Original Research

HDAC9 gene is associated with stroke risk in a Chinese population

Yan Han[1,*], Wenzhu Sun[2,3,*], Li Wang[4,5,*], Sha Tao[4,5], Lu Tian[4,5], Yong Hao[1], Wei Zhang[6,7], Shuai Wu[1], Shixu Li[1], Huihui Lv[1], S Lilly Zheng[4,5], Jielin Sun[4,5] and Jianfeng Xu[2,3,4,5]

[1]Department of Neurology, Changhai Hospital, Second Military Medical University, Shanghai 200433, China; [2]State Key Laboratory of Genetic Engineering, School of Life Sciences, Fudan University, Shanghai 200433, China; [3]Center for Genetic Epidemiology, School of Life Sciences, Fudan University, Shanghai 200433, China; [4]Center for Cancer Genomics, Wake Forest University School of Medicine, Winston-Salem, NC 27157, USA; [5]Center for Genomics and Personalized Medicine Research, Wake Forest University School of Medicine, Winston-Salem, NC 27157, USA; [6]Department of Neurology, Chinese People's Liberation Army General Hospital, Beijing 100853, China; [7]Department of Neurology, Chinese People's Liberation Army 307 Hospital, Beijing 100071, China

*The first three authors contributed equally to this work.

Corresponding author: Jianfeng Xu. Email: jxu@wakehealth.edu

Abstract

A recent Genome-Wide Association study (GWAS) identified rs11984041 on *HDAC9* gene to be significantly associated with stroke in a Caucasian population. However, whether *HDAC9* gene also plays a role in other ethnicities is unknown. The current study was conducted to investigate the association of single nucleotide polymorphisms (SNPs) on *HDAC9* gene and stroke risk in a Chinese population. Sixteen tagging SNPs for *HDAC9* gene were genotyped in a Chinese population of 279 stroke cases and 984 controls from Changhai Hospital in Shanghai, China. The candidate region of *HDAC9* investigated was on a haplotype block located between two recombination hotspots in the tail region of *HDAC9*, the only region harbouring significantly associated SNPs based on the previous GWAS. rs11984041 was not polymorphic in Chinese population. Two other SNPs, rs2389995 and rs2240419 on *HDAC9* of chromosome 7q21.1, were significantly associated with large-vessel stroke risk, with *P* values of 0.035 and 0.042, respectively (Table 3). Individuals with risk allele (A) for rs2389995 and (T) for rs2240419 had increased risk of stroke (odds ratio [OR] = 1.33, 95% confidence interval [CI]: 1.01–1.75; and OR = 1.29, 95% CI: 1.02–1.63), respectively. Multivariate analysis including both SNPs in the logistic regression model revealed independent association effects of each SNP, with a $P = 0.013$ and 0.010, respectively. The results from our studies indicate *HDAC9* gene is significantly associated with large-vessel stroke risk in Chinese population. However, SNPs on *HDAC9* gene display heterogeneity effects across different ethnic populations. Additional studies are needed to further evaluate our results.

Keywords: stroke, *HDAC9*, large vessel, SNP, Chinese, ischaemic

***Experimental Biology and Medicine* 2013; 238: 842–847. DOI: 10.1177/1535370213494650**

Introduction

Stroke (cerebrovascular disease) is one of the leading causes of death and adult chronic disability internationally. The prevalence of stroke increases as the proportion of older populations internationally increases. Conventional risk factors such as hypertension explain a substantial portion of stroke risk, however, much remains unknown.[1] Twin and familial studies have shown genetic components to contribute a substantial susceptibility to stroke.[2,3] Eighty percentage of all stroke cases are ischaemic, the three most common types being large-vessel, cardioembolic and small vessel (lacunar) stroke.[4] Genetic components have been shown to influence large-vessel atherosclerosis and small vessel disease (SVD) more than cardioembolic atherosclerosis.[5] However, differences in genetic susceptibility between ethnicities have yet to be fully elaborated.

A recent Genome-Wide Association study (GWAS) study reported a variant of *HDAC9* (rs11984041) associated with large-vessel ischaemic stroke in populations with European ancestry ($P < 1 \times 10^{-7}$) with an estimated effect size (odds ratio [OR] = 1.38, 95% confidence interval [CI] = 1.22–1.57).[4] However to date, no comprehensive studies have been conducted to evaluate the effect of this single nucleotide polymorphism (SNP) in other ethnicities including the Chinese population, for which stroke is also the second leading cause of death, close behind malignant tumours with a high incidence rate.[6,7] Interestingly, rs11984041 is not polymorphic in the Chinese population indicating that potential genetic heterogeneity may exist across ethnicities. Therefore, fine-mapping studies are needed to evaluate

the role of the *HDAC9* gene and stroke risk in the Chinese population.

In this study, we aimed to determine the association of 16 tagging SNPs (tSNPs) selected from the *HDAC9* gene with stroke risk in a Chinese population. Our study represents one of the first efforts to elucidate the effect of *HDAC9* with stroke in the Chinese population.

Materials and methods

Study subjects

All subjects were of Chinese Han ancestry. A total of 279 stroke cases and 984 controls were included in the study. Cases were recruited at Changhai Hospital of Second Military Medical University in Shanghai, China. Inclusion criteria included cases diagnosed with stroke between July 2011 and June 2012. Community was used as controls in the current study. All the controls were collected from April 2010 to November 2010 in Shanghai. Four Shanghai districts and counties were selected via random cluster sampling methods. For each selected district or county, one or two communities were selected. Inclusion criteria for controls were as follows: (i) age of 40–79 years old, with full self-care skill and has been a resident in Shanghai for more than 1 year and (ii) clear consciousness, able to provide a blood sample and complete a medical examination. A total of 1587 adult males participated in the control sample recruitment and 1539 (96.98) completed all the questionnaires and examinations needed. Due to funding limitations, 1008 subjects were randomly selected for genotyping and 984 were successfully genotyped. This study was approved by the Institutional Ethics Review Board at Changhai Hospital and written informed consent was completed with each individual. Population characteristics are summarized in Table 1.

SNP selection

In the previous GWAS study, *HDAC9* variants with association signals ($P < 0.05$) were located between two recombination hotspots that encompass the tail end.[4] Therefore, the last haplotype block located between two recombination hotspots for the CHB (Chinese Beijing) population was chosen for our study. Based on HapMap phase 3 data on the CHB population (release #27, NCBI build 36), this includes a region from 18,935 to 19,004 kb on chromosome 7 (Figure 1). Aggressive-tagging methods (two- and three-marker) implemented in Haploview software[8] were used to select tSNPs. A total of 16 tSNPs were selected based on Hapmap phase 3 (release 27 in the CHB + JPT population).

Genotyping

The 16 SNPs were genotyped for all study subjects using the MassARRAY iPLEX system (Sequenom, Inc., San Diego, CA, USA) at Fudan University in Shanghai, China. Two duplicates and two water samples were included in each 96-well plate as polymerase chain reaction-negative controls. All assays were performed in a blinded fashion. Genotyping missing rates were 2.12% for all samples.

Table 1 Clinical and demographic characteristics of all subjects

Characteristics	Stroke cases $N = 279$[a]	Controls With genotypes ($N = 984$)
Age (mean (SD))	67 (11.65)	61 (9.00)
Male (%)	183 (66.60)	984 (100)
Large vessel (%)	237 (84.95)	NA
Cardioembolic stroke (%)	35 (12.54)	NA
Lacunar stroke (%)	7 (2.51)	NA
Medical history		
Hypertension (%)	183 (62.28)	NA
Diabetes (%)	199 (74.25)	NA
Hyperlipidaemia (%)	143 (53.36)	NA

[a]No subtype information was available for five case subjects.

Statistical analysis

The genotype distributions for all tSNPs were tested for Hardy-Weinberg equilibrium (HWE). The main effects of the tSNPs were estimated using a logistic regression model, assuming an additive model of inheritance, adjusting for age. SNP–SNP interaction was evaluated by including both SNPs and an multiplicative interaction term in the logistic model, adjusting for age. All analyses were conducted using PLINK software.[9] *P* values were two-tailed. An alpha of 0.05 was used to claim nominal statistical significance.

Results

Demographic and clinical information for 279 cases and 984 controls with genotypes are presented in Table 1. A total of 237 large-vessel stroke cases comprised the majority of the total stroke cases (84.95%); small-vessel stroke comprising 12.54% and cardioembolic 2.51% of the total cases. Mean age was higher for cases compared with controls and was adjusted in the association test.

Genotype distributions for all SNPs were in HWE in both control and case groups ($P > 0.05$, data not shown). All SNPs had missing rates smaller than 5% (data not shown). The association results for the 16 tSNPS with all stroke cases are presented in Table 2. None of the 16 SNPs was significantly associated with overall stroke risk ($P > 0.05$).

Further analysis, limited to large-vessel stroke cases only, revealed two SNPs, rs2389995 and rs2240419 to be statistically associated with large-vessel stroke risk, with *P* values of 0.035 and 0.042, respectively. Risk allele 'A' of rs2389995 was associated with a 1.33-fold increased risk with stroke (95% CI: 1.01–1.75) while risk allele 'T' of rs2240419 was associated with a 1.29-fold of increased risk of stroke (95% CI: 1.02–1.63). Low linkage disequilibrium (LD) was suggested between the two SNPs, as indicated by an R^2 value of 0.006, and a D' value of 0.07 (Figure 1). Multivariate analysis with both SNPs in the logistic model, with adjusting for age, also indicated the independent association effects ($P = 0.013$ and 0.010, respectively). SNP–SNP interaction was tested but no significant gene–gene interaction effect was found

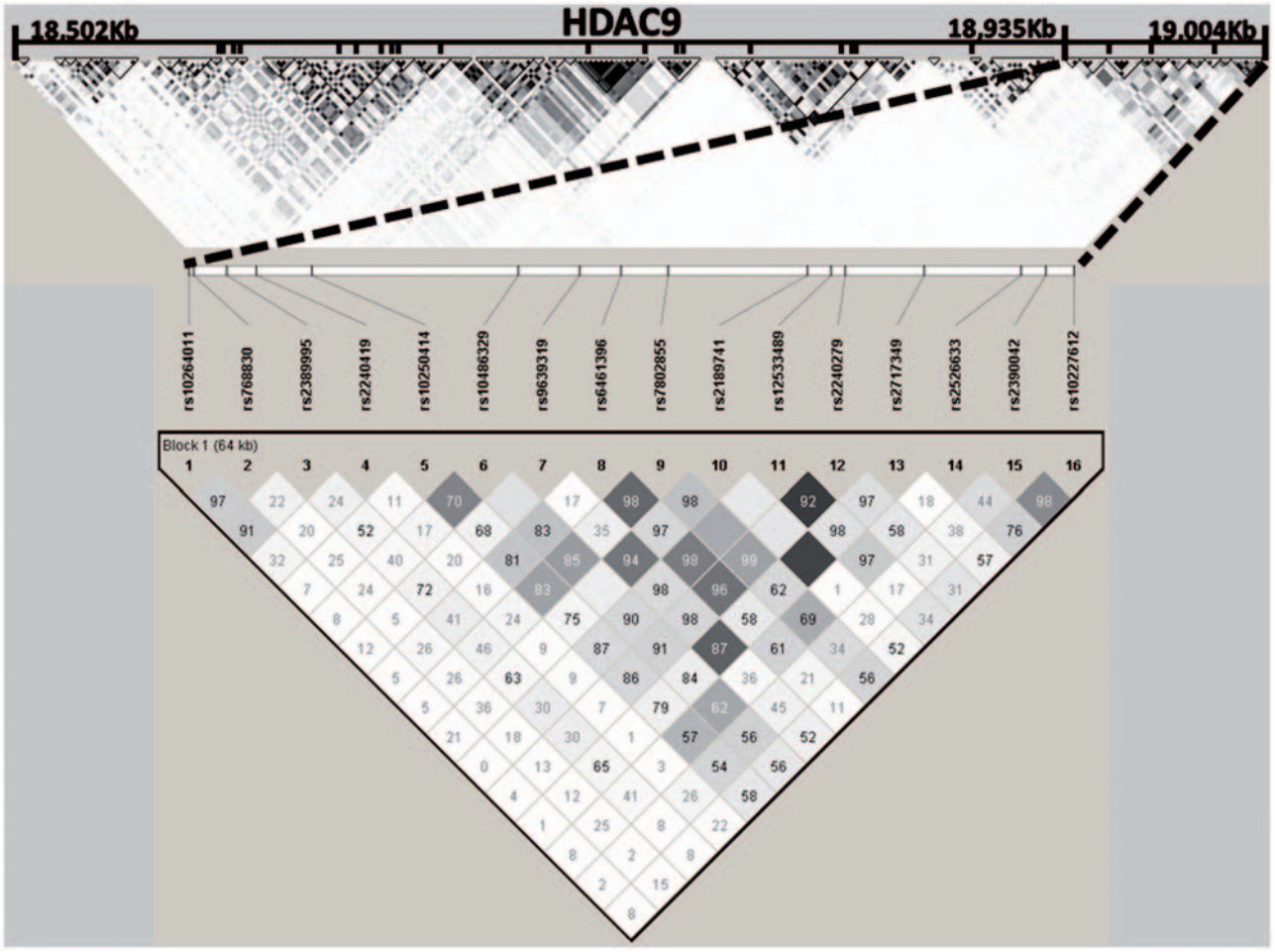

Figure 1 Linkage disequilibrium (LD) structure for *HDAC9* gene. The top panel indicates the LD map based on Hapmap phase 3 data on the CHB population (Release #27, NCBI build 36) using Haploview software. The bottom panel indicates the LD structure based on our own study population including 279 stroke cases and 984 controls. The numbers showed in the plot were the values for pairwise D′

($P > 0.05$ for the multiplicative interaction term, data not shown). Additional analysis, limited to large-vessel stroke for men only showed similar results (Supplementary Table 1).

Discussion

In this study, we systematically investigated the association of 16 tSNPs selected from the last block encompassing the tail end of *HDAC9* with stroke risk in a Chinese population of 279 cases of stroke and 984 controls. Two SNPs, rs2389995 and rs2240419 were significantly associated with large-vessel ischaemic stroke risk in the Chinese population ($P = 0.035$ and 0.042, respectively).

In a previous European (CEU) GWAS study, *HDAC9* was reported as significantly associated with large-vessel ischaemic stroke (OR = 1.42[1.28–1.57]) from a meta-analysis study ($P = 1.87 \times 10^{-11}$). However, this SNP was found to be nonpolymorphic in the Chinese and Japanese (CHB) population based on HapMap data. All variants with an association signal in the region of HDAC9 were located in the tail region of HDAC9 between two recombination hotspots. For this reason, tSNPs encompassing the tail region of HDAC9 between two recombination hotspots were chosen to determine possible association with large-vessel ischaemic stroke in the Chinese population.

Two SNPs were significantly associated with large-vessel stroke. The two significant SNPs, rs2389995 and rs2240419, are located 25,442 and 23,271 bp upstream of the previously reported SNP of rs11984041, respectively. Both of the two SNPs are located on the intron 21, while rs11984041 was located in intron 24. Interestingly, previously reported GWAS SNP rs11984041, which was significantly associated with large-vessel stroke in populations of European ancestry, was found to be nonpolymorphic in the CHB + JPT population while rs2389995 and rs2240419 were found to be nonpolymorphic in the CEU population supporting the concept that SNPs confer heterogeneity effects among different ethnic populations for stroke risk. In addition, although *HDAC9* plays a role in stroke risk in Caucasian and Chinese populations, the effect sizes, 1.42, 1.33, and 1.29 for rs11984041, rs2389995 and rs2240419, respectively, are

Table 2 Fine-mapping results for SNPs located in *HDAC9* In a Chinese stroke population

SNP	BP	Alleles[a]	Risk allele	RAF		OR (95% CI)	*P* value
				Case	Control		
rs10264011	18970232	A > T	T	0.443	0.439	1.01 (0.82–1.24)	0.922
rs768830	18970607	A > G	A	0.822	0.819	1.03 (0.79–1.34)	0.833
rs2389995	18973018	A > G	A	0.819	0.784	1.25 (0.97–1.62)	0.083
rs2240419	18975189	C > T	T	0.278	0.240	1.22 (0.98–1.53)	0.080
rs10250414	18979194	G > A	G	0.759	0.751	1.10 (0.86–1.41)	0.457
rs10486329	18994192	T > A	A	0.259	0.254	1.04 (0.83–1.30)	0.733
rs9639319	18998650	G > T	G	0.797	0.787	1.08 (0.84–1.38)	0.538
rs6461396	19001691	T > C	C	0.477	0.461	1.10 (0.90–1.35)	0.338
rs7802855	19005077	C > T	C	0.612	0.585	1.15 (0.94–1.41)	0.183
rs2189741	19015263	G > T	G	0.882	0.861	1.19 (0.88–1.61)	0.266
rs12533489	19016933	G > C	C	0.312	0.309	1.05 (0.85–1.30)	0.657
rs2240279	19018009	T > C	C	0.354	0.332	1.16 (0.94–1.43)	0.164
rs2717349	19023736	C > G	C	0.847	0.818	1.22 (0.93–1.59)	0.153
rs2526633	19030751	G > C	C	0.310	0.297	1.02 (0.83–1.27)	0.840
rs2390042	19032534	C > T	C	0.592	0.566	1.13 (0.92–1.39)	0.229
rs10227612	19034579	T > G	G	0.408	0.374	1.17 (0.95–1.43)	0.140

RAF: risk allele frequency; OR: odds ratio.

[a]Major allele > minor allele.

P value and 95% CI calculated based on logistic regression and analysis with adjustment of age.

Table 3 Fine-mapping results for SNPs located in *HDAC9* in a Chinese large-vessel stroke population.

SNP	BP	Alleles	Risk allele	RAF		OR (95% CI)[a]	*P* value[a]
				Case	Control		
rs10264011	18970232	A > T	T	0.451	0.439	1.05 (0.85–1.31)	0.644
rs768830	18970607	A > G	A	0.819	0.819	1.02 (0.77–1.34)	0.906
rs2389995	18973018	A > G	A	0.827	0.784	1.33 (1.01–1.75)	**0.042**
rs2240419	18975189	C > T	T	0.288	0.240	1.29 (1.02–1.63)	**0.036**
rs10250414	18979194	G > A	G	0.756	0.751	1.09 (0.84–1.41)	0.522
rs10486329	18994192	T > A	A	0.270	0.254	1.01 (0.80–1.28)	0.910
rs9639319	18998650	G > T	G	0.797	0.787	1.08 (0.83–1.39)	0.594
rs6461396	19001691	T > C	C	0.464	0.461	1.06 (0.86–1.31)	0.593
rs7802855	19005077	C > T	C	0.600	0.585	1.10 (0.88–1.35)	0.415
rs2189741	19015263	G > T	G	0.878	0.861	1.15 (0.83–1.56)	0.401
rs12533489	19016933	G > C	G	0.696	0.691	1.01 (0.81–1.27)	0.914
rs2240279	19018009	T > C	C	0.346	0.332	1.12 (0.90–1.40)	0.303
rs2717349	19023736	C > G	C	0.846	0.818	1.22 (0.92–1.61)	0.178
rs2526633	19030751	G > C	C	0.326	0.297	1.10 (0.88–1.37)	0.416
rs2390042	19032534	C > T	C	0.587	0.566	1.11 (0.89–1.37)	0.349
rs10227612	19034579	T > G	G	0.404	0.374	1.14 (0.92–1.41)	0.240

[a]*P* value and 95% CI were calculated based on logistic regression and analysis with adjustment of age.

different, further indicating genetic heterogeneity effect for SNPs on *HDAC9* gene with stroke risk.

Evidence has shown that the previous reported SNP, rs11984041, did not have the same effect in each stroke subtype, indicating a genetic heterogeneity across stroke subtypes at this SNP. Particularly, rs11984041 showed greatest effect for large-vessel disease while contributing much smaller effect sizes for SVD and cardiovascular embolism (CE).[4] This effect is also suggested in our study, as the analysis for overall stroke cases versus controls did not show statistically significant association for rs2389995 and rs2240419, while the analysis limiting to large-vessel strokes revealed significant association. However, since the majority of our stroke samples are composed of large-vessel

subtype (87%), we do not have sufficient power to estimate a precise magnitude of association in the other two subtypes (SVD and CE). Therefore, further studies with more samples on SVD and CE subtypes are needed to evaluate the association of SNPs on *HDAC9* gene.

HDAC9 is a part of the *HDAC* gene family which regulates gene expression patterns by encoding proteins that deacetylate histones, resulting in chromatin condensation and transcriptional repression. Its dynamic effects are governed and reversible by histone acetyltransferases. Vertebrate *HDACs* are divided into two categories, class I, II and 3, based on their homology to yeast genes *Rpd3*, *HdaI*, *Sir2*, respectively.[10,11] HDAC9, a class II HDAC, is a signal-responsive transcriptional repressor which is down-regulated upon denervation with upregulation of chromatin acetylation and has been linked to suppression of cardiac hypertrophy.[12] Lack of HDAC9 in *in vivo* models has resulted in an increased sensitivity to cardiac stress signals.[13] HDAC9, having two isoforms designated *HDAC9* and *HDAC9*a, has been shown to suppress myogenesis through myocyte enhancing transcriptional factor 2 (MEF2)[14] and heart development.[13] However, *HDAC9* involvement with systemic arteries has yet to be reported. The isoform *HDAC9* has been shown to have more localization in the brain, heart and pancreas[14] and has been the target of our association study attempt. Due to its localization in the brain, an alternative method of increased stroke risk would be that HDACs play the dual role of activators and repressors of cardiac hypertrophy.[11] HDACs have been linked to the suppression of the various genes relevant to the protection against ischaemic insult and neuronal survival. Forced *HDAC9* expression has been shown to prevent upregulation of activity-dependent genes and chromatin acetylation through MEF2 and class I HDACs.[15] HDAC inhibitors have been asserted as possible treatments for stroke.[11] Despite the above potential supporting evidence for HDAC9's role in stroke development, further studies must be conducted in order to determine the biological mechanism of *HDAC9* for stroke risk.

There are some limitations to our study that one should be made aware of. First, due to our relatively small sample size, we only have sufficient power to detect SNPs that confer relatively modest to large effect sizes. For example, our samples allow us to have an 80% power to detect association for an SNP with an OR of 1.50 (1.36) at an alpha level of 0.05, assuming an MAF of 0.1 (0.2). Another limitation is that none of the identified SNPs reached Bonferroni correction significance. However, Bonfferonni correction may represent an over stringent criteria since not all of the SNPs tested are independent due to the LD structure. Besides, we observed more than the expected number of SNPs at a nominal P value of 0.05. Under the null hypothesis, we expect to observe one SNP to be significant at the P value of 0.05 (16×0.05). However, we observed two independent significant SNPs at this P value level. The excess of observed associated SNPs may indicate a true association signal between *HDAC9* gene and large-vessel stroke.

In conclusion, through a fine mapping study, we identified two SNPs on *HDAC9* associated with large-vessel ischaemic stroke risk in the Chinese population. Our findings also demonstrate genetic heterogeneity exists between Chinese versus European American populations. Future studies with large sample sizes need to be conducted in the Chinese population in order to further evaluate our findings.

Author contributions: JS and JX were involved in conception and design; YH, WS and LW conducted analysis and interpretation; WS, YH, WZ, SW, HL and SZ were involved in data collection; ST and LT performed statistical analysis; YH, LW, WS, JS, and JX wrote the article; JX obtained funding and took overall responsibility.

ACKNOWLEDGEMENTS

The authors thank all the study subjects who participated in this study. The study is supported by an intramural grant by Fudan University to J Xu.

REFERENCES

1. Singh RB, Suh I, Singh VP, Chaithiraphan S, Laothavorn P, Sy RG, Babilonia NA, Rahman ARA, Sheikh S, Tomlinson B, Sarraf-Zadigan N. Hypertension and stroke in Asia: prevalence, control and strategies in developing countries for prevention. *J Hum Hypertens* 2000;**14**:749–63
2. Liao DP, Myers R, Hunt S, Shahar E, Paton C, Burke G, Province M, Heiss G. Familial history of stroke and stroke risk - The Family Heart Study. *Stroke* 1997;**28**:1908–12
3. Bejot Y, Benatru I, Rouaud O, Fromont A, Besancenot JP, Moreau T, Giroud M. Epidemiology of stroke in Europe: geographic and environmental differences. *J Neurol Sci* 2007;**262**:85–8
4. Bellenguez C, Bevan S, Gschwendtner A, Spencer CCA, 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 Y-C, Woo D, Nalls MA, Mueller-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 M-G, Goel A, Helgadottir A, Blackwell JM, Bramon E, Brown MA, Casas JP, Corvin A, Duncanson A, Jankowski J, Mathew CG, Palmer CNA, 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 CLM, Rothwell PM, Dichgans M, Donnelly P, Markus HS, Isgc, Wtccc. Genome-wide association study identifies a variant in HDAC9 associated with large vessel ischemic stroke. *Nat Genet* 2012;**44**:328–33
5. Jerrard-Dunne P, Cloud G, Hassan A, Markus HS. Evaluating the genetic component of ischemic stroke subtypes - A family history study. *Stroke* 2003;**34**:1364–9
6. Liu J, Cheng J, Peng J, Han S, Yu L, Nie S. Effects of polymorphisms of heat shock protein 70 gene on ischemic stroke, and interaction with smoking in China. *Clin Chim Acta* 2007;**384**:64–8
7. Jia Q, Liu L, Wang Y. Risk factors and prevention of stroke in the Chinese Population. *J Stroke Cerebrovasc Dis* 2011;**20**:395–400
8. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005;**21**:263–5
9. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, de Bakker PIW, Daly MJ, Sham PC. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007;**81**:559–75
10. Bertos NR, Wang AH, Yang XJ. Class II histone deacetylases: structure, function, and regulation. *Biochem Cell Biol-Biochim Biol Cell* 2001;**79**:243–52

11. Antos CL, McKinsey TA, Dreitz M, Hollingsworth LM, Zhang CL, Schreiber K, Rindt H, Gorczynski RJ, Olson EN. Dose-dependent blockade to cardiomyocyte hypertrophy by histone deacetylase inhibitors. *J Biol Chem* 2003;**278**:28930–7
12. Zhang CL, McKinsey TA, Chang SR, Antos CL, Hill JA, Olson EN. Class II histone deacetylases act as signal-responsive repressors of cardiac hypertrophy. *Cell* 2002;**110**:479–88
13. Chang SR, McKinsey TA, Zhang CL, Richardson JA, Hill JA, Olson EN. Histone deacetylases 5 and 9 govern responsiveness of the heart to a subset of stress signals and play redundant roles in heart development. *Mol Cell Biol* 2004;**24**:8467–76
14. Zhou XB, Marks PA, Rifkind RA, Richon VM. Cloning and characterization of a histone deacetylase, HDAC9. *Proc Natl Acad Sci U S A* 2001;**98**:10572–7
15. Mejat A, Ramond F, Bassel-Duby R, Khochbin S, Olson EN, Schaeffer L. Histone deacetylase 9 couples neuronal activity to muscle chromatin acetylation and gene expression. *Nat Neurosci* 2005;**8**:313–21

(Received January 28, 2013, Accepted April 8, 2013)