Association of MicroRNA Biogenesis Genes Polymorphisms with Risk of Large Artery Atherosclerosis Stroke

Mengmeng Wang1,3 · Jinyu Gu2 · Chong Shen4 · Wuzhuang Tang5 · Xiaoru Xing6 · Zhizhong Zhang1,2 · Xinfeng Liu1,2

Received: 18 November 2020 / Accepted: 9 February 2021 / Published online: 23 February 2021
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC part of Springer Nature 2021

Abstract

Convincing evidence has shown that microRNAs (miRNAs) are involved in the pathogenesis of stroke. This study aimed to examine whether miRNA biogenesis genes polymorphisms are associated with risk of large artery atherosclerosis (LAA) stroke. Three polymorphisms (DROSHA rs10719 T>C, RAN rs3803012 A>G, and PIWIL1 rs10773771 C>T) were screened by certain criteria. A total of 1,785 (710 cases and 1,075 controls) study subjects were included in this study. We found that rs10773771 CC genotype was associated with a decreased risk of LAA stroke (CC vs. TT/CT: OR 0.63, 95% CI 0.46–0.86, \(P = 3 \times 10^{-3}\)). In silico analysis suggested that rs10773771 can change the mRNA secondary structure of PIWIL1 and affect the binding of the miRNAs and regulatory motifs to the 3′-UTR of PIWIL1. Expression quantitative trait loci analysis showed that rs10773771 could change the expression of PIWIL1 in human skin (\(P = 1.534 \times 10^{-10}\)) and thyroid tissues (\(P = 4.869 \times 10^{-6}\)). These findings suggested that PIWIL1 rs10773771 may be associated with a decreased risk of LAA stroke.

Keywords Polymorphism · Stroke · MicroRNA biogenesis genes · Subtype

Introduction

Stroke is a complex disease caused by environmental and genetic factors. Existing studies have shown that genetic factors play a substantial role in the risk of stroke (Bak et al. 2002; Jerrard-Dunne et al. 2003). About 80% of stroke is ischemic stroke (IS). The most common subtypes of IS are large artery atherosclerosis (LAA), cardioembolic stroke (CES), and small vessel disease (SVD) (Jerrard-Dunne et al. 2003). The genetic heritability varies by different subtypes of IS, of which LAA has the highest heritability of all IS subtypes (40.3%) (Bevan et al. 2012).

MicroRNAs (miRNAs) are a class of small noncoding RNAs that can bind to the 3′-untranslated regions (3′-UTRs) in mRNAs regulating their expression (Ambros 2004). Increasing evidences suggest that miRNAs have been involved in the pathogenesis of many diseases, such as cancer (Lu et al. 2005), schizophrenia (Beveridge et al. 2010), Parkinson’s diseases (Kim et al. 2007), and stroke (Khoshnam et al. 2017). The biosynthesis of miRNAs is a complex process involving multiple miRNA biogenesis genes (Ambros 2004). Briefly, primary miRNA is produced in the nucleus by RNA polymerase II. Then, the primary miRNA is further processed into the precursor miRNA.
Hypertension was defined as systolic pressure ≥ 140 mmHg or diastolic pressure ≥ 90 mmHg on at least three separate occasions or currently taking anti-hypertensive medications or has a history of hypertension. Diabetes mellitus was defined as fasting glucose ≥ 126 mg/dl or taking anti-diabetic medications or has a history of diabetes.

We conducted this case–control study to investigate whether there is an association between three known miRNA biogenesis genes polymorphisms (DROSHA rs10719 T > C, RAN rs3803012 A > G, and PIWIL1 rs10773771 C > T) and LAA stroke risk.

Materials and Methods

Study Subjects

This case-control study was approved by the institutional review boards of the local participating hospitals (2017NZGKJ-041). Written informed consent was obtained from all participants. This is a case–control study designed with samples from the Chinese Han population. Our study comprised 710 LAA patients and 1,075 controls. All LAA patients had a focal neurologic deficit lasting > 24 h, which were confirmed by computed tomography (CT) or magnetic resonance imaging (MRI). All LAA patients were diagnosed according to the Trial of Org 10172 in Acute Stroke Treatment (TOAST) criteria (Adams et al. 1993). All patients with tumor, autoimmune diseases or hemorrhagic diseases were excluded. The healthy controls were recruited from those presenting at local hospitals for health physical examinations during the same period. The controls did not have neurological diseases and cerebrovascular disease as well as cardiovascular diseases according to clinical examinations or history taken. Risk factors of LAA stroke patients and controls were determined according to established criteria. Hypertension was defined as systolic pressure ≥ 140 mmHg and diastolic pressure ≥ 90 mmHg on at least three separate occasions or currently taking anti-hypertensive medications or has a history of hypertension. Diabetes mellitus was defined as fasting glucose ≥ 126 mg/dl or taking anti-diabetic medications or has a history of diabetes.

SNP Selection and Genotyping

The single-nucleotide polymorphisms (SNPs) were selected from eight known miRNA biogenesis genes (DROSHA, DGCR8, RAN, XPOS, DICER, PIWIL1, GEMIN3, GEMIN4). The selection criteria were as follows: (1) in the 3′-UTR region; (2) the minor allele frequency of each SNP ≥ 0.05 in 1000 Genome CHB (Han Chinese in Beijing) data on the GRCh37 reference assembly (www.1000genomes.org/); (3) each SNP can influence the miRNA binding predicted by miRanda (John et al. 2004) and miRbase (Griffiths-Jones et al. 2008) database using the SNPinfo Web Server (https://snpinfo.niehs.nih.gov/snpinfo/snpfunc.html) (Xu and Taylor 2009). A total of three SNPs (DROSHA rs10719, RAN rs3803012, and PIWIL1 rs10773771) met the selection criteria described above and were included in this study. Genomic DNA was isolated from whole blood. Genotyping for DROSHA rs10719, RAN rs3803012, and PIWIL1 rs10773771 was conducted using SNPscan technology, with technical assistance from Center for Genetic and Genomic Analysis, Genesky Biotechnologies Inc. (Shanghai, China) (Supplementary Figs. 1–3). For quality control and validation purposes, a random 5% of sample was repeated genotyping to check for consistency, and the results were 100% consistent.

Statistical Analysis

Categorical variables were presented as number (percentage) and continuous variables as mean (SD, standard deviation). Categorical variables were compared with Chi-square test and continuous variables with Student’s t test. The association between microRNA biogenesis genes polymorphisms and LAA stroke was estimated with the odds ratio (OR) and 95% confidence interval (CI) by multivariate logistic regression model, adjusting for age, sex, hypertension and diabetes. The Hardy–Weinberg equilibrium, the association between these polymorphisms and LAA stroke were calculated by SNPStats web tool (http://bioinfo.iconcologia.net/SNPstats) (Sole et al. 2006). Goodness of fit was evaluated using Akaike information criterion (AIC) for selecting the best genetic model to the SNPs (Sole et al. 2006).

All reported P values are two-sided. P < 0.05 was deemed as statistical significance unless otherwise specified. All analyses were performed with IBM SPSS Statistics version 23.0 (Armonk, NY: IBM Corp.).

In Silico Analysis

To evaluate the potential function of the 3′-UTR SNP rs10773771, we conducted in silico analysis using RNAfold Web server (http://rna.tbi.univie.ac.at/) (Hofacker 2003), SNPinfo Web server (https://snpinfo.niehs.nih.gov/snpinfo/snpfunc.html) (Xu and Taylor 2009) and HaploReg v4.1 (https://pubs.broadinstitute.org/mammals/haploreg/haploreg.php) (Ward and Kellis 2016; Kheradpour and Kellis 2014).
Expression Quantitative Trait Loci (eQTL)

For SNP rs10773771, we queried publicly available eQTL database GTEx V7 (https://www.gtexportal.org/home/) (The Genotype-Tissue Expression (GTEx) project 2013).

### Results

**miRNA Biogenesis Genes Polymorphisms and the Risk of LAA Stroke**

The clinical characteristics of LAA patients and controls in the study are summarized in Table 1. Distribution of the genotypes in the control groups was not deviated from Hardy–Weinberg equilibrium ($P > 0.05$ for all 3 SNPs). As shown in Table 2, no significant associations between rs10719, rs3803012 and LAA stroke were detected in the multivariate logistic regression analysis. However, the rs10773771 CC genotype was significantly associated with a decreased risk of LAA stroke in both codominant ($P = 0.014$) and recessive model ($P = 3 \times 10^{-3}$). According to the AIC values, recessive model is the best-fit model for rs10773771. In the recessive model, compared with TT/CT genotype, rs10773771 CC genotype was associated with significantly decreased risk of LAA stroke (OR 0.63, 95% CI 0.46–0.86, $P = 3 \times 10^{-3}$). Further stratification analysis suggested that this decreased risk was only observed in males (OR 0.63, 95% CI 0.42–0.96, $P = 0.032$), those aged above 66

### Table 1 Baseline information for the study participants

| Variables | Controls $N = 1075$ | Cases $N = 710$ | $P$ |
|-----------|---------------------|----------------|-----|
| Age (year, SD) | 65.79 (8.28) | 66.32 (9.93) | 0.242 |
| Sex | | | 0.460 |
| Male | 630 (58.6) | 429 (60.4) | |
| Female | 445 (41.4) | 281 (39.6) | |
| Hypertension | | | |
| No | 577 (53.7) | 192 (27.0) | < 0.001 |
| Yes | 498 (46.3) | 518 (73.0) | |
| Diabetes | | | < 0.001 |
| No | 1039 (96.7) | 527 (74.2) | |
| Yes | 36 (3.3) | 183 (25.8) | |

### Table 2 Association between SNPs and LAA stroke

| Model | Control (n, %) | Case (n, %) | COR | 95%CI | $P$ | AOR | 95%CI | $P^*$ | AIC |
|-------|----------------|-------------|-----|-------|-----|-----|-------|-------|-----|
| rs10719 | | | | | | | | | |
| Codominant | | | | | | | | | |
| T/T | 545 (50.7) | 374 (52.7) | 1.00 | 1.00 | 2123.6 |
| T/C | 435 (40.5) | 277 (39.0) | 0.93 | 0.76–1.13 | 0.464 | 0.83 | 0.67–1.03 | 0.091 |
| C/C | 95 (8.8) | 59 (8.3) | 0.91 | 0.64–1.28 | 0.577 | 0.83 | 0.57–1.21 | 0.333 |
| Dominant | | | | | | | | | |
| T/T | 545 (50.7) | 374 (52.7) | 1.00 | 1.00 | 2121.6 |
| T/C–C/C | 530 (49.3) | 336 (47.3) | 0.92 | 0.76–1.12 | 0.413 | 0.83 | 0.67–1.02 | 0.074 |
| Recessive | | | | | | | | | |
| T/T–T/C | 980 (91.2) | 651 (91.7) | 1.00 | 1.00 | 2124.4 |
| C/C | 95 (8.8) | 59 (8.3) | 0.94 | 0.67–1.31 | 0.698 | 0.90 | 0.62–1.30 | 0.574 |
| rs10773771 | | | | | | | | | |
| Codominant | | | | | | | | | |
| T/T | 380 (35.4) | 261 (36.8) | 1.00 | 1.00 | 2117.7 |
| C/T | 520 (48.4) | 371 (52.2) | 1.04 | 0.85–1.28 | 0.718 | 1.06 | 0.85–1.33 | 0.593 |
| C/C | 175 (16.3) | 78 (11.0) | 0.65 | 0.48–0.88 | 0.006 | 0.65 | 0.47–0.92 | 0.014 |
| Dominant | | | | | | | | | |
| T/T | 380 (35.4) | 261 (36.8) | 1.00 | 1.00 | 2124.6 |
| C/T–C/C | 695 (64.7) | 449 (63.2) | 0.94 | 0.77–1.15 | 0.543 | 0.96 | 0.77–1.19 | 0.715 |
| Recessive | | | | | | | | | |
| T/T–C/T | 900 (83.7) | 632 (89.0) | 1.00 | 1.00 | 2116.0 |
| C/C | 175 (16.3) | 78 (11.0) | 0.63 | 0.48–0.84 | 0.002 | 0.63 | 0.46–0.86 | 0.003 |
| rs3803012 | | | | | | | | | |
| Codominant | | | | | | | | | |
| A/A | 964 (89.7) | 660 (93.0) | 1.00 | 1.00 | 2124.4 |
| G/A | 110 (10.2) | 50 (7.0) | 0.66 | 0.47–0.94 | 0.021 | 0.78 | 0.53–1.12 | 0.179 |
| G/G | 1 (0.1) | 0 (0.0) | | | | | | | |

*COR crude odd ratio, AOR adjusted odd ratio

*Adjusted by age, sex, hypertension, and diabetes
In this study, we evaluated the associations between three miRNA biogenesis genes polymorphisms (DROSHA rs10719 T>C, RAN rs3803012 A>G, and PIWIL1 rs10773771 C>T) and the risk of LAA stroke. To the best of our knowledge, this is the first study to report the association between rs10773771 and LAA stroke. These findings indicated that genetic variant at PIWIL1 contribute to the development of LAA stroke.

There is mounting evidence that miRNAs have been involved in the pathogenesis of stroke (Khoshnam et al. 2017; Rink and Khanna 2011; Mirzaei et al. 2018). PIWIL1 plays an important role in stem cell renewal, division and RNA silencing (Hutvagner and Simard 2008). Moreover, it can also take part in the process of microRNA biogenesis that pre-miRNAs are processed to produce the mature miRNAs (Bartel 2004). Given the function of PIWIL1 in miRNA biogenesis and the involvement of miRNA in the pathogenesis of stroke, SNPs in PIWIL1 may affect the risk of stroke through regulating its function. The SNP rs10773771, which reached statistical significance in this study, is located in the 3′-UTR of PIWIL1. Furthermore, we performed a number of complementary in silico analysis to predict the potential function of the 3′-UTR SNP rs10773771. The SNP rs10773771 can change the mRNA secondary structure of PIWIL1. Additionally, the rs10773771 can influence the combination of miRNAs to the 3′-UTR of PIWIL1 using the SNPInfo web server. The in-silico results suggested that the rs10773771 could affect the combination of three miRNAs (hsa-miR-1264, hsa-miR-340, hsa-miR-590-3p; Table 3) to the 3′-UTR of PIWIL1. In addition, based on the HaploReg v4.1, we identified that rs10773771 is likely to alter the binding affinity of regulatory motifs HNF1, Irx, Pou2f2, Sp100 and TATA (Table 4).

**eQTL Analysis**

Based on gene expression data extracted from GTEx, rs10773771 was identified to be an eQTL for PIWIL1 in human skin ($P=1.534 \times 10^{-10}$; Fig. 1b) and thyroid tissues ($P=4.869 \times 10^{-6}$; Fig. 1c).

**Discussion**

In this study, we evaluated the associations between three miRNA biogenesis genes polymorphisms (DROSHA rs10719 T>C, RAN rs3803012 A>G, and PIWIL1 rs10773771 C>T) and the risk of LAA stroke. To the best of our knowledge, this is the first study to report the association between rs10773771 and LAA stroke. These findings indicated that genetic variant at PIWIL1 contribute to the development of LAA stroke.
of stroke, such as smoking and hyperlipoidemia were not included in the study, which may introduce additional biases. (2) Our study was a hospital-based case–control study, and the selection bias cannot be eliminated.

In summary, we have identified an association between rs10773771 and LAA stroke. The underlying mechanisms remain to be elucidated in the future.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10571-021-01057-8.

Acknowledgements This study was supported by National Natural Science Foundation of China (Grant Nos. 81771285 and 81530038).

Author contributions MW: study design, statistical analysis, and manuscript drafting. JG: study design, statistical analysis, manuscript drafting. CS: data collection. WT: statistical analysis. XX: statistical analysis. ZZ: study design, interpretation of results, and critical revision of manuscript. XL: study design, interpretation of results, and critical revision of manuscript.

Compliance with Ethical Standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical Approval This study was approved by the institutional review boards of the local participating hospitals.

Informed consent Informed consent was obtained from all participants.

References

Adams HP, Bendixen BH, Kappelle LJ, Biller J, Love BB, Gordon DL, Marsh EE (1993) Classification of subtype of acute ischemic stroke. Definitions for use in a multicenter clinical trial. TOAST. Trial of Org 10172 in Acute Stroke Treatment. Stroke 24(1):35–41
Ambros V (2004) The functions of animal microRNAs. Nature 431(7006):350–355. https://doi.org/10.1038/nature02871
Bak S, Gaist D, Sindrup SH, Skytte A, Christensen K (2002) Genetic liability in stroke: a long-term follow-up study of Danish twins. Stroke 33(3):769–774
Bartel DP (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. Cell 116(2):281–297
Bevan S, Taylor M, Adib-Samii P, Malik R, Paul NL, Jackson C, Farrall M, Rothwell PM, Sudlow C, Dickhans M, Markus HS (2012) Genetic heritability of ischemic stroke and the contribution of previously reported candidate gene and genomewide associations. Stroke 43(12):3161–3167. https://doi.org/10.1161/STROKEAHA.112.665760
