Genetic polymorphisms and related risk factors of ischemic stroke in a Mongolian population in China

Q. Wu¹, H. Wu¹, Yu.L. Orlov², G. Geggentana¹, W. Huo¹, A.O. Bragin², N. Wu¹, S. Suyalatu¹, F. Zhao¹, J. Zhao¹, L.E. Tabikhanova², M. Chen³, H. Bai⁴

1 Inner Mongolia University for the Nationalities, Tongliao, China
2 Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia
3 Novosibirsk State University, Novosibirsk, Russia
4 Northern Virginia Community College, Virginia, USA
5 Department of Bioinformatics, College of Life Sciences, Zhejiang University, Hangzhou, China

Ischemic stroke is caused by an interruption in the flow of blood to the brain and a risk factor for death and disability. Recent genome-wide association studies have identified more than 40 common sequence variants associated with ischemic stroke. However, the results are not always the same in populations with different genetic backgrounds. In the present study, we evaluated a hypothesis that a North Asian population living in a geographic area with unusually harsh environmental conditions would develop unique genetic risks. We investigated the candidate genes for ischemic stroke and risk factors in a Chinese Mongolian population which has not been explored previously.

A total of 167 stroke cases and 176 controls were included in the study. Genotyping was performed by amplicon sequencing. The association was detected with single nucleotide polymorphisms (SNPs) located within genes NINJ2 (rs12425791) and ALDH2 (rs2238151) as well as intergenic rs9536591 were significantly associated with ischemic stroke, of which SNP rs12425791 of the NINJ2 gene was the strongest association. ALDH2 gene encodes mitochondrial aldehyde dehydrogenase, involved in the oxidative pathway of alcohol metabolism. Sex, age, body mass index and high blood pressure might be the risk factors. The current work also demonstrated genetic heterogeneity exists between Chinese and other populations. Our study provided the new insights into the genetic basis and environmental factors of ischemic stroke in Mongolian population.

Key words: population genetics; human genome; SNPs; Chinese Mongolian population; ischemic stroke.

Genетический полиморфизм и сопутствующие факторы риска ишемического инсульта в монгольской популяции Китая

Ч. У¹, Х. У¹, Ю.Л. Орлов², Г. Гегентана¹, В. Хуо¹, А.О. Брагин², Н. У¹, С. Суялату¹, Ф. Жао¹, Ж. Жао¹, Л.Э. Табиханова², М. Чен³, Х. Бай⁴

1 Национальный университет Внутренней Монголии, Тунхуа, Китай
2 Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
3 Новосибирский государственный университет, Новосибирск, Россия
4 Колледж Северной Вирджинии, Вирджиния, США
5 Отдел биоинформатики, Колледж естественных наук, Университет Чжецзянь, Ханьчоу, Китай

Ишемический инсульт, вызванный прерыванием потока крови в мозг, представляет серьезную медицинскую проблему, являясь фактором риска смерти и потери трудоспособности. Недавние исследования геномных ассоциаций (genome-wide association studies – GWAS) определили более 40 общих вариантов геномных последовательностей, связанных с ишемическим инсультом. Однако результаты не могут быть одинаково применимы для различных мировых популяций. В данной работе мы рассматриваем гипотезу о том, что население Северной Азии, живущее в суровых условиях окружающей среды, может иметь уникальный спектр генов риска развития заболеваний сердечно-сосудистой системы. Исследованы генов-кандидаты предрасположенности к ишемическому инсульту в популяции монголов Китая, на которых не была изучена ранее. В исследование были включены в общей сложности 167 больных с инсультом и 176 здоровых лиц. Генотипирование выполнено с помощью секвенирования ампликонов. Обнаружена ассоциация развития ишемического инсульта с одиночными нуклеотидными полиморфизмами в генах NINJ2 (rs12425791) и ALDH2 (rs2238151), а также в межгенном участке rs9536591. Среди генов-кандидатов наибольший вклад в развитие инсульта вносит NINJ2 (rs12425791). Ген ALDH2 кодирует митохондриальный фермент альдегиддегидрогеназу, вовлеченную в оксидативный путь метаболизма алкоголя. Пол, возраст, индекс массы тела и высокое кровяное давление также могут быть факторами риска. Данная работа показала существующую гетерогенность между китайской и другими популяционными выборками. Наше исследование предлагает новый взгляд на взаимодействие генотипа и факторов окружающей среды при развитии ишемического инсульта в монгольском населении.

Ключевые слова: популяционная генетика; геном человека; одинонуклеотидные замены; монгольская популяция Китая; ишемический инсульт.
Ischemic stroke, accounts for about 80% of the stroke population (Dichgans, 2007; Mozaffarian et al., 2016), is a leading cause of mortality and morbidity and presents a serious and growing threat to public health (Krupinski et al., 1994; Lai et al., 1994). Ischemic stroke is a multi-factorial disorder (Hassan, Markus, 2000; Goldstein et al., 2006) associated closely with conventional vascular risk factors and genetic factors which have been accepted as important risk contributors to the development of ischemic stroke (Dichgans, 2007). Animal and epidemiological study has indicated that ischemic stroke has obvious genetic predisposition (Hassan, Markus, 2000).

With the advances of study on ischemic stroke, many susceptibility genes related to stroke have been identified using candidate gene association studies and genome-wide association studies (GWAS) (Matar et al., 2007; Ding et al., 2011; Wan et al., 2011; Bellenguez et al., 2012). The most susceptibility genes for ischemic stroke mainly are related with lipid metabolism, inflammation, renin-angiotensin-aldosterone system, homocysteine metabolism, nitric oxide synthase, and heart disease (Hassan, Markus, 2000; Lindsberg, Grau, 2003; Rubattu et al., 2004; Komitopoulou et al., 2006; Adibhatla, Hatcher, 2008). Besides, numerous potential factors attribute to the development of ischemic stroke, such as age (Wolf et al., 1992; Brown et al., 1996), sex (Brown et al., 1996; Sacco et al., 1998), low birth weight (Barker, Lackland, 2003), race (Gorelick, 1998; Sacco et al., 1998; Wan et al., 2011).

However, the assessing the susceptibility genes remains problematic. This is in part limited by the number of approaches currently available (Hassan, Markus, 2000). And, the ischemic stroke is a polygenic disease with a number of different phenotypes which have different genetic profiles (Hassan, Markus, 2000). Therefore, it is necessary for the discovery of new susceptibility loci in different ethnic groups. The problem of ethnic background predisposition to human diseases including type 2 diabetes is widely discussed (Bai et al., 2015; Nabodita, Sher, 2016; Tiis et al., 2016).

To provide new insights in clinical diagnosis and therapy, the aim of present study was to test 15 candidate SNPs of ischemic stroke and explore genetic risks of ischemic stroke in a relatively large Mongolian ethnicity.

Methods

Ethics Statement. This study was approved by the Institutional Review Board of the Affiliated hospital of Inner Mongolia University for the Nationalities and complied with the Declaration of Helsinki. The written informed consent was obtained from each participant.

Study participants. The subjects included 343 individuals of Chinese Mongolian ethnicity, of whom 167 were patients affected with ischemic stroke at various stages of the disease and 176 were normal controls. Control individuals were not related to the ischemic stroke patients or each other. Nonstroke, healthy, non-relatives persons were selected as control group from the same region.

Peripheral blood samples from all participants were collected into heparinized tubes. Body mass index (BMI), systolic blood pressure (SBP), diastolic blood pressure (DBP), total cholesterol (TC), triglyceride (TG), high density lipoprotein cholesterol (HDL-C) and low density lipoprotein cholesterol (LDL-C) were measured at the same time for all participants (Table 1). The significance tests were performed to evaluate the case-control difference in sex, age and seven clinical parameters. Population attributable risks were evaluated by dividing the individuals into 5 groups: 30–39 years, 40–49 years, 50–59 years, 60–69 years and above 70 years old. Sex, age, BMI, high blood pressure (HBP) were selected as risk factors for ischemic stroke. The associations between the risk factors with ischemic stroke were tested using logistic regression.

A multivariate unconditional logistic regression model was employed to analyze the high risk factors of ischemic stroke.

Selection of SNPs. We selected a list of SNPs previously found to be associated with ischemic stroke based on NHGRI GWAS catalog (www.genome.gov/gwastudies, http://www.ebi.ac.uk/gwas/). Candidate SNPs were initially selected with the following considerations: (1) SNPs found to be associated with ischemic stroke in an Asian sample were given higher priority; and (2) subsequently SNPs found to be associated with multiple studies were included. We were able to genotype 15 SNPs located in or near 15 candidate genes.

DNA extraction and sequencing. DNA was extracted from all peripheral blood samples. We estimated the concentration of isolated genomic DNA using Qubit dsDNA BR Assay Kit (Invitrogen, USA), and the DNA solution was further diluted to a concentration of 10 ng/μL. We designed the targeted sequencing primers and redesigned the primer sets with dispersed or weak electrophoretic bands. To prepare the chip array, we used a multisample nanodispenser (WaferGen, USA) to disperse DNA and primers into SmartChip MyDesign Chip (WaferGen, USA). Following the polymerase chain reaction (PCR) amplification, we purified PCR products through Agencourt AMPure XP-medium beads to get mixed Illumina pair-end libraries. Insert sizes were calculated by Agilent 2100 bioanalyzer (Agilent, USA) and concentrations were estimated by Real Time PCR. Sequencing was performed on either Illumina MiSeq. All sequencing steps were in strict accordance with Illumina recommended protocols.

Data processing and statistical analysis. In brief, the final sequencing depth reached >200x, and the length of pair-end reads was 100 bp. Reads with an average base quality of ≥20 were kept for further analysis, that corresponds to common practice of sequencing data analysis (Orlov, 2014; Ignatieva et al., 2015). After filtering, we carried out assembly using Burrows-Wheeler Aligner (version 0.5.9) (Li, Durbin, 2009). We designed the targeted sequencing primers and redesigned the primer sets with Illumina MiSeq. All sequencing steps were in strict accordance with Illumina recommended protocols.

DNA extraction and sequencing. DNA was extracted from all peripheral blood samples. We estimated the concentration of isolated genomic DNA using Qubit dsDNA BR Assay Kit (Invitrogen, USA), and the DNA solution was further diluted to a concentration of 10 ng/μL. We designed the targeted sequencing primers and redesigned the primer sets with dispersed or weak electrophoretic bands. To prepare the chip array, we used a multisample nanodispenser (WaferGen, USA) to disperse DNA and primers into SmartChip MyDesign Chip (WaferGen, USA). Following the polymerase chain reaction (PCR) amplification, we purified PCR products through Agencourt AMPure XP-medium beads to get mixed Illumina pair-end libraries. Insert sizes were calculated by Agilent 2100 bioanalyzer (Agilent, USA) and concentrations were estimated by Real Time PCR. Sequencing was performed on either Illumina MiSeq. All sequencing steps were in strict accordance with Illumina recommended protocols.

DNA extraction and sequencing. DNA was extracted from all peripheral blood samples. We estimated the concentration of isolated genomic DNA using Qubit dsDNA BR Assay Kit (Invitrogen, USA), and the DNA solution was further diluted to a concentration of 10 ng/μL. We designed the targeted sequencing primers and redesigned the primer sets with dispersed or weak electrophoretic bands. To prepare the chip array, we used a multisample nanodispenser (WaferGen, USA) to disperse DNA and primers into SmartChip MyDesign Chip (WaferGen, USA). Following the polymerase chain reaction (PCR) amplification, we purified PCR products through Agencourt AMPure XP-medium beads to get mixed Illumina pair-end libraries. Insert sizes were calculated by Agilent 2100 bioanalyzer (Agilent, USA) and concentrations were estimated by Real Time PCR. Sequencing was performed on either Illumina MiSeq. All sequencing steps were in strict accordance with Illumina recommended protocols.

DNA extraction and sequencing. DNA was extracted from all peripheral blood samples. We estimated the concentration of isolated genomic DNA using Qubit dsDNA BR Assay Kit (Invitrogen, USA), and the DNA solution was further diluted to a concentration of 10 ng/μL. We designed the targeted sequencing primers and redesigned the primer sets with dispersed or weak electrophoretic bands. To prepare the chip array, we used a multisample nanodispenser (WaferGen, USA) to disperse DNA and primers into SmartChip MyDesign Chip (WaferGen, USA). Following the polymerase chain reaction (PCR) amplification, we purified PCR products through Agencourt AMPure XP-medium beads to get mixed Illumina pair-end libraries. Insert sizes were calculated by Agilent 2100 bioanalyzer (Agilent, USA) and concentrations were estimated by Real Time PCR. Sequencing was performed on either Illumina MiSeq. All sequencing steps were in strict accordance with Illumina recommended protocols.
Популяционная генетика

Table 1. Variables and assignment

<table>
<thead>
<tr>
<th>Variable</th>
<th>Assigning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, years</td>
<td>Continuous variable</td>
</tr>
<tr>
<td>BMI, kg/m²</td>
<td>Continuous variable</td>
</tr>
<tr>
<td>HBP, mmHg</td>
<td>Continuous variable</td>
</tr>
<tr>
<td>TC, mmol/L</td>
<td>Continuous variable</td>
</tr>
<tr>
<td>TG, mmol/L</td>
<td>Continuous variable</td>
</tr>
<tr>
<td>HDL-C, mmol/L</td>
<td>Continuous variable</td>
</tr>
<tr>
<td>LDL-C, mmol/L</td>
<td>Continuous variable</td>
</tr>
</tbody>
</table>

Note: Using the cutting standard of obesity in Chinese patients: BMI ≥ 28 kg/m², Male’s waist ≥ 90 cm, Female’s waist ≥ 85 cm.

Table 2. Clinical characteristics of Mongolian population

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Controls</th>
<th>Cases</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Samples, n</td>
<td>176</td>
<td>167</td>
<td></td>
</tr>
<tr>
<td>Sex (Male/Female)</td>
<td>49/127</td>
<td>96/71</td>
<td>0.000</td>
</tr>
<tr>
<td>Age, years</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BMI, kg/m²</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HBP, mmHg</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TC, mmol/L</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TG, mmol/L</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HDL-C, mmol/L</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LDL-C, mmol/L</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: Values for continuous variables denote mean ± standard error of mean (SEM).

Table 3. Multivariate logistic regression analysis

<table>
<thead>
<tr>
<th>Variable</th>
<th>B</th>
<th>SE</th>
<th>Wald</th>
<th>P-value</th>
<th>OR</th>
<th>95 % CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td>–0.98</td>
<td>0.27</td>
<td>13.51</td>
<td>2.37E–04</td>
<td>0.38</td>
<td>0.22–0.63</td>
</tr>
<tr>
<td>Age</td>
<td>0.05</td>
<td>0.01</td>
<td>11.88</td>
<td>5.67E–04</td>
<td>1.05</td>
<td>1.02–1.07</td>
</tr>
<tr>
<td>BMI</td>
<td>0.90</td>
<td>0.28</td>
<td>10.18</td>
<td>1.42E–03</td>
<td>2.45</td>
<td>1.41–4.25</td>
</tr>
<tr>
<td>HBP</td>
<td>1.95</td>
<td>0.3</td>
<td>43.43</td>
<td>4.39E–11</td>
<td>7.05</td>
<td>3.95–12.61</td>
</tr>
</tbody>
</table>

SBP and DBP. We tested association with diabetes related quantitative traits (TC, HDL-C, LDL-C, and TG) across both ischemic stroke cases and controls using linear regression with the age, sex, BMI, and ischemic stroke status as covariates. All quantitative trait measures were normalized by quantile normalization and the normalized values were used in the analyses. Formal statistical tests, including 95 % confidence intervals (CI), were performed using SPSS 13.0.

Differences in population structure between the healthy and control sample were estimated by comparing risk allele frequency and the Wright’s fixation index (FST) using plink.

Results

After rigorous sample and marker level quality control filtering, genotypes of 15 SNPs on 343 individuals (including with 167 ischemic stroke cases and 176 stroke ethnically matched controls) were kept for subsequent analyses. Clinical characteristics of the sample are summarized in Table 2. Sex, age, BMI, SBP and DBP were significant difference between patients and normal (Table 2, Figure). Results of this study suggested that age, high blood pressure (HBP) are major risk factors for ischemic stroke in Chinese Mongolian population (Table 3). Table 4 presents the association results.
between the 15 SNPs and ischemic stroke status. Of the 15 SNPs tested, SNPs located within \textit{NINJ2} (rs12425791, 95 % CI = 1.08–2.15) were significantly associated with ischemic stroke. As a major risk factor of cerebral infarction and intracerebral hemorrhage (Izzo et al., 2003; Fields et al., 2004), hypertension contributes to the attack of stroke (Lewington et al., 2002). Along with the increasing age, blood pressure increase also (Burt et al., 1995). The control of blood pressure is an effective method for prevention of outbreak, development and cure of ischemic stroke (Black et al., 2001; Whelton et al., 2002; Ong et al., 2007). Based on our analysis, the typical clinical features of patients with ischemic stroke have obvious high blood pressure in Mongolian population, further prompt the Chinese Mongolian stroke may be caused by accelerated atherosclerosis.

Several studies have evaluated the association between rs12425791 and risk of ischemic stroke, but the results remain controversial. SNPs rs12425791 locates on chromosome 12p13, close to \textit{NINJ2} gene which encodes ninjurin-2 protein. Ninjurin-2 protein, also named nerve injury-induced protein, is induced expression by nerve injury in Schwann protein. Ninjurin-2 protein. Ninjurin-2 protein, also named nerve injury-induced protein, is induced expression by nerve injury in Schwann cells (Araki, Milbrandt, 2000). The GWAS analyses showed that rs12425791 is associated with an increased risk of stroke for black and Dutch population (Ikram et al., 2009) and Asian population (Li et al., 2012; Lian et al., 2012). However, the contradictory studies about association between rs12425791 or \textit{NINJ2} gene and ischemic stroke also emerge in Chinese Han population (Chen et al., 2010; Ding et al., 2011; Tong et al., 2011; Wan et al., 2014). These studies suggest that the association between \textit{NINJ2} gene and ischemic stroke may be caused by other underlying factors.
et al., 2011; Gu et al., 2013; Xie et al., 2013). In our study, we found variants rs12425791 of the NINJ2 gene to be strongest associated with ischemic stroke in Chinese Mongolian ethnicity, and A alleles increase the risk of ischemic stroke, indicating the rs12425791 of the NINJ2 gene might play function in the insult of ischemic stroke. 

ALDH2 gene locates in chromosome 12q24.2 and encodes mitochondrial aldehyde dehydrogenase (ALDH-E2). ALDH-E2 involves in the oxidative pathway of alcohol metabolism (Smith, 1986; Hempel et al., 1987). Substantial evidence indicated that alcohol abuse is sever risk factor for stroke (Gill et al., 1986, 1991; Hillbom et al., 1999; Klatsky et al., 2001; Mazzaglia et al., 2001). ALDH2 polymorphisms play a pivotal role on hypertension (Hasi et al., 2011; Wang et al., 2013; Yokoyama et al., 2013; Hu et al., 2014), heart disease (Gu, Li, 2014; Mizuno et al., 2015) and stroke (Yao et al., 2011; Lai et al., 2012). The single base mutation (ALDH2*2) of ALDH2, the predominant allele in East Asian populations (around 70 %) (Chen, Yin, 2008), is responsible for acute alcohol-flushing reaction in Asians (Crabb, 1990). Indeed, Mongolian drinks more than other ethnic groups in China (Cochrane et al., 2003). In the present study, SNP rs2238151 of ALDH2 gene was associated with ischemic stroke in Chinese Mongolian population, indicating rs2238151 take action in ischemic stroke by alcohol metabolism.

In conclusion, we identified SNP rs2238151 of ALDH2 gene and rs12425791 of NINJ2 gene associated with ischemic stroke risk in the Chinese Mongolian population. Our discovery also demonstrated genetic heterogeneity exists between Chinese and other populations. Meanwhile, this study has limitations, we had limited power to detect associations with small effect sizes and associations with rare variants.

The studies of genes associated with the human deceases such as predisposition to alcohol have been started at the Institute of Cytology and Genetics SB RAS earlier on animal models (Morozova, Popova, 2010). Such problems of computer modeling in animal genetics were discussed at BGRS-SB-2016 conference (Orlov et al., 2016) that continue series of collaborative works between the authors from China and Russia (Bai et al., 2015).

Acknowledgments
We thank our study individuals for their generous participation in this study. Authors are grateful to L.P. Osipova and V.N. Babenko for science discussion. This study was supported by the National Natural Science Foundation of China (81060098 and 31402106), NSFC-RFBR Cooperation and Exchange Projects (8151130050 and RFBR), Scientific Research Project of Inner Mongolia University for Nationalities (NMD1321) and Science and Technology Program Project of Inner Mongolia Autonomous Region (Y Y150002). Work by YLO, AOB and LET was supported by ICG SB RAS budget project 0324-2016-0003 and RFBR project 15-54-53091.

References


Genetic polymorphisms and related risk factors of ischemic stroke in a Mongolian population in China
Q. Wu, H. Wu, Yu.L. Orlov … L.E. Tabikhanova, M. Chen, H. Bai

Lindsberg P.J., Grau A. Inflammation and infections as risk factors for ischemic stroke. Stroke. 2003;34(10):2518-2523. DOI 10.1161/01.STR.000008915.51603.CC.
Генетические факторы риска ишемического инсульта в популяции монголов Китая

Ч. У, Х. У, Ю.Л. Орлов, Л.Э. Табиханова, М. Чен, Х. Бай

2017
21 • 5

Популяционная генетика
Вавиловский журнал генетики и селекции • 2017 • 21 • 5


