	70 (11.4%)	36 (13.7%)	1.23 (0.80-1.90)		1.79 (0.78-4.12)	
rs2236406						
TT	249 (40.5%)	88 (33.5%)	1.00	0.014	1.00	0.096
TC	296 (48.1%)	127 (48.3%)	1.21 (0.88-1.67)		1.79 (0.99-3.24)	
CC	70 (11.4%)	48 (18.2%)	1.94 (1.25-3.01)		2.12 (0.83-5.42)	
Dominant model						
TT	249 (40.5%)	88 (33.5%)	1.00	0.049	1.00	0.033
TC+CC	366 (59.5%)	175 (66.5%)	1.35 (1.00-1.83)		1.85 (1.04-3.26)	
Recessive model						
TT+TC	545 (88.6%)	215 (81.8%)	1.00	0.0076	1.00	0.34
CC	70 (11.4%)	48 (18.2%)	1.74 (1.17-2.59)		1.52 (0.64-3.60)	
rs532436						
GG	380 (61.8%)	162 (61.6%)	1.00	1.00	1.00	0.40
GA	207 (33.7%)	89 (33.8%)	1.01 (0.74-1.37)		0.66 (0.36-1.21)	
AA	28 (4.5%)	12 (4.6%)	1.01 (0.50-2.03)		0.82 (0.24-2.83)	
Dominant model						
GG	380 (61.8%)	162 (61.6%)	1.00	0.96	1.00	0.19
GA+AA	235 (38.2%)	101 (38.4%)	1.01 (0.75-1.36)		0.68 (0.39-1.21)	

Recessive model						
GG+GA	587 (95.5%)	251 (95.4%)	1.00	0.99	1.00	0.92
AA	28 (4.5%)	12 (4.6%)	1.00 (0.50-2.00)		0.94 (0.28-3.18)	
rs76221407						
AA	527 (85.7%)	228 (86.7%)	1.00	0.75	1.00	1
AG	87 (14.2%)	34 (12.9%)	0.90 (0.59-1.38)		1.03 (0.49-2.17)	
GG	1 (0.2%)	1 (0.4%)	2.31 (0.14-37.12)		0.54 (0.00-NA)	
Dominant model						
AA	527 (85.7%)	228 (86.7%)	1.00	0.69	1.00	0.93
AG+GG	88 (14.3%)	35 (13.3%)	0.92 (0.60-1.40)		1.03 (0.49-2.17)	
Recessive model						
AA+AG	614 (99.8%)	262 (99.6%)	1.00	0.55	1.00	0.97
GG	1 (0.2%)	1 (0.4%)	2.34 (0.15-37.61)		0.54 (0.00-NA)	

[a]Adjusted for age, gender, hypertension, diabetes, hyperlipidemia, smoking, drinking and NIHSS.

Table 3. Stratification and interaction analysis for *PTCH1* rs2236406 with short-term functional outcome after stroke

Variables	Patients mRS (0-2)/(3-6)	TC+CC mRS (0-2)/(3-6)	TT mRS (0-2)/(3-6)	OR (95% CI)[a]	Interaction *P* value
Age (years)					
<60	265/77	159/51	106/26	1.95 (0.76-4.97)	0.89
≥60	350/186	207/124	143/62	1.79 (0.88-3.64)	
Gender					
Male	424/170	244/112	180/58	1.95 (0.99-3.85)	0.77
Female	191/93	122/63	69/30	1.62 (0.58-4.58)	
Hypertension					
No	141/55	76/37	65/18	6.26 (1.55-25.28)	0.05
Yes	474/208	290/138	184/70	1.43 (0.76-2.67)	
Diabetes					
No	394/137	229/93	165/44	2.40 (1.11-5.17)	0.31
Yes	221/126	137/82	84/44	1.32 (0.57-3.10)	
Hyperlipidemia					
No	368/170	207/115	161/55	2.02 (0.97-4.17)	0.70
Yes	247/93	159/60	88/33	1.61 (0.65-3.95)	
Smoking					
No	362/148	214/95	148/53	1.51 (0.70-3.23)	0.43
Yes	253/115	152/80	101/35	2.36 (1.01-5.47)	
Drinking					
No	481/195	288/132	193/63	1.89 (0.96-3.72)	0.90
Yes	134/68	78/43	56/25	1.74 (0.61-4.97)	
TOAST					
LAA	293/190	174/124	119/66	1.56 (0.77-3.17)	0.44
SVO	322/73	192/51	130/22	2.48 (0.96-6.38)	

[a]Adjusted for age, gender, hypertension, diabetes, hyperlipidemia, smoking, drinking and NIHSS.

results revealed that after adjusting for potential covariates, TC/CC genotype was significantly associated with poor recovery after stroke (OR = 1.85, 95% CI = 1.04-3.26).

Subgroup analyses based on age, gender, smoking and drinking status, diabetes, hyperlipidemia, hypertension and IS subtypes are listed in **Table 3**. We found that in patients with

Table 4. Clinical characteristics of patients grouped by ischemic stroke recurrence

Variables	Patients N = 940	Recurrence N = 139	Log-rank *P*
Age (years)			
≥60	582	97	0.02
<60	358	42	
Gender			
Male	634	94	0.78
Female	306	45	
Hypertension			
No	211	26	0.13
Yes	729	113	
Diabetes			
No	560	76	0.10
Yes	380	63	
Hyperlipidemia			
No	574	83	0.92
Yes	366	56	
Smoking			
No	551	84	0.29
Yes	389	55	
Drinking			
No	724	108	0.54
Yes	216	31	
TOAST			
LAA	527	91	0.02
SVO	413	48	

smoking habit, the poor outcome risk for rs2236406 TC/CC genotype was significantly increased (OR = 2.36, 95% CI = 1.01-5.74). Additionally, the increased poor outcome risk was more evident in the participants without hypertension or diabetes (OR = 6.26, 95% CI = 1.55-25.28; OR = 2.40, 95% CI = 1.11-5.17). Thus, subsequent interaction analysis of rs2236406 with hypertension, diabetes and smoking was conducted under the dominant model. The interaction between rs2236406 and hypertension was marginally significant ($P_{interaction}$ = 0.05). Compared with non-hypertensive patients with TT genotype, hypertensive patients with TT genotype had a 3.70-fold increased risk of poor outcome (OR = 3.70, 95% CI = 1.03-13.23). In addition, no statistically significant interaction was detected between diabetes and smoking status and rs2236406 ($P_{interaction}$ = 0.31 and $P_{interaction}$ = 0.43).

Table 5. Association between GWAS-linked hot SNPs and stroke recurrence risk

Genotype	Patients N = 940	Recurrence N = 139	Log-rank *P*
rs10744777			
CC	881	127	
TC	51	12	0.28
TT	1	0	0.66
rs1842681			
GG	835	121	
AG	101	17	0.47
AA	4	1	0.67
rs2107595			
GG	423	62	
GA	403	60	0.92
AA	114	17	0.79
rs2236406			
TT	365	48	
TC	448	67	0.60
CC	127	24	0.13
rs532436			
GG	569	70	
GA	325	59	0.02
AA	46	10	0.06
GA+AA	371	69	0.01
rs76221407			
AA	805	120	
AG	133	19	0.99
GG	2	0	0.56

GWAS-linked hot SNPs and recurrence risk of stroke

The baseline characteristics of 940 IS patients for recurrence risk analysis are listed in **Table 4**. Among them, 582 patients were older than 60 years, and 527 were patients with large-artery atherosclerotic stroke. Until the end of the follow-up period, 139 participants had stroke recurrence. As shown in **Table 4**, age and stroke classification were significantly associated with stroke recurrence risk (P<0.05).

Table 5 displays the relationship of GWAS-linked hot SNPs with recurrence risk of stroke. Among these six selected SNPs, we observed that *ABO* rs532436 was significantly associated with the risk of stroke recurrence (GA vs. GG: P = 0.02; GA+AA vs. GG: P = 0.01, **Figure 1**), while other five SNPs (rs10744777, rs1842681, rs2107595, rs2236406 and rs76221407)

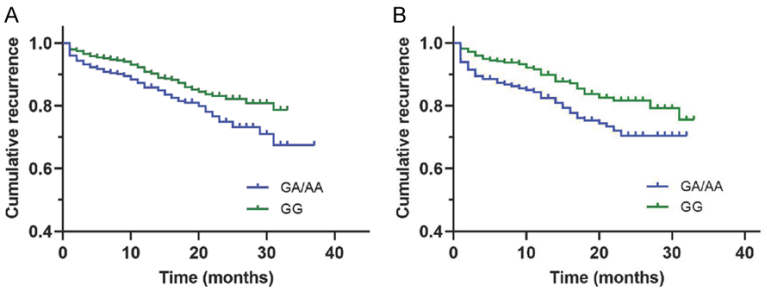

Figure 1. Kaplan-Meier plots for stroke recurrence risk based on *ABO* rs532436 genotypes. A. Recurrence plot under dominant model. B. Recurrence plot for large-artery atherosclerotic stroke.

Table 6. Association between *ABO* rs532436 and stroke recurrence risk

Genotype	Patients	Recurrence	HR (95% CI)[a]	*P*[a]
Genotypic Model				
GG	569	70		
GA	325	59	1.46 (1.03-2.07)	0.03
AA	46	10	2.07 (1.05-4.09)	0.04
Dominant model				
GA+AA vs. GG	371/569	69/70	1.53 (1.09-2.13)	0.01
Recessive model				
AA vs. GA+GG	46/894	10/129	1.84 (0.96-3.54)	0.07

[a]Adjusted for age, gender, hypertension, diabetes, hyperlipidemia, smoking and drinking.

were not related. Then, using Cox regression to evaluate the relationship between rs532436 and recurrence risk, we found that rs532436 was an independent prognostic factor for stroke recurrence (**Table 6**). Further subgroup analysis based on TOAST classification indicated that rs532436 GA/AA genotype was associated with an increased recurrence risk for large-artery atherosclerotic stroke but not small-vessel disease (LAA: $P = 0.01$, SVO: $P = 0.49$, **Figure 1**).

In the stratified analyses by the clinical characteristics of IS patients, we observed the increased stroke recurrence for rs532436 GA/AA genotype in females as well as elderly participants (HR = 1.88, 95% CI = 1.04-3.41; HR = 1.52, 95% CI = 1.02-2.27, **Table 7**). In addition, the increased risk of stroke recurrence was more pronounced among patients with hypertension and diabetes (HR = 1.62, 95% CI = 1.12-2.34; HR = 1.72, 95% CI = 1.04-2.83), and patients without smoking and drinking habits (HR = 1.87, 95% CI = 1.21-2.88; HR = 1.90, 95% CI = 1.30-2.77).

Functional annotation

We first performed in silico prediction of the potential function for these two significant SNPs (rs2236406 and rs532436) by using HaploReg V4.1 (**Table 8**). In various tissues and cell lines, these two polymorphisms were located in the regions with four histone modifications (H3K4me3, H3K9ac, H3K4me1 and H3K27ac), which were the markers reflecting the transcriptional status of promoters and enhancers. They were also located in DNase hypersensitive sites and could change the regulatory motifs (rs2236406: Glis2, LF-A1, Spz1_1 and ZBTB7A_known2; rs532436: ERalpha-a_disc4, Rad21_disc5 and VDR_4). In addition, rs2236406 was located at the binding site of the transcription factor GATA2. Taken together, these findings indicated that rs2236406 and rs532436 may be involved in regulating gene expression.

To verify the above findings, we then conducted an expression quantitative trait loci (eQTL) analysis using GTEx dataset. As summarized in **Table 8**, we found a significant association between rs2236406 and RP11-43505.5 (antisense to PTCH1) levels in tibial artery ($P = 2.1E-07$) and tibial nerve ($P = 2.2E-08$). Moreover, a significant association was observed between rs532436 and ABO expression in adrenal gland ($P = 2.6E-06$) and whole blood ($P = 2.6E-09$).

Discussion

To our knowledge, this is the first study to explore the association between six GWAS-related hot SNPs (*ALDH2* rs10744777, *HDAC9* rs2107595, *ABO* rs532436, *PATJ* rs762214-07, *LOC105372028* rs1842681 and *PTCH1* rs2236406) and stroke recovery and recurrence in a northern Chinese population.

Our study showed that *PTCH1* rs2236406 was significantly related with functional outcome

Table 7. Stratified analysis for *ABO* rs532436 with recurrence risk of ischemic stroke

Variable	Genotype (recurrence/patients)		HR (95% CI)[a]	*P*[a]
	GA+AA	GG		
Total	69/371	70/569	1.53 (1.09-2.13)	0.01
Age				
<60	20/132	22/226	1.61 (0.87-2.98)	0.13
≥60	49/239	48/343	1.52 (1.02-2.27)	0.04
Gender				
Male	45/246	49/388	1.36 (0.91-2.05)	0.14
Female	24/125	21/181	1.88 (1.04-3.41)	0.04
Hypertension				
No	10/78	16/133	1.25 (0.56-1.01)	0.59
Yes	59/293	54/436	1.62 (1.12-2.34)	0.01
Diabetes				
No	36/222	40/338	1.34 (0.85-2.11)	0.20
Yes	33/149	30/231	1.72 (1.04-2.83)	0.03
Hyperlipidemia				
No	39/217	44/357	1.51 (0.98-2.34)	0.06
Yes	30/154	26/212	1.62 (0.95-2.76)	0.08
Smoking				
No	46/221	38/330	1.87 (1.21-2.88)	<0.01
Yes	23/150	32/239	1.15 (0.66-1.99)	0.63
Drinking				
No	58/279	50/445	1.90 (1.30-2.77)	<0.01
Yes	11/92	20/124	0.67 (0.32-1.41)	0.29
TOAST				
LAA	47/212	44/315	1.69 (1.12-2.56)	0.01
SVO	22/159	26/254	1.22 (0.69-2.17)	0.49

[a]Adjusted for age, gender, hypertension, diabetes, hyperlipidemia, smoking and drinking.

after stroke, but not the recurrence risk. A 35% increased poor outcome risk was observed among IS patients with TC/CC genotype within rs2236406 compared with those with TT genotype, indicating that C allele could be considered as an unfavorable factor for stroke outcome. Moreover, the results of multivariate logistic regression model revealed that rs2236406 TC/CC genotype was independently associated with poor recovery. Subsequent subgroup analyses showed that non-hypertensive, non-diabetic and smoking patients with rs2236406 TC/CC genotype had an increased risk of poor functional outcome after stroke. Moreover, a marginally significant interaction between rs2236406 TC/CC and hypertension was detected. In total, these findings support the role of *PTCH1* rs2236406 polymorphism as a novel genetic biomarker and potential therapeutic target for stroke recovery in Chinese population. In addition, we cannot replicate the association between *PATJ* rs76221407 and *LOC105372028* rs1842681 and stroke recovery. This discrepancy could be attributed to genetic heterogeneity across different racial and geographical groups (G allele frequency for rs76221407: 0.06 in East Asians and 0.01 in Europeans, A allele frequency for rs1842681: 0.07 in East Asians and 0.22 in Europeans; data from 1000 Genomes Project).

How then can we explain the association between *PTCH1* rs2236406 polymorphism and stroke recovery? First, Patched-1 (PTCH1, encoded by *PTCH1* gene) is a 12-pass transmembrane protein that acts as the receptor for Hedgehog (Hh) ligands, including Sonic Hedgehog (Shh) [27]. When Shh binds to PTCH1, PTCH1 will be inactivated, thereby unleashing the activity of Smoothened, and further allowing transcription factor Gli1 to initiate target gene transcription [28]. Thus, PTCH1 is considered to be an important participant in the Shh signaling pathway. Moreover, growing studies have suggested that Shh signaling can not only play an active role in CNS neurogenesis, but also exhibit neuroprotective effects in cerebral ischemia [29, 30]. In oxygen/glucose deprivation cells and rat brains following middle cerebral artery occlusion, the inhibition of Shh signaling pathway can exacerbate neuronal apoptosis, and increase infarction volume and neurobehavioral deficits [31, 32]. Similarly, Sims et al. found that hypoxia induced proliferation of neural progenitor cells in vitro, and this proliferation was enhanced by adding recombinant Shh or blocked by cyclopamine (a specific inhibitor of Shh pathway), indicating that Shh signaling

Table 8. Functional annotation for *PTCH1* rs2236406 and *ABO* rs532436

SNP	SNP region	Gene	Haploreg					GTEx		
			Promoter markers (H3K4me3/H3K9ac)	Enhancer markers (H3K4me1/H3K27ac)	DNase	Protein bound	Motif change	Gene expression	Tissue	*P* value
rs2236406	Intronic region	*PTCH1*	28 tissues	91 tissues	7 tissues	GATA2	Glis2, LF-A1, Spz1_1, ZBTB7A_known2	RP11-435O5.5 (antisense to PTCH1)	Tibial artery Tibial nerve	2.1E-07 2.2E-08
rs532436	Intronic region	*ABO*	81 tissues	77 tissues	5 tissues	-	ERalpha-a_disc4, Rad21_disc5, VDR_4	ABO	Adrenal gland Whole blood	2.6E-06 2.6E-09

might be involved in injury remodeling [33]. Second, rs2236406 was located in the transcriptionally active region, and may alter the regulatory motif and binding ability of transcription factor. In addition, eQTL analysis showed that it can change the expression of RP11-43505.5 (antisense to PTCH1) in nerves and blood vessels. We speculate RP11-43505.5 may interact with PTCH1 to regulate Shh signaling pathway. Further functional studies are required to confirm our hypothesis and investigate the underlying biological mechanism.

Regarding the recurrence risk of stroke, for the first time, our results demonstrated that *ABO* rs532436 was statistically associated with stroke recurrence, but not the other five GWAS-related SNPs (rs10744777, rs2107595, rs76-221407, rs1842681 and rs2236406). Compared with GG genotype, a 53% increased risk for IS recurrence was found in participants with GA/AA genotype within rs532436, which indicated that A allele may be a risk factor for IS recurrence. Furthermore, Cox regression results demonstrated that rs532436 GA/AA genotype was independently associated with recurrence risk. Subsequent subgroup analysis according to stroke TOAST subtypes showed a more obvious effect of rs532436 GA/AA genotype on LAA subtype. Among LAA stroke patients, A allele carriers had a 69% increment in the risk of stroke recurrence compared with those with GG genotype. Additionally, we performed the stratified analyses based on clinical characteristics and observed an increased recurrent risk for patients with rs532436 GA/AA genotype among females, elders and participants with comorbidities. Taken together, the above findings suggest that *ABO* rs532436 polymorphism has the potential for predicting IS recurrence.

ABO rs532436 polymorphism was found to be related to stroke recurrence in our study, which could be explained by following reasons. First, the *ABO* gene encodes glycosyltransferases, which transfer specific sugar residues to the H-antigen and characterize the ABO groups [34]. Recent evidences indicated that ABO blood group could determine serum von Willebrand factor (vWF) and soluble E-selectin levels [35-37]. Moreover, Williams et al. found that elevated vWF increased the risk for stroke recurrence among 2100 participants from the VISP (Vitamin Intervention for Stroke Prevention) clinical trial, indicating that vWF played a vital role in thrombus formation during cerebrovascular damage [38]. As for E-selectin, it is considered to be a marker of endothelial dysfunction [39]. Up-regulation of E-selectin could contribute to endothelial inflammation, atherosclerosis and subsequent ischemic stroke risk [40, 41]. Second, our eQTL analysis demonstrated that rs532436 could regulate the expression of ABO in whole blood and adrenal gland. Also, rs532436 has been identified to be genome-wide associated with circulating vWF levels in the VISP population (beta = -0.33, P = 1.01E-21) [38]. In addition, Qi et al. performed a GWAS focusing on serum soluble E-selectin concentration and found the strongest association with rs651007, which was almost in perfect linkage disequilibrium with rs532436 among Asians and Europeans (r^2 = 0.99 and 0.99, respectively) [42].

Our study has several advantages. Firstly, it recruited a large cohort of IS patients with prospective data collection. Next, the selected SNPs came from the high-quality GWAS on stroke susceptibility and functional outcome, making it easier for us to discover novel biomarkers and therapeutic targets for stroke prognosis. Then, we collected a wide array of clinical data, including age, gender, comorbidities, smoking and drinking status, allowing us to assess the independent influence of GWAS-linked hot SNPs on stroke recovery and recurrence. Lastly, regarding putative functional annotation, our findings are biologically sound. Meanwhile, some limitations of the present study should be noted. The follow-up period was relatively short, which prevented an evaluation on long-term prognosis of ischemic stroke. Besides, our study only enrolled the participants with the most common stroke subtypes (SVO and LAA), limiting the general application to the other three less common subtypes. Finally, individual data on medications and lifestyle interventions were unavailable. In future studies, more detailed information for each patient should be required in order to make more accurate adjusted estimates.

Conclusion

In summary, our results indicated that *PTCH1* rs2236406 can predict the functional outcome

after ischemic stroke, while *ABO* rs532436 can forecast the recurrence risk of ischemic stroke. More researches are warranted to verify our findings and investigate the underlying molecular mechanisms.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (Grant No. 81400950, 81501006), Natural Science Foundation of Liaoning Province (Grant No. 2019-MS-365, 2019-MS-364).

Disclosure of conflict of interest

None.

Address correspondence to: Dr. Xu Liu, Department of Neurology, First Affiliated Hospital of China Medical University, No. 155 North Nanjing Street, Shenyang 110001, Liaoning, China. Fax: +86-24-83282515; E-mail: valentine1120@126.com

References

[1] Donnan GA, Fisher M, Macleod M and Davis SM. Stroke. Lancet 2008; 371: 1612-1623.

[2] Feigin VL, Norrving B and Mensah GA. Global burden of stroke. Circ Res 2017; 120: 439-448.

[3] Ovbiagele B, Goldstein LB, Higashida RT, Howard VJ, Johnston SC, Khavjou OA, Lackland DT, Lichtman JH, Mohl S, Sacco RL, Saver JL and Trogdon JG; American Heart Association Advocacy Coordinating Committee and Stroke Council. Forecasting the future of stroke in the United States: a policy statement from the American Heart Association and American Stroke Association. Stroke 2013; 44: 2361-2375.

[4] Benjamin EJ, Muntner P, Alonso A, Bittencourt MS, Callaway CW, Carson AP, Chamberlain AM, Chang AR, Cheng S, Das SR, Delling FN, Djousse L, Elkind MSV, Ferguson JF, Fornage M, Jordan LC, Khan SS, Kissela BM, Knutson KL, Kwan TW, Lackland DT, Lewis TT, Lichtman JH, Longenecker CT, Loop MS, Lutsey PL, Martin SS, Matsushita K, Moran AE, Mussolino ME, O'Flaherty M, Pandey A, Perak AM, Rosamond WD, Roth GA, Sampson UKA, Satou GM, Schroeder EB, Shah SH, Spartano NL, Stokes A, Tirschwell DL, Tsao CW, Turakhia MP, VanWagner LB, Wilkins JT, Wong SS and Virani SS; American Heart Association Council on Epidemiology and Prevention Statistics Committee and Stroke Statistics Subcommittee. Heart disease and stroke statistics-2019 update: a report from the American heart association. Circulation 2019; 139: E56-E528.

[5] Meschia JF, Worrall BB and Rich SS. Genetic susceptibility to ischemic stroke. Nat Rev Neurol 2011; 7: 369-378.

[6] Zhu R, Zhao Y, Xiao T, Wang Q and Liu X. Association between microRNA binding site polymorphisms in immunoinflammatory genes and recurrence risk of ischemic stroke. Genomics 2020; 112: 2241-2246.

[7] Dichgans M and Markus HS. Genetic association studies in stroke: methodological issues and proposed standard criteria. Stroke 2005; 36: 2027-2031.

[8] Tan MS, Jiang T, Tan L and Yu JT. Genome-wide association studies in neurology. Ann Transl Med 2014; 2: 124.

[9] Kilarski LL, Achterberg S, Devan WJ, Traylor M, Malik R, Lindgren A, Pare G, Sharma P, Slowik A, Thijs V, Walters M, Worrall BB, Sale MM, Algra A, Kappelle LJ, Wijmenga C, Norrving B, Sandling JK, Rönnblom L, Goris A, Franke A, Sudlow C, Rothwell PM, Levi C, Holliday EG, Fornage M, Psaty B, Gretarsdottir S, Thorsteinsdottir U, Seshadri S, Mitchell BD, Kittner S, Clarke R, Hopewell JC, Bis JC, Boncoraglio GB, Meschia J, Ikram MA, Hansen BM, Montaner J, Thorleifsson G, Stefanson K, Rosand J, de Bakker PI, Farrall M, Dichgans M, Markus HS and Bevan S; 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 2014; 83: 678-685.

[10] 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 2016; 15: 174-184.

[11] Malik R, Traylor M, Pulit SL, Bevan S, Hopewell JC, Holliday EG, Zhao W, Abrantes P, Amouyel P, Attia JR, Battey TWK, Berger K, Boncoraglio GB, Chauhan G, Cheng YC, Chen WM, Clarke R, Cotlarciuc I, Debette S, Falcone GJ, Ferro JM, Gamble DM, Ilinca A, Kittner SJ, Kourkoulis CE, Lemmens R, Levi CR, Lichtner P, Lindgren A, Liu J, Meschia JF, Mitchell BD, Oliveira SA, Pera J, Reiner AP, Rothwell PM, Sharma P, Slowik A, Sudlow CLM, Tatlisumak T, Thijs V, Vicente AM, Woo D, Seshadri S, Saleheen D, Rosand J, Markus HS, Worrall BB and Dichgans M. Low-frequency and common genetic variation in ischemic stroke. Neurology 2016; 86: 1217-1226.

[12] Malik R, Chauhan G, Traylor M, Sargurupremraj M, Okada Y, Mishra A, Rutten-Jacobs L,

Giese AK, van der Laan SW, Gretarsdottir S, Anderson CD, Chong M, Adams HHH, Ago T, Almgren P, Amouyel P, Ay H, Bartz TM, Benavente OR, Bevan S, Boncoraglio GB, Brown RD Jr, Butterworth AS, Carrera C, Carty CL, Chasman DI, Chen WM, Cole JW, Correa A, Cotlarciuc I, Cruchaga C, Danesh J, de Bakker PIW, DeStefano AL, den Hoed M, Duan Q, Engelter ST, Falcone GJ, Gottesman RF, Grewal RP, Gudnason V, Gustafsson S, Haessler J, Harris TB, Hassan A, Havulinna AS, Heckbert SR, Holliday EG, Howard G, Hsu FC, Hyacinth HI, Ikram MA, Ingelsson E, Irvin MR, Jian X, Jiménez-Conde J, Johnson JA, Jukema JW, Kanai M, Keene KL, Kissela BM, Kleindorfer DO, Kooperberg C, Kubo M, Lange LA, Langefeld CD, Langenberg C, Launer LJ, Lee JM, Lemmens R, Leys D, Lewis CM, Lin WY, Lindgren AG, Lorentzen E, Magnusson PK, Maguire J, Manichaikul A, McArdle PF, Meschia JF, Mitchell BD, Mosley TH, Nalls MA, Ninomiya T, O'Donnell MJ, Psaty BM, Pulit SL, Rannikmäe K, Reiner AP, Rexrode KM, Rice K, Rich SS, Ridker PM, Rost NS, Rothwell PM, Rotter JI, Rundek T, Sacco RL, Sakaue S, Sale MM, Salomaa V, Sapkota BR, Schmidt R, Schmidt CO, Schminke U, Sharma P, Slowik A, Sudlow CLM, Tanislav C, Tatlisumak T, Taylor KD, Thijs VNS, Thorleifsson G, Thorsteinsdottir U, Tiedt S, Trompet S, Tzourio C, van Duijn CM, Walters M, Wareham NJ, Wassertheil-Smoller S, Wilson JG, Wiggins KL, Yang Q, Yusuf S; AFGen Consortium; Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium; International Genomics of Blood Pressure (iGEN-BP) Consortium; INVENT Consortium; STARNET, Bis JC, Pastinen T, Ruusalepp A, Schadt EE, Koplev S, Björkegren JLM, Codoni V, Civelek M, Smith NL, Trégouët DA, Christophersen IE, Roselli C, Lubitz SA, Ellinor PT, Tai ES, Kooner JS, Kato N, He J, van der Harst P, Elliott P, Chambers JC, Takeuchi F, Johnson AD; 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, Sanghera DK, Melander O, Jern C, Strbian D, Fernandez-Cadenas I, Longstreth WT Jr, Rolfs A, Hata J, Woo D, Rosand J, Pare G, Hopewell JC, Saleheen D, Stefansson K, Worrall BB, Kittner SJ, Seshadri S, Fornage M, Markus HS, Howson JMM, Kamatani Y, Debette S and Dichgans M. Multiancestry genome-wide association study of 520,000 subjects identifies 32 loci associated with stroke and stroke subtypes. Nature Genetics 2018; 50: 524-537.

[13] Mola-Caminal M, Carrera C, Soriano-Tárraga C, Giralt-Steinhauer E, Díaz-Navarro RM, Tur S, Jiménez C, Medina-Dols A, Cullell N, Torres-Aguila NP, Muiño E, Rodríguez-Campello A, Ois A, Cuadrado-Godia E, Vivanco-Hidalgo RM, Hernandez-Guillamon M, Solé M, Delgado P, Bustamante A, García-Berrocoso T, Mendióroz M, Castellanos M, Serena J, Martí-Fàbregas J, Segura T, Serrano-Heras G, Obach V, Ribó M, Molina CA, Alvarez-Sabín J, Palomeras E, Freijo M, Font MA, Rosand J, Rost NS, Gallego-Fabrega C, Lee JM, Heitsch L, Ibanez L, Cruchaga C, Phuah CL, Lemmens R, Thijs V, Lindgren A, Maguire J, Rannikmae K, Sudlow CL, Jern C, Stanne TM, Lorentzen E, Muñoz-Narbona L, Dávalos A, López-Cancio E, Worrall BB, Woo D, Kittner SJ, Mitchell BD, Montaner J, Roquer J, Krupinski J, Estivill X, Rabionet R, Vives-Bauzá C, Fernández-Cadenas I and Jiménez-Conde J. PATJ low frequency variants are associated with worse ischemic stroke functional outcome. Circ Res 2019; 124: 114-120.

[14] Söderholm M, Pedersen A, Lorentzen E, Stanne TM, Bevan S, Olsson M, Cole JW, Fernandez-Cadenas I, Hankey GJ, Jimenez-Conde J, Jood K, Lee JM, Lemmens R, Levi C, Mitchell BD, Norrving B, Rannikmäe K, Rost NS, Rosand J, Rothwell PM, Scott R, Strbian D, Sturm JW, Sudlow C, Traylor M, Thijs V, Tatlisumak T, Woo D, Worrall BB, Maguire JM, Lindgren A and Jern C; International Stroke Genetics Consortium, the NINDS-SiGN Consortium, and the Genetics of Ischaemic Stroke Functional Outcome (GISCOME) Network. Genome-wide association meta-analysis of functional outcome after ischemic stroke. Neurology 2019; 92: e1271-e1283.

[15] Adams HP Jr, Bendixen BH, Kappelle LJ, Biller J, Love BB, Gordon DL and Marsh EE 3rd. 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 1993; 24: 35-41.

[16] Traylor M, Malik R, Nalls MA, Cotlarciuc I, Radmanesh F, Thorleifsson G, Hanscombe KB, Langefeld C, Saleheen D, Rost NS, Yet I, Spector TD, Bell JT, Hannon E, Mill J, Chauhan G, Debette S, Bis JC, Longstreth WT Jr, Ikram MA, Launer LJ, Seshadri S; METASTROKE, UK Young Lacunar DNA Study, NINDS Stroke Genetics Network, Neurology Working Group of the CHARGE Consortium, Hamilton-Bruce MA, Jimenez-Conde J, Cole JW, Schmidt R, Słowik A, Lemmens R, Lindgren A, Melander O, Grewal RP, Sacco RL, Rundek T, Rexrode K, Arnett DK, Johnson JA, Benavente OR, Wassertheil-Smoller S, Lee JM, Pulit SL, Wong Q, Rich SS, de Bakker PI, McArdle PF, Woo D, Anderson CD, Xu H, Heitsch L, Fornage M, Jern C, Ste-

fansson K, Thorsteinsdottir U, Gretarsdottir S, Lewis CM, Sharma P, Sudlow CL, Rothwell PM, Boncoraglio GB, Thijs V, Levi C, Meschia JF, Rosand J, Kittner SJ, Mitchell BD, Dichgans M, Worrall BB and Markus HS; International Stroke Genetics Consortium. Genetic variation at 16q24.2 is associated with small vessel stroke. Ann Neurol 2017; 81: 383-394.

[17] Cheng YC, Stanne TM, Giese AK, Ho WK, Traylor M, Amouyel P, Holliday EG, Malik R, Xu H, Kittner SJ, Cole JW, O'Connell JR, Danesh J, Rasheed A, Zhao W, Engelter S, Grond-Ginsbach C, Kamatani Y, Lathrop M, Leys D, Thijs V, Metso TM, Tatlisumak T, Pezzini A, Parati EA, Norrving B, Bevan S, Rothwell PM, Sudlow C, Slowik A, Lindgren A, Walters MR; WTCCC-2 Consortium, Jannes J, Shen J, Crosslin D, Doheny K, Laurie CC, Kanse SM, Bis JC, Fornage M, Mosley TH, Hopewell JC, Strauch K, Müller-Nurasyid M, Gieger C, Waldenberger M, Peters A, Meisinger C, Ikram MA, Longstreth WT Jr, Meschia JF, Seshadri S, Sharma P, Worrall B, Jern C, Levi C, Dichgans M, Boncoraglio GB, Markus HS, Debette S, Rolfs A, Saleheen D and Mitchell BD. Genome-wide association analysis of young-onset stroke identifies a locus on chromosome 10q25 Near HABP2. Stroke 2016; 47: 307-316.

[18] 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 2016; 15: 695-707.

[19] Traylor M, Mäkelä KM, Kilarski LL, Holliday EG, Devan WJ, Nalls MA, Wiggins KL, Zhao W, Cheng YC, Achterberg S, Malik R, Sudlow C, Bevan S, Raitoharju E; METASTROKE, International Stroke Genetics Consortium, Wellcome Trust Case Consortium 2 (WTCCC2), Oksala N, Thijs V, Lemmens R, Lindgren A, Slowik A, Maguire JM, Walters M, Algra A, Sharma P, Attia JR, Boncoraglio GB, Rothwell PM, de Bakker PI, Bis JC, Saleheen D, Kittner SJ, Mitchell BD, Rosand J, Meschia JF, Levi C, Dichgans M, Lehtimäki T, Lewis CM and Markus HS. A novel MMP12 locus is associated with large artery atherosclerotic stroke using a genome-wide age-at-onset informed approach. PLoS Genet 2014; 10: e1004469.

[20] Traylor M, Farrall M, Holliday EG, Sudlow C, Hopewell JC, Cheng YC, Fornage M, Ikram MA, Malik R, Bevan S, Thorsteinsdottir U, Nalls MA, Longstreth W, Wiggins KL, Yadav S, Parati EA, Destefano AL, Worrall BB, Kittner SJ, Khan MS, Reiner AP, Helgadottir A, Achterberg S, Fernandez-Cadenas I, Abboud S, Schmidt R, Walters M, Chen WM, Ringelstein EB, O'Donnell M, Ho WK, Pera J, Lemmens R, Norrving B, Higgins P, Benn M, Sale M, Kuhlenbäumer G, Doney AS, Vicente AM, Delavaran H, Algra A, Davies G, Oliveira SA, Palmer CN, Deary I, Schmidt H, Pandolfo M, Montaner J, Carty C, de Bakker PI, Kostulas K, Ferro JM, van Zuydam NR, Valdimarsson E, Nordestgaard BG, Lindgren A, Thijs V, Slowik A, Saleheen D, Paré G, Berger K, Thorleifsson G; Australian Stroke Genetics Collaborative, Wellcome Trust Case Control Consortium 2 (WTCCC2), Hofman A, Mosley TH, Mitchell BD, Furie K, Clarke R, Levi C, Seshadri S, Gschwendtner A, Boncoraglio GB, Sharma P, Bis JC, Gretarsdottir S, Psaty BM, Rothwell PM, Rosand J, Meschia JF, Stefansson K, Dichgans M and Markus HS; 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 2012; 11: 951-962.

[21] International Stroke Genetics Consortium (ISGC); Wellcome Trust Case Control Consortium 2 (WTCCC2), Bellenguez C, Bevan S, Gschwendtner A, Spencer CC, Burgess AI, Pirinen M, Jackson CA, Traylor M, Strange A, Su Z, Band G, Syme PD, Malik R, Pera J, Norrving B, Lemmens R, Freeman C, Schanz R, James T, Poole D, Murphy L, Segal H, Cortellini L, Cheng YC, Woo D, Nalls MA, Müller-Myhsok B, Meisinger C, Seedorf U, Ross-Adams H, Boonen S, Wloch-Kopec D, Valant V, Slark J, Furie K, Delavaran H, Langford C, Deloukas P, Edkins S, Hunt S, Gray E, Dronov S, Peltonen L, Gretarsdottir S, Thorleifsson G, Thorsteinsdottir U, Stefansson K, Boncoraglio GB, Parati EA, Attia J, Holliday E, Levi C, Franzosi MG, Goel A, Helgadottir A, Blackwell JM, Bramon E, Brown MA, Casas JP, Corvin A, Duncanson A, Jankowski J, Mathew CG, Palmer CN, Plomin R, Rautanen A, Sawcer SJ, Trembath RC, Viswanathan AC, Wood NW, Worrall BB, Kittner SJ, Mitchell BD, Kissela B, Meschia JF, Thijs V, Lindgren A, Macleod MJ, Slowik A, Walters M, Rosand J, Sharma P, Farrall M, Sudlow CL, Rothwell PM, Dichgans M, Donnelly P and Markus HS. Genome-wide association study identifies a variant in HDAC9 associated with large vessel ischemic stroke. Nat Genet 2012; 44: 328-333.

[22] Holliday EG, Maguire JM, Evans TJ, Koblar SA, Jannes J, Sturm JW, Hankey GJ, Baker R, Golledge J, Parsons MW, Malik R, McEvoy M, Biros E, Lewis MD, Lincz LF, Peel R, Oldmeadow C, Smith W, Moscato P, Barlera S, Bevan S, Bis JC, Boerwinkle E, Boncoraglio GB, Brott TG, Brown RD Jr, Cheng YC, Cole JW, Cotlarciuc I,

Devan WJ, Fornage M, Furie KL, Grétarsdóttir S, Gschwendtner A, Ikram MA, Longstreth WT Jr, Meschia JF, Mitchell BD, Mosley TH, Nalls MA, Parati EA, Psaty BM, Sharma P, Stefansson K, Thorleifsson G, Thorsteinsdottir U, Traylor M, Verhaaren BF, Wiggins KL, Worrall BB, Sudlow C, Rothwell PM, Farrall M, Dichgans M, Rosand J, Markus HS, Scott RJ, Levi C and Attia J; 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 2012; 44: 1147-1151.

[23] Solé X, Guinó E, Valls J, Iniesta R and Moreno V. SNPStats: a web tool for the analysis of association studies. Bioinformatics (Oxford, England) 2006; 22: 1928-1929.

[24] Ward LD and Kellis M. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. Nucleic Acids Res 2012; 40: D930-934.

[25] eGTEx Project. Enhancing GTEx by bridging the gaps between genotype, gene expression, and disease. Nat Genet 2017; 49: 1664-1670.

[26] GTEx Consortium. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. Science 2015; 348: 648-660.

[27] Marigo V, Davey RA, Zuo Y, Cunningham JM and Tabin CJ. Biochemical evidence that patched is the Hedgehog receptor. Nature 1996; 384: 176-179.

[28] Nachtergaele S, Whalen DM, Mydock LK, Zhao Z, Malinauskas T, Krishnan K, Ingham PW, Covey DF, Siebold C and Rohatgi R. Structure and function of the Smoothened extracellular domain in vertebrate Hedgehog signaling. Elife 2013; 2: e01340.

[29] Rahi S and Mehan S. Understanding abnormal SMO-SHH signaling in autism spectrum disorder: potential drug target and therapeutic goals. Cell Mol Neurobiol 2020; [Epub ahead of print].

[30] Liu L, Zhao B, Xiong X and Xia Z. The neuroprotective roles of sonic hedgehog signaling pathway in ischemic stroke. Neurochem Res 2018; 43: 2199-2211.

[31] Yin S, Bai X, Xin D, Li T, Chu X, Ke H, Han M, Chen W, Li X and Wang Z. Neuroprotective effects of the sonic hedgehog signaling pathway in ischemic injury through promotion of synaptic and neuronal health. Neural Plasticity 2020; 2020: 8815195.

[32] Ji H, Miao J, Zhang X, Du Y, Liu H, Li S and Li L. Inhibition of sonic hedgehog signaling aggravates brain damage associated with the downregulation of Gli1, Ptch1 and SOD1 expression in acute ischemic stroke. Neurosci Lett 2012; 506: 1-6.

[33] Sims JR, Lee SW, Topalkara K, Qiu J, Xu J, Zhou Z and Moskowitz MA. Sonic hedgehog regulates ischemia/hypoxia-induced neural progenitor proliferation. Stroke 2009; 40: 3618-3626.

[34] Yamamoto F, Clausen H, White T, Marken J and Hakomori S. Molecular genetic basis of the histo-blood group ABO system. Nature 1990; 345: 229-233.

[35] Jenkins PV and O'Donnell JS. ABO blood group determines plasma von Willebrand factor levels: a biologic function after all? Transfusion 2006; 46: 1836-1844.

[36] Ward S, O'Sullivan J and O'Donnell JS. The relationship between ABO blood group, von Willebrand factor and primary hemostasis. Blood 2020; 136: 2864-2874.

[37] Paterson AD, Lopes-Virella MF, Waggott D, Boright AP, Hosseini SM, Carter RE, Shen E, Mirea L, Bharaj B, Sun L and Bull SB. Genome-wide association identifies the ABO blood group as a major locus associated with serum levels of soluble E-selectin. Arterioscler Thromb Vasc Biol 2009; 29: 1958-1967.

[38] Williams SR, Hsu FC, Keene KL, Chen WM, Dzhivhuho G, Rowles JL 3rd, Southerland AM, Furie KL, Rich SS, Worrall BB and Sale MM; GARNET (The Genomics and Randomized Trials Network) Collaborative Research Group. Genetic drivers of von willebrand factor levels in an ischemic stroke population and association with risk for recurrent stroke. Stroke 2017; 48: 1444-1450.

[39] Hackman A, Abe Y, Insull W, Pownall H, Smith L, Dunn K, Gotto AM and Ballantyne CM. Levels of soluble cell adhesion molecules in patients with dyslipidemia. Circulation 1996; 93: 1334-1338.

[40] Ma S, Tian XY, Zhang Y, Mu C, Shen H, Bismuth J, Pownall HJ, Huang Y and Wong WT. E-selectin-targeting delivery of microRNAs by microparticles ameliorates endothelial inflammation and atherosclerosis. Sci Rep 2016; 6: 22910.

[41] Prugger C, Luc G, Haas B, Morange PE, Ferrieres J, Amouyel P, Kee F, Ducimetiere P and Empana JP. Multiple biomarkers for the prediction of ischemic stroke. Arterioscler Thromb Vasc Biol 2013; 33: 659-666.

[42] Qi L, Cornelis MC, Kraft P, Jensen M, Van Dam RM, Sun Q, Girman CJ, Laurie CC, Mirel DB, Hunter DJ, Rimm E and Hu FB. Genetic variants in ABO blood group region, plasma soluble E-selectin levels and risk of type 2 diabetes. Hum Mol Genet 2010; 19: 1856-1862.