

Genetic polymorphisms in pri-let-7a-2 are associated with ischemic stroke risk in a Chinese Han population from Liaoning, China: a case-control study

<https://doi.org/10.4103/1673-5374.301019>

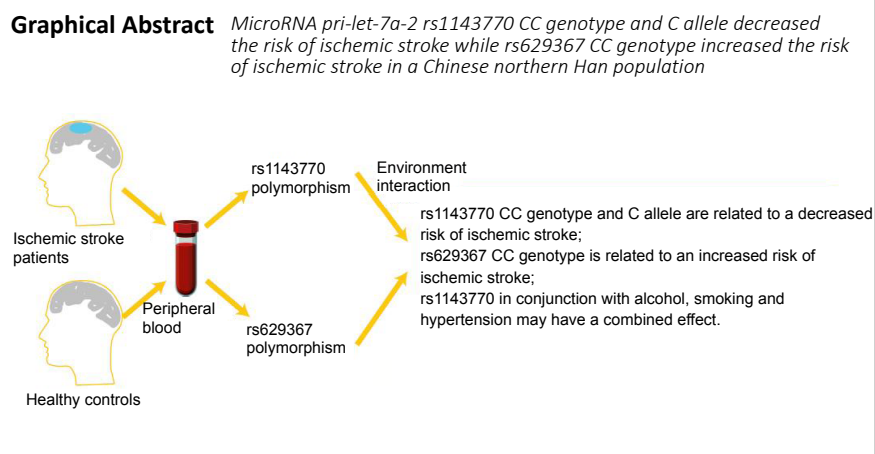
Date of submission: July 20, 2020

Date of decision: August 26, 2020

Date of acceptance: October 23, 2020

Date of web publication: December 12, 2020

Yu-ye Wang^{1,*}, He-yu Zhang^{2,#}, Wen-juan Jiang^{1,3}, Fang Liu¹, Lei Li¹, Shu-min Deng¹, Zhi-yi He¹, Yan-zhe Wang^{1,3,*}



Abstract

Ischemic stroke is a complicated disease, and its pathogenesis has been attributed to the occurrence of genetic polymorphisms. Evidence has suggested that the microRNA let-7a is involved in the pathogenesis of ischemic stroke. Pri-miRNA is the primary transcript, which undergoes several processing steps to generate pre-miRNA and, later, mature miRNAs. In this case-control study, we analyzed the distribution of pri-let-7a-2 variants in patients at a high risk for ischemic stroke and the interactions of pri-let-7a-2 variants and environmental factors. Blood samples and clinical information were collected from 1086 patients with ischemic stroke and 836 healthy controls between December 2013 and December 2015 at the First Affiliated Hospital of China Medical University. We found that the rs1143770 CC genotype and the C allele were associated with a decreased risk of ischemic stroke, whereas the rs629367 CC genotype was associated with an increased risk for ischemic stroke. Moreover, these two single-nucleotide polymorphisms were in linkage disequilibrium in this study sample. We analyzed gene-environment interactions and found that rs1143770 exerted a combined effect on the pathogenesis of ischemic stroke, together with alcohol use, smoking, and a history of hypertension. Therefore, the detection of pri-let-7a-2 polymorphisms may increase the awareness of ischemic stroke risk. This study was approved by the Institutional Ethics Committee of the First Affiliated Hospital of China Medical University, China (approval No. 2012-38-1) on February 20, 2012, and was registered with the Chinese Clinical Trial Registry (registration number: ChiCTR-COC-17013559) on December 27, 2017.

Key Words: case-control study; Chinese Han population; ischemic stroke; pri-microRNA; pri-let-7a-2; risk factors; rs1143770; rs629367; single-nucleotide polymorphism; single-nucleotide polymorphism-environment interaction

Chinese Library Classification No. R741; R394.5; Q343.1+5

Introduction

Stroke was identified as the primary cause of disability-adjusted life-years and years of life lost at the national level in China in 2017. Among 34 province-level administrative units in China, the age-standardized years of life lost per 100,000 population due to stroke in Liaoning Province was

significantly higher than that at the national level, and the ratio of observed age-standardized disability-adjusted life-years to expected life-years per 100 000 population associated with stroke in Liaoning Province was 1.81, ranking third in China (Zhou et al., 2019). Ischemic stroke (IS) is a complicated disease that involves both genetic and environmental

¹Department of Neurology, Key Laboratory for Neurological Big Data of Liaoning Province, The First Affiliated Hospital of China Medical University, Shenyang, Liaoning Province, China; ²Department of Neurology, The First Affiliated Hospital, Sun Yat-sen University; Guangdong Provincial Key Laboratory of Diagnosis and Treatment of Major Neurological Diseases, National Key Clinical Department and Key Discipline of Neurology, Guangzhou, Guangdong Province, China;

³Department of Pharmacology, School of Pharmacy, China Medical University, Shenyang, Liaoning Province, China

*Correspondence to: Yan-Zhe Wang, MD, PhD, yanzhewangcmu@126.com.

<https://orcid.org/0000-0003-1330-6298> (Yan-Zhe Wang)

#Both authors contributed equally to this study.

Funding: This study was supported by the National Natural Science Foundation of China, No. 81901189, and the Natural Science Foundation of Liaoning Province of China, No. 2019-BS-147 (both to YZW).

How to cite this article: Wang YY, Zhang HY, Jiang WJ, Liu F, Li L, Deng SM, He ZY, Wang YZ (2021) Genetic polymorphisms in pri-let-7a-2 are associated with ischemic stroke risk in a Chinese Han population from Liaoning, China: a case-control study. *Neural Regen Res* 16(7):1302-1307.

factors (Hatakeyama et al., 2020; Kaiser and Weat, 2020). The investigations of genetic variants associated with IS onset risk could help to clarify IS pathogenesis and provide new management and prevention methods for IS. Although previous studies have examined the roles played by single-nucleotide polymorphisms (SNPs) in stroke (Wang et al., 2016a, 2017, 2019; Zhang et al., 2016; Yang et al., 2020), the genetic variants correlated with IS susceptibility have not yet been clarified in the Chinese population.

MicroRNAs (miRNAs) are small, single-stranded RNAs of approximately 22 nucleotides in length. miRNAs inhibit post-transcriptional gene expression by interacting with target messenger RNAs (mRNAs) to induce their translational suppression or degradation (Bartel, 2004; Plasterk, 2006). The let-7 family is one of the most abundant miRNA families found in the encephalon (Rehfeld et al., 2015), and let-7a, a member of the let-7 family, has been shown to inhibit the expression of proinflammatory cytokines, decrease brain edema, and improve neurological deficits in mice with intracerebral hemorrhage (Yang et al., 2018). Pri-miRNA is the primary transcript, which undergoes several processing steps to generate pre-miRNA and, later, mature miRNA. In humans, chromosomes 9, 11, and 12 encode the let-7a precursors let-7a-1, let-7a-2, and let-7a-3, respectively, which all produce the same mature miRNA, let-7a (Lee et al., 2016). Although many previous studies have extensively investigated the biological roles of the let-7 family, few studies have addressed the functions of genetic variations in members of this family. Previous studies have also demonstrated that SNPs in miRNAs, including those found in primary (pri-), precursor, and mature miRNAs, can influence target selection or the expression level of miRNAs (Calin et al., 2005; Duan et al., 2007), which could result in increased disease susceptibility. In our study, using data from the National Center for Biotechnology Information bioinformatics databases, we identified two SNPs (pri-let-7a-2 rs629367 and rs1143770, www.ncbi.nlm.nih.gov/snp) with minor allele frequencies $\geq 5\%$ in the Chinese population, which were identified as tagSNPs in the 1000 Genomes Project database (www.1000genomes.org). To date, these two SNPs in pri-let-7a-2 have been correlated with either altered risk or prognosis for several diseases (Zhou et al., 2013; Xu et al., 2014; Shin et al., 2016; Fang et al., 2018; Heidari et al., 2018; Wang and Bi, 2018). However, their potential roles in IS onset and prognosis remain unclear.

Because IS is a multifactorial disease, and several studies have reported that gene-environment interactions between certain gene variants and risk factors are associated with IS risk (Wu et al., 2017; Feng et al., 2019; Liu et al., 2019), we also investigated whether environmental factors (such as alcohol use, smoking, and hypertension) could influence the risk of stroke among individuals carrying specific genetic variants.

A hospital-based, case-control study, which enrolled 1086 IS patients and 836 healthy control subjects, was performed to investigate the effects of two SNPs identified in pri-let-7a-2 (rs1143770 and rs629367; both located on chromosome 11) on IS risk. We aimed to investigate the associations between pri-let-7a-2 genetic polymorphisms at these loci and IS risk and to assess the gene-environment interactions in a Liaoning Province Han population.

Participants and Methods

Participants

Our study included 1086 patients and 836 age- and sex-matched healthy controls. Patient data were collected from December 2013 to December 2015 at the First Affiliated Hospital of China Medical University. The inclusion criteria for patients were: (1) sudden onset of focal neurological deficits; (2) deficits persisted for more than 24 hours; and (3) available brain imaging of the corresponding infarction. Patients were

excluded if they were diagnosed with transient ischemic attack, cardioembolism, cerebral trauma, cerebrovascular malformations, coagulation dysfunction, autoimmune diseases, tumors, chronic infectious diseases, or diseases of other systems. Controls were enrolled if they showed no evidence of stroke or other neurological diseases and had no diagnosed diseases of other systems (Wang et al., 2016a, 2017, 2019). Informed consent (**Additional File 1**) was signed by the entire study population. The severity of stroke and prognosis were measured using the National Institutes of Health Stroke Scale (Brott et al., 1989) at admission and the modified Rankin Scale (Bamford et al., 1989) 10 days after admission. The National Institutes of Health Stroke Scale evaluates the level of consciousness, gaze, visual field, facial palsy, motor arm, motor leg, limb ataxia, sensory, language dysarthria, and neglect. The modified Rankin Scale scores were defined as follows: 0, completely asymptomatic; 1, had symptoms but no significant dysfunction; 2, mild disability; 3, moderate disability; 4, severe disability; and 5, serious disability. The Institutional Ethical Committee of the First Affiliated Hospital of China Medical University approved this study on February 20, 2012 (No. 2012-38-1; **Additional File 2**), and the research was conducted according to the Code of Ethics of the World Medical Association (*Declaration of Helsinki*). The study protocol was registered with the Chinese Clinical Trial Registry (registration number ChiCTR-COC-17013559) on December 27, 2017. The writing and editing of this article were performed in accordance with the STrengthening the Reporting of OBservational studies in Epidemiology (STROBE) Statement. A flow chart can be found in **Figure 1**.

SNP selection

SNPs were searched using the National Center for Biotechnology Information bioinformatics databases, including the SNP database (dbSNP; www.ncbi.nlm.nih.gov/snp/?term=dbSNP). We selected SNPs according to the following criteria: a) located in the primary precursor area of the let-7 family (± 600 bp); (b) minor allele frequency > 0.05 in the Chinese Han population; and (c) identified as tagSNPs in the 1000 Genomes Project database. Two SNPs located in pri-let-7a-2 were identified: rs1143770 and rs629367. SNP ID numbers and detailed sequence data were extracted from the public dbSNP database (www.ncbi.nlm.nih.gov/SNP).

DNA extraction and genotyping

Genomic DNA was extracted from 200 μ L of ethylenediaminetetraacetic acid-anticoagulated peripheral blood using a DNA Purification Kit (Promega, Madison, WI, USA), and a SNaPshot reaction was performed, as previously described (Wang et al., 2016a, 2017). A SNaPshot Multiplex Kit (Applied Biosystems Co., Ltd., Foster City, CA, USA) was used for the genotyping analysis. **Table 1** shows the sequences of the polymerase chain reaction primers used in this study. The experimental results were analyzed on an ABI 3130XL DNA sequence detector and in GeneMapper 4.0 (Applied Biosystems Co., Ltd.).

Statistical analysis

Assuming a genotypic relative risk for a recessive model of 2, a minor allele frequency of 0.15, a 1.88% population prevalence of IS, and a type I error probability of 0.05, in the sample size of 1086 patient samples and 836 healthy controls, we would be unable to reject the null hypothesis in which the odds ratio (OR) = 1 with a power of 99.99%. We compared the distributions of demographic variables between patients and healthy controls using Pearson's chi-square test and examined differences in the genotypes and risk factors associated with various alleles and haplotypes. A goodness-of-fit chi-square test was used to test the Hardy-Weinberg equilibrium of each genotype. ORs and 95% confidence intervals (CIs) were

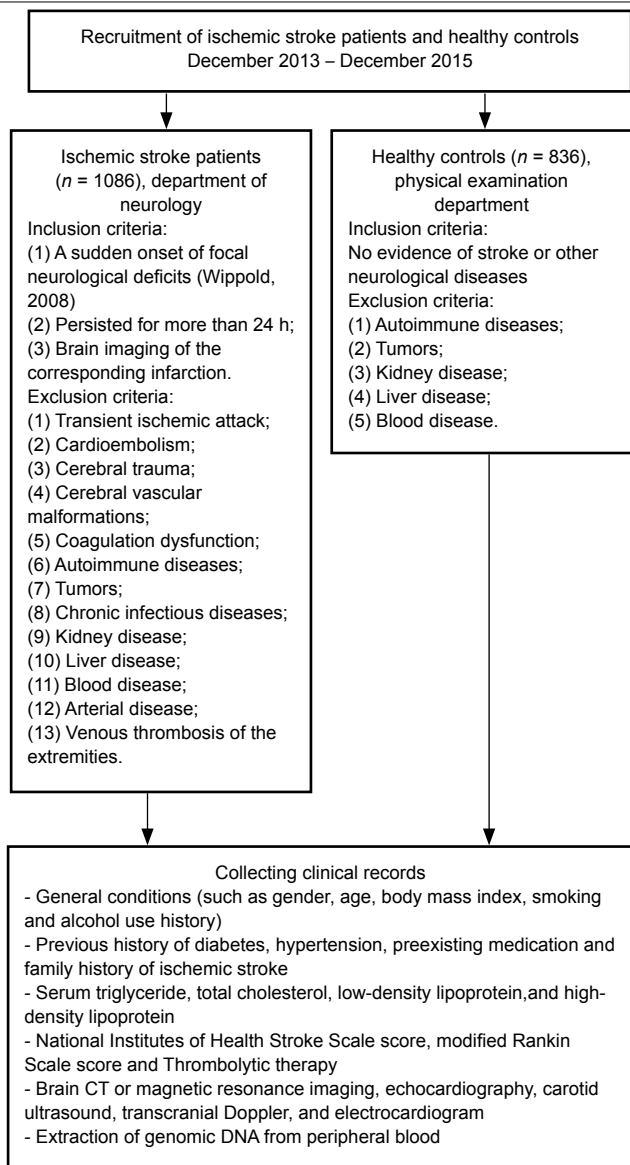


Figure 1 | Study flow chart.

calculated using unconditional logistic regression to estimate the association between IS and each genotype. Based on the acquired frequencies of the two SNPs, the SHEsis analysis platform was used to calculate the linkage disequilibrium indices (D' and r^2) and deduce the haplotype frequencies (Shi and He, 2005; Li et al., 2009). When $D' > 0.8$ or $r^2 > 0.4$, the result was considered positive, and the two calculated SNPs were considered to be in linkage disequilibrium in the population. Gene-environment and gene-gene interactions associated with IS risk were evaluated using a logistic regression model. Confounding factors, including sex, body mass index, age, hypertension, diabetes, alcohol use, smoking, and hyperlipidemia, were controlled during the genotype interaction analysis. All statistical analyses were two-sided and were performed using SPSS software, version 16.0 (SPSS Institute, Chicago, IL, USA). $P < 0.05$ was considered significant.

Results

Clinical characteristics of IS patients and controls

To investigate the potential associations between the rs1143770 and rs629367 SNPs and the risk of IS, clinical information and blood specimens were collected and analyzed from 1086 patients and 836 sex- and age-matched controls.

Table 1 | Primer sequences of pri-let-7a-2 single-nucleotide polymorphism

Single-nucleotide polymorphism	Primer sequence	Product size (bp)
rs1143770	Forward: 5'-CTT GGG ACT GGC CTT CTT GAA C-3' Reverse: 5'-CAC TGT GTT TTG GTT TAA GCC ATT TCT-3'	241
rs629367	Forward: 5'-CCA GGC AAG ACC AGA AAA CCT G-3' Reverse: 5'-AAT GTG AGG CTT TGA GGG ATT CAG-3'	212

The clinical characteristics of the patients and controls are summarized in **Table 2**. No significant differences in age ($P = 0.345$) or sex ($P = 0.835$) were observed between healthy controls and patients. However, compared with the control group, conventional risk factors for IS, such as body mass index, hypertension, diabetes mellitus, alcohol use, smoking, and family history of IS, were remarkably more prevalent in the patient group, and the serum total cholesterol, triglyceride, and low-density lipoprotein cholesterol levels were significantly higher in the patient group, whereas high-density lipoprotein cholesterol levels were significantly lower in the patient group ($P < 0.001$).

Allele and genotype distributions among IS patients and controls

All allele distributions conformed to the Hardy-Weinberg equilibrium (rs1143770: $P = 0.09$; rs629367: $P = 0.16$). The allele and genotype frequencies for rs1143770 and rs629367 among the 1086 patients with IS and the 836 control participants are shown in **Table 3**.

As presented in **Table 3**, individuals with the CC genotype of rs1143770 had a reduced IS risk compared with individuals with the homozygous wild-type TT genotype (adjusted OR = 0.586, 95% CI: 0.428–0.802, $P = 0.001$). Individuals with the CC genotype of rs1143770 had a significantly decreased risk compared with individuals with the TT + TC genotype in the recessive model (adjusted OR = 0.627, 95% CI: 0.481–0.816, $P = 0.001$). Moreover, the rs1143770 C allele was significantly correlated with a decreased risk of IS compared with the T allele (adjusted OR = 0.853, 95% CI: 0.750–0.969, $P = 0.014$). However, no significant differences between IS patients and controls were found in the dominant model ($P > 0.05$). The CC genotype of rs629367 was correlated with an increased IS risk compared with the AA genotype (adjusted OR = 1.665, 95% CI: 1.025–2.706, $P = 0.04$) and the AA + AC genotype (adjusted OR = 1.716, 95% CI: 1.064–2.768, $P = 0.027$) in the recessive model.

Haplotype distribution of IS patients and controls

The two SNPs were in linkage disequilibrium in this study population ($D' = 0.86$). Three of the four possible haplotypes had frequencies > 0.03 , among both controls and IS cases, and were included in further haplotype analyses. These three haplotypes represented 97.5% of the chromosomes in IS cases and 99.7% of the chromosomes in controls. No difference in overall haplotype distribution was identified between controls and cases. The haplotype analysis showed no significant differences in haplotype frequency, although a trend toward the increased frequency of the C-C haplotype of rs1143770-rs629367 was observed among patients compared with healthy controls (adjusted OR = 1.126, 95% CI: 0.990–1.280, $P = 0.07$; **Table 4**).

SNP-environment analysis of IS patients and controls

To assess prospective gene-environment interactions, correlations between the identified rs1143770 and rs629367 polymorphisms and exposure to the environmental factors

Table 2 | Characteristics of and risk factors for ischemic stroke

Variable	Patients (n = 1086)	Controls (n = 836)	P-value
Age (≤/ > 60 yr)	423 (39.0)/663 (61.0)	308 (36.8)/528 (63.2)	0.345
Sex (male/female)	608 (56.0)/478 (44.0)	636 (57.1)/478 (42.9)	0.835
BMI (≤/ > 22.9 kg/m ²)	555 (51.1)/531 (48.9)	540 (64.6)/296 (35.4)	< 0.001
Diabetes mellitus	313 (28.8)	62 (7.4)	< 0.001
Hypertension	706 (65.0)	164 (19.6)	< 0.001
History of smoking	439 (40.4)	129 (15.4)	< 0.001
History of alcohol use	219 (20.2)	97 (11.6)	< 0.001
Family history of ischemic stroke	93 (8.6)	36 (4.3)	< 0.001
NIHSS score*	7.64±3.98	–	–
mRS score*	2.85±1.02	–	–
Preexisting medication (aspirin)	123 (11.3)/963 (88.7)	–	–
Preexisting medication (statins)	61 (5.6)/420 (94.4)	–	–
Thrombolytic therapy	259 (23.8)/827 (76.2)	–	–
Triglyceride (≤/ > 1.7 mM)	666 (61.3)/420 (38.7)	574 (68.7)/262 (31.3)	< 0.001
Triglyceride (≤/ > 5.72 mM)	832 (76.6)/254 (23.4)	721 (86.2)/115 (13.8)	< 0.001
High-density lipoprotein (>/ ≤ 0.91 mM)	1006 (92.6)/80 (7.4)	808 (96.7)/28 (3.3)	< 0.001
Low-density lipoprotein (≤/ > 3.64 mM)	880 (81.0)/206 (19.0)	737 (88.2)/99 (11.8)	< 0.001
Fasting blood glucose (</ ≥ 7.0 mM)	648 (59.7)/438 (40.3)	742 (88.8)/94 (11.2)	< 0.001

*Data are expressed as the mean ± SD; other data are expressed as number (percentage) and analyzed by unconditional logistic regression. BMI: Body mass index; NIHSS: National Institutes of Health Stroke Scale; mRS: modified Rankin Scale.

were calculated using a logistic regression model. The results from the analysis of patients versus controls revealed that rs1143770 interacted with environmental factors (**Table 5**). When individuals carrying the wild-type TT genotype of rs1143770 without alcohol use were used as a reference, individuals in the patient group carrying the TC and CC genotypes with alcohol use presented increased risks of IS (adjusted except for alcohol use OR = 1.606, 95% CI: 1.094–2.358, *P* = 0.016 and adjusted except for alcohol use OR = 1.973, 95% CI: 1.168–3.332, *P* = 0.011, respectively; **Table 5**). The adjusted ORs for the risk of IS, calculated for the gene–environment interactions between the TT, CT, and CC genotypes of rs1143770 and smoking, were 3.419 (95% CI: 2.206–5.299; *P* < 0.001), 2.896 (95% CI: 2.063–4.064; *P* < 0.001), and 4.055 (95% CI: 2.503–6.509; *P* < 0.001), respectively (**Table 5**). The adjusted ORs for the risk of IS calculated for the gene–environment interactions between the TT, CT, and CC genotypes of rs1143770 and hypertension were 4.573 (95% CI: 3.089–6.770; *P* < 0.001), 6.744 (95% CI: 4.776–9.523; *P* < 0.001), and 4.588 (95% CI: 3.028–6.951; *P* < 0.001), respectively (**Table 5**).

Discussion

Stroke is one of the three leading causes of death worldwide, with a high incidence and a high fatality rate, and represents a tremendous social burden; therefore, better understanding which populations are susceptible to stroke is necessary for the prevention of stroke-associated damage (Chen et al., 2019). Recently, an increasing number of studies have shown that noncoding RNAs, especially miRNAs, play a key role in the development and occurrence of stroke (Liu et al., 2013; Zhao et al., 2013; Yang et al., 2017; Sonoda et al., 2019). The let-7 family is one of the most abundant miRNA families found in the brain (Rehfeld et al., 2015). Let-7a promotes microglial M2 polarization by targeting casein kinase 2 interacting protein

Table 3 | Allele and genotype frequencies of genetic polymorphisms among ischemic stroke patients and controls and their main effects on stroke risk

Single-nucleotide polymorphism	Patients (n = 1086)	Controls (n = 836)	OR (95% CI) ^a	P-value ^a
rs1143770				
Genotype				
TT (ref)	301(27.7)	208(24.9)	1.00 (ref.)	
TC	573(52.8)	422(50.5)	0.903 (0.699–1.168)	0.438
CC	212(19.5)	206(24.6)	0.586 (0.428–0.802)	0.001
Dominant effect				
TT (ref)	301(27.7)	208(24.9)	1.00 (ref.)	
TC+CC	785(72.3)	628(75.1)	0.797 (0.625–1.017)	0.068
Recessive effect				
TT+TC (ref)	874(80.5)	630(75.4)	1.00 (ref.)	
CC	212(19.5)	206(24.6)	0.627 (0.481–0.816)	0.001
rs1143770 allele				
T (ref)	1175(58.4)	838(54.5)	1.00 (ref.)	
C	997(41.6)	834(45.5)	0.853 (0.750–0.969)	0.014
rs629367				
genotype				
AA (ref)	665(61.2)	496(59.3)	1.00 (ref.)	
AC	348(32.0)	302(36.1)	0.922 (0.732–1.160)	0.487
CC	73(6.7)	38(4.5)	1.665 (1.025–2.706)	0.04
Dominant effect				
AA (ref)	665(61.2)	496(59.3)	1.00 (ref.)	
AC+CC	421(38.8)	340(40.7)	0.999 (0.802–1.245)	0.994
Recessive effect				
AA+AC (ref)	1013(93.3)	798(95.5)	1.00 (ref.)	
CC	73(6.7)	38(4.5)	1.716 (1.064–2.768)	0.027
rs629367 allele				
A (ref)	1678(77.3)	1294(77.4)	1.00 (ref.)	
C	494(22.7)	378(22.6)	1.008 (0.865–1.174)	0.92

^aAdjusted OR (95% CI) and *P*-value: Adjusted for age, sex, body mass index (BMI), diabetes mellitus, hypertension, history of smoking, history of alcohol use, and family history of ischemic stroke. Data are expressed as the number (percentage) and were analyzed by unconditional logistic regression. *P*-value is compared with the reference (ref) value for each type of data. CI: Confidence interval; SNP: single-nucleotide polymorphism; OR: odds ratio.

Table 4 | Haplotype frequencies in ischemic stroke patients and controls and their relationships with stroke risk

Haplotype	Patients (n = 1086)	Controls (n = 836)	OR (95% CI)	P-value
T-A	439 (20.2)	372 (22.3)	0.909 (0.778–1.063)	0.231
T-C	558 (25.7)	462 (27.6)	0.934 (0.809–1.079)	0.355
C-A	55 (2.5)	6 (0.3)	–	–
C-C	1120 (51.6)	832 (49.8)	1.126 (0.990–1.280)	0.07

The first letter represents the genotype of the rs1143770 allele, and the second letter represents the genotype of the rs629367. No difference in the overall haplotype distribution was observed between patients and controls (global test *P* = 0.189). CI: Confidence interval; OR: odds ratio.

1 following intracerebral hemorrhage (Yang et al., 2018) and suppresses α -synuclein-induced microglial inflammation by interacting with signal transducer and activator of transcription-3 in Parkinson's disease (Zhang et al., 2019). However, the potential role of let-7a in IS remains unclear. In the present study, we analyzed a northern Chinese Han sample consisting of 1086 cases and 836 controls without significant differences in age or sex and investigated the roles played by the rs1143770 and rs629367 SNPs in pri-let-7a-2 in IS.

Pri-miRNA is the primary miRNA transcript, and undergoes several processing steps to generate the mature miRNAs that are capable of interfering with messenger RNA targets. Calin et al. (2005) determined that a germline mutation in the primary

Table 5 | Association between rs1143770 polymorphisms and the risk of ischemic stroke, stratified by different environmental risks

Environment risk	rs1143770 genotype	Patients (n = 1086)	Controls (n = 836)	OR (95% CI) *	P-value
Alcohol use					
Status					
No	TT	253 (29.2)	183 (24.8)	1.00 (ref)	–
	TC	462 (53.3)	372 (50.3)	0.898 (0.711–1.135)	0.369
	CC	152 (17.5)	184 (24.9)	0.598 (.0448–0.796)	< 0.001
Yes	TT	48 (21.9)	25 (25.8)	1.389 (0.826–2.335)	0.215
	TC	111 (50.7)	50 (51.5)	1.606 (1.094–2.358)	0.016
	CC	60 (27.4)	22 (22.7)	1.973 (1.168–3.332)	0.011
Smoking					
Status					
No	TT	106 (16.4)	181 (25.6)	1.00 (ref)	–
	TC	358 (55.3)	351 (46.9)	0.975 (0.756–1.258)	0.847
	CC	183 (28.3)	175 (24.8)	0.560 (0.408–0.769)	< 0.001
Yes	TT	118 (26.9)	33 (25.6)	3.419 (2.206–5.299)	< 0.001
	TC	215 (49.0)	71 (55.0)	2.896 (2.063–4.064)	< 0.001
	CC	106 (24.1)	25 (19.4)	4.055 (2.503–6.569)	< 0.001
Hypertension					
Status					
No	TT	62 (16.3)	158 (23.5)	1.00 (ref)	–
	TC	195 (51.3)	350 (52.1)	0.716 (0.534–0.960)	0.026
	CC	123 (32.4)	164 (24.4)	0.486 (0.334–0.707)	< 0.001
Yes	TT	150 (21.2)	42 (25.6)	4.573 (3.089–6.770)	< 0.001
	TC	378 (53.5)	72 (43.9)	6.744 (4.776–9.523)	< 0.001
	CC	178 (25.2)	50 (30.5)	4.588 (3.028–6.951)	< 0.001

Data are expressed as the number (percentage) and were analyzed by logistic regression. P-value is compared with the reference (ref) value for each type of data. CI: Confidence interval; OR: odds ratio.

precursor of miR-16-1-miR-15a resulted in decreased miRNA expression levels, both *in vitro* and *in vivo*, which could lead to the development of chronic lymphocytic leukemia. In addition, rs3809783 (A > T) within pri-miR-10a has been correlated with recurrent spontaneous abortion by affecting mature miR-10a production (Li et al., 2016). It is reasonable to speculate that, sequence variations in pri-let-7a-2 could also influence the expression levels of let-7a, which could alter susceptibility to many diseases, including cancer (Mohr and Mott, 2015; Lee et al., 2016) and several neurological diseases (Yang et al., 2018; Zhang et al., 2019). Genetic polymorphisms in pri-let-7a-2 (rs1143770: C > T; rs629347: C > A) are believed to be involved in the development of several diseases (Zhou et al., 2013; Xu et al., 2014; Shin et al., 2016; Fang et al., 2018; Heidari et al., 2018; Wang and Bi, 2018). The CT/TT genotypes of rs1143770 have been associated with significantly increased risk for diabetic nephropathy (Zhou et al., 2013) and papillary thyroid carcinoma (Heidari et al., 2018). In addition, rs1143770 CT or TT has been associated with significantly improved overall survival and disease-free survival in patients with surgically resected non-small cell lung carcinoma (Shin et al., 2016). A previous study reported that the CC genotype of rs629347 was associated with increased risks of gastric cancer and atrophic gastritis and poor survival among gastric carcinoma patients (Xu et al., 2014). Based on the results reported by Wang et al., patients with primary liver cancer who carry the AC + CC genotype of rs629367 have worse prognosis after transcatheter arterial chemoembolization than patients with other genotypes (Wang and Bi, 2018). In addition, according to both the additive and dominant models, the rs629367 C allele may increase lung cancer risk in patients younger than 57 years old (Fang et al., 2018). Our findings suggested that the rs1143770 CC genotype and C allele are associated with reduced IS risk, whereas the rs629367 CC genotype is correlated with an increased IS risk, which is consistent with previous results.

Evidence has suggested potential mechanisms through which rs1143770 and rs629367 might affect IS occurrence. rs1143770 is a biologically plausible modulator of IS risk because variations at this location could result in partial loss of transcription factor cyclic adenosine monophosphate response element-binding protein (Zhou et al., 2013). The activation of the cyclic adenosine monophosphate/cyclic adenosine monophosphate response element-binding protein pathway has been reported to promote neuroplasticity (Wang et al., 2018) and ischemic post-conditioning (Wu et al., 2015), which are believed to be involved in the underlying mechanisms associated with effective IS treatment strategies. The rs629379 polymorphism most likely modulates IS risk through effects on mature let-7a production. *In vivo* studies have demonstrated that mature let-7a expression in subjects with the AA, CA, and CC genotypes of rs629367 was associated with a significant and gradual increase in the risk of atrophic gastritis after controlling for disease factors (tumor size, Borrmann type, TNM stage, depth of invasion and lymphatic metastasis). *In vitro* studies have shown that SGC7901 cells transfected with the pCMV-MIR-pri-let-7a-2 rs629367-C allele exhibited higher expression levels of mature let-7a than cells transfected with the A allele (Xu et al., 2014). Additionally, let-7 knockdown has been reported to inhibit the activation of the p38 mitogen-activated protein kinase and c-Jun N-terminal kinase signaling pathways, which reduced apoptosis and inflammatory reactions and exerted a neuroprotective effect in ischemia/reperfusion model rats (Wang et al., 2016b). Identifying these genetic polymorphisms may help predict the risk of IS, although further research remains necessary to explain the relationships between these SNPs and the occurrence/risk of IS and to study the biological functions of pri-let-7a-2 in IS.

Furthermore, through analyses of gene–environment interactions, we determined that when combined with alcohol use, smoking, and hypertension, rs1143770 may play an interactive role in the pathogenesis of IS. Alcohol use increases IS risk and is significantly correlated with the rs1143770 TC/CC genotype, whereas smoking and hypertension are significantly correlated with a high risk for IS regardless of genotype. When stratified by smoking exposure, the rs1143770 CC genotype was associated with a higher increased IS risk than the TT or TC genotypes, whereas when stratified by hypertension, the rs1143770 TC genotype had the highest OR. Individuals with specific genotypes should pay close attention to controlling their alcohol intake, quitting smoking, and maintaining healthy blood pressure. Therefore, the detection of pri-let-7a-2 promoter polymorphisms in patients with specific diseases or habits (such as hypertension, alcoholism, and smoking) may increase their awareness of their IS risk. High-risk patients are also recommended to undergo regular checkups and change daily routines to decrease the potential adverse effects associated with IS.

However, owing to some limitations, our conclusions should be interpreted with caution. First, our sample size was relatively small, and we lacked a large-scale survey because the participants in our analysis were limited to the northern Chinese Han population. Whether these results can be extended to other populations requires further research. Second, the study lacked *in vivo* and *in vitro* experiments to determine whether the genetic variations identified in this study affected IS risk by regulating the expression or functions of host genes. Therefore, validation studies that include large numbers of participants remain necessary. We also require further studies in which biological functional analyses are performed and studies performed in other ethnic populations to better understand the roles played by rs1143770 and rs629379 polymorphisms in the modulation of IS risk.

In summary, the results of our study suggested that pri-let-7a-2 polymorphisms are significantly correlated with the risk of IS in the northern Chinese Han population. We further

found that SNP-environment interactions may play an integral role in IS. In addition, our work indicated that alcohol use interacted with genetic variants in patients with IS, suggesting that possible interactions with alcohol use and various genomic loci should be validated in other cohorts. Moreover, the underlying mechanisms through which these two SNPs affect let-7a should be thoroughly investigated to better understand their functions in IS pathogenesis. Given the increasing socioeconomic burden associated with IS-induced disability and lethality, strategies designed to facilitate the early monitoring of high-risk groups and improve IS prevention will contribute greatly to solving this global problem. The presence of pri-let-7a-2 polymorphisms may indicate an increased risk of IS, and high-risk populations may be advised to undergo regular checkups to prevent IS occurrence.

Acknowledgments: Thank nurses from Department of Neurology, The First Affiliated Hospital of China Medical University for blood collection.

Author contributions: Conceptualization, writing review & editing, fundraising: YZW; investigation: WJJ, FL, LL, SMD; resources: YYW, HYZ, WJJ, FL, LL, SMD; methodology: YYW, HYZ, YZW; software, formal analysis, writing original draft: YYW, HYZ; validation: WJJ, FL, LL, SMD, YZW; supervision: ZYH. All authors approved the final version of the manuscript.

Conflicts of interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Financial support: This study was supported by the National Natural Science Foundation of China, No. 81901189, and the Natural Science Foundation of Liaoning Province of China, No. 2019-BS-147 (both to YZW). The funding sources had no role in study conception and design, data analysis or interpretation, paper writing or deciding to submit this paper for publication.

Institutional review board statement: The study was approved by the Institutional Ethical Committee of the First Affiliated Hospital of China Medical University, China on February 20, 2012 (approval No. 2012-38-1), and registered with the Chinese Clinical Trial Registry (registration number: ChiCTR-COC-17013559) on December 27, 2017.

Declaration of participant consent: The authors certify that they have obtained the consent forms from participants. In the form, participants have given their consent for their images and other clinical information to be reported in the journal. The participants understand that their names and initials will not be published.

Reporting statement: The writing and editing of the article were performed in accordance with the STrengthening the Reporting of OBservational Studies in Epidemiology (STROBE) Statement.

Biostatistics statement: The statistical methods of this study were reviewed by the epidemiologist of The First Affiliated Hospital of China Medical University, China.

Copyright license agreement: The Copyright License Agreement has been signed by all authors before publication.

Data sharing statement: Datasets analyzed during the current study are available from the corresponding author on reasonable request.

Plagiarism check: Checked twice by iThenticate.

Peer review: Externally peer reviewed.

Open access statement: This is an open access journal, and articles are distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as appropriate credit is given and the new creations are licensed under the identical terms.

Open peer reviewer: Diana Amantea, University of Calabria, Italy.

Additional files:

Additional file 1: Informed consent form (Chinese).

Additional file 2: Hospital ethics approval (Chinese).

Additional file 3: Open peer review report 1.

Additional file 4: Original data of the study.

References

Bamford JM, Sandercock PA, Warlow CP, Slattery J (1989) Interobserver agreement for the assessment of handicap in stroke patients. *Stroke* 20:828.

Bartel DP (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 116:281-297.

Brott T, Adams HP, Jr., Olinger CP, Marler JR, Barsan WG, Biller J, Spilker J, Holleran R, Eberle R, Hertzberg V, et al. (1989) Measurements of acute cerebral infarction: a clinical examination scale. *Stroke* 20:864-870.

Calin GA, Ferracin M, Cimmino A, Di Leva G, Shimizu M, Wojcik SE, Iorio MV, Visone R, Sever NI, Fabbri M, Iuliano R, Palumbo T, Pichiiorri F, Roldo C, Garzon R, Sevignani C, Rassenti L, Alder H, Volinia S, Liu CG, et al. (2005) A MicroRNA signature associated with prognosis and progression in chronic lymphocytic leukemia. *N Engl J Med* 353:1793-1801.

Chen Y, Wang J, Fan Y, Qin C, Xia X, Johnson J, Kallen AN (2019) Absence of the long noncoding RNA H19 results in aberrant ovarian STAR and progesterone production. *Mol Cell Endocrinol* 490:15-20.

Duan R, Pak C, Jin P (2007) Single nucleotide polymorphism associated with mature miR-125a alters the processing of pri-miRNA. *Hum Mol Genet* 16:1124-1131.

Fang C, Li XP, Chen YX, Wu NY, Yin JY, Zhang W, Zhou HH, Liu ZQ (2018) Functional miRNA variants affect lung cancer susceptibility and platinum-based chemotherapy response. *J Thorac Dis* 10:3329-3340.

Feng C, Yang Y, Yang S, Tu X, Wang Y, Song Y, Hui R, Zhang W (2019) Effect of gene-gene and gene-environment interaction on the risk of first-ever stroke and poststroke death. *Mol Genet Genomic Med* 7:e846.

Hatakeyama M, Ninomiya I, Kanazawa M (2020) Angiogenesis and neuronal remodeling after ischemic stroke. *Neural Regen Res* 15:16-19.

Heidari Z, Mohammadpour-Gharehbagh A, Eskandari M, Harati-Sadegh M, Salimi S (2018) Genetic polymorphisms of miRNA let-7a-2 and pri-mir-34b/c are associated with an increased risk of papillary thyroid carcinoma and clinical/pathological features. *J Cell Biochem doi: 10.1002/jcb.28152*.

Kaiser EE, West FD (2020) Large animal ischemic stroke models: replicating human stroke pathophysiology. *Neural Regen Res* 15:1377-1387.

Lee H, Han S, Kwon CS, Lee D (2016) Biogenesis and regulation of the let-7 miRNAs and their functional implications. *Protein Cell* 7:100-113.

Li Y, Wang XQ, Zhang L, Lv XD, Su X, Tian S, Liu CM, Ma X, Xia HF (2016) A SNP in pri-miR-10a is associated with recurrent spontaneous abortion in a Han-Chinese population. *Oncotarget* 7:8208-8222.

Li Z, Zhang Z, He Z, Tang W, Li T, Zeng Z, He L, Shi Y (2009) A partition-ligation-combination-subdivision EM algorithm for haplotype inference with multiallelic markers: update of the SHEsis (<http://analysis.bio-x.cn>). *Cell Res* 19:519-523.

Liu H, Luo D, Qiu Y, Huang Y, Chen C, Song X, Gao L, Zhou Y (2019) The effect of AMBP SNPs, their haplotypes, and gene-environment interactions on the risk of atherothrombotic stroke among the Chinese population. *Genet Test Mol Biomarkers* 23:487-494.

Liu X, Li F, Zhao S, Luo Y, Kang J, Zhao H, Yan F, Li S, Ji X (2013) MicroRNA-124-mediated regulation of inhibitory member of apoptosis-stimulating protein of p53 family in experimental stroke. *Stroke* 44:1973-1980.

Mohr AM, Mott JL (2015) Overview of microRNA biology. *Semin Liver Dis* 35:3-11.

Plasterk RH (2006) Micro RNAs in animal development. *Cell* 124:877-881.

Rehfeld F, Rohde AM, Nguyen DT, Wulczyn FG (2015) Lin28 and let-7: ancient milestones on the road from pluripotency to neurogenesis. *Cell Tissue Res* 359:145-160.

Shi YY, He L (2005) SHEsis, a powerful software platform for analyses of linkage disequilibrium, haplotype construction, and genetic association at polymorphism loci. *Cell Res* 15:97-98.

Shin KM, Jung DK, Hong MJ, Kang HJ, Lee WK, Yoo SS, Lee SY, Cha SI, Lee J, Kim CH, Seok Y, Cho S, Son JW, Lee EB, Jheon S, Kim YT, Park JY (2016) The pri-let-7a-2 rs1143770C>T is associated with prognosis of surgically resected non-small cell lung cancer. *Gene* 577:148-152.

Sonoda T, Matsuzaki J, Yamamoto Y, Sakurai T, Aoki Y, Takizawa S, Niida S, Ochiya T (2019) Serum MicroRNA-Based Risk Prediction for Stroke. *Stroke* 50:1510-1518.

Wang G, Bi C (2018) Correlations of pri-Let-7 gene polymorphisms with the recurrence and metastasis of primary liver cancer after transcatheter arterial chemoembolization. *Pathol Res Pract* 214:667-672.

Wang H, Gaur U, Xiao J, Xu B, Xu J, Zheng W (2018) Targeting phosphodiesterase 4 as a potential therapeutic strategy for enhancing neuroplasticity following ischemic stroke. *Int J Biol Sci* 14:1745-1754.

Wang Y, Yin X, Li L, Deng S, He Z (2016a) Association of apolipoprotein C3 genetic polymorphisms with the risk of ischemic stroke in the northern Chinese Han population. *PLoS One* 11:e0163910.

Wang Y, Liu F, Li L, Deng S, He Z (2017) The association between apolipoprotein A1-C3-A5 gene cluster promoter polymorphisms and risk of ischemic stroke in the northern Chinese Han population. *J Int Med Res* 45:2042-2052.

Wang YZ, Zhang HY, Liu F, Li L, Deng SM, He ZY (2019) Association between PPARG genetic polymorphisms and ischemic stroke risk in a northern Chinese Han population: a case-control study. *Neural Regen Res* 14:1986-1993.

Wang ZK, Liu FF, Wang Y, Jiang XM, Yu XF (2016b) Let-7a gene knockdown protects against cerebral ischemia/reperfusion injury. *Neural Regen Res* 11:262-269.

Wippold FJ 2nd. (2008) Focal neurologic deficit. *AJNR Am J Neuroradiol* 29:1998-2000.

Wu H, Yang SF, Dai J, Qiu YM, Miao YF, Zhang XH (2015) Combination of early and delayed ischemic postconditioning enhances brain-derived neurotrophic factor production by upregulating the ERK-CREB pathway in rats with focal ischemia. *Mol Med Res* 12:6427-6434.

Wu Z, Huang Y, Huang J, Fan L (2017) Impact of CRP gene and additional gene-smoking interaction on ischemic stroke in a Chinese Han population. *Neurol Res* 39:442-447.

Xu Q, Dong Q, He C, Liu W, Sun L, Liu J, Xing C, Li X, Wang B, Yuan Y (2014) A new polymorphism biomarker rs629367 associated with increased risk and poor survival of gastric cancer in Chinese by up-regulated miRNA-let-7a expression. *PLoS One* 9:e95249.

Yang W, Ma F, Wang L, He X, Zhang H, Zheng J, Wang Y, Jin T, Yuan D, He Y (2020) The association analysis between CYP24A1 genetic polymorphisms and the risk of ischemic stroke in Chinese Han population. *Brain Behav* 10:e01503.

Yang X, Tang X, Sun P, Shi Y, Liu K, Hassan SH, Stetler RA, Chen J, Yin KJ (2017) MicroRNA-15a/16-1 antagomir ameliorates ischemic brain injury in experimental stroke. *Stroke* 48:1941-1947.

Yang Z, Jiang X, Zhang J, Huang X, Zhang X, Wang J, Shi H, Yu A (2018) Let-7a promotes microglia M2 polarization by targeting CK1P-1 following ICH. *Immunol Lett* 202:1-7.

Zhang J, Dongwei Z, Zhang Z, Xinhui Q, Kunwang B, Guohui L, Jian D (2019) miR-let-7a suppresses α -Synuclein-induced microglia inflammation through targeting STAT3 in Parkinson's disease. *Biochem Biophys Res Commun* 519:740-746.

Zhang L, Yang J, Xue Q, Yang D, Lu Y, Guang X, Zhang W, Ba R, Zhu H, Ma X (2016) An rs13293512 polymorphism in the promoter of let-7 is associated with a reduced risk of ischemic stroke. *J Thromb Thrombolysis* 42:610-615.

Zhao H, Wang J, Gao L, Wang R, Liu X, Gao Z, Tao Z, Xu C, Song J, Ji X, Luo Y (2013) MiRNA-424 protects against permanent focal cerebral ischemia injury in mice involving suppressing microglia activation. *Stroke* 44:1706-1713.

Zhou J, Peng R, Li T, Luo X, Peng H, Zha H, Yin P, Wen L, Zhang Z (2013) A potentially functional polymorphism in the regulatory region of let-7a-2 is associated with an increased risk for diabetic nephropathy. *Gene* 527:456-461.

Zhou M, Wang H, Zeng X, Yin P, Zhu J, Chen W, Li X, Wang L, Wang L, Liu Y, Liu J, Zhang M, Qi J, Yu S, Afshin A, Gakidou E, Glenn S, Krish VS, Miller-Petrie MK, Montjoy-Venning WC, et al. (2019) Mortality, morbidity, and risk factors in China and its provinces, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet* 394:1145-1158.

P-Reviewer: Amantea D; C-Editor: Zhao M; S-Editors: Yu J, Li CH; L-Editors: Giles L, Yu J, Song LP; T-Editor: Jia Y