

RESEARCH

Open Access



# *HTRA1* methylation in peripheral blood as a potential marker for the preclinical detection of stroke: a case–control study and a prospective nested case–control study

Chunlan Liu<sup>1†</sup>, Mengxia Li<sup>1†</sup>, Qiming Yin<sup>1</sup>, Yao Fan<sup>2</sup>, Chong Shen<sup>1\*</sup> and Rongxi Yang<sup>1\*</sup>

## Abstract

**Background:** Stroke is the leading cause of mortality in China. DNA methylation has essential roles in multiple diseases, but its association with stroke was barely studied. We hereby explored the association between blood-based HTRA serine protease 1 (*HTRA1*) methylation and the risk of stroke.

**Results:** The association was discovered in a hospital-based case–control study (cases/controls = 190:190) and further validated in a prospective nested case–control study including 139 cases who developed stroke within 2 years after recruitment and 144 matched stroke-free controls. We observed stroke-related altered *HTRA1* methylation and expression in both case–control study and prospective study. This blood-based *HTRA1* methylation was associated with stroke independently from the known risk factors and mostly affected the older population. The prospective results further showed that the altered *HTRA1* methylation was detectable 2 years before the clinical determination of stroke and became more robust with increased discriminatory power for stroke along with time when combined with other known stroke-related variables [onset time  $\leq 1$  year: area under the curve (AUC) = 0.76].

**Conclusions:** In our study, altered *HTRA1* methylation was associated with stroke at clinical and preclinical stages and thus may provide a potential biomarker in the blood for the risk evaluation and preclinical detection of stroke.

**Keywords:** *HTRA1*, Methylation, Stroke, Preclinical detection, Marker

## Background

Stroke is a leading cause of mortality and a major contributor to neurological disability in low-income and middle-income countries [1, 2]. In 2019, stroke was responsible for 6.55 million deaths and 143.23 million disability-adjusted life years (DALYs) worldwide [2]. In China, there were

3.94 million new stroke cases and 2.19 million deaths with the highest DALYs of 45.9 million lost in 2019 [3].

Stroke is a complex disease attributed to genetic and environmental factors [4]. According to the genome-wide association studies, multiple genetic factors have only explained 5–10% of the risk for stroke [5, 6]. In recent years, attention has been paid to epigenetic factors and their role in the pathophysiological process of stroke [7]. DNA methylation is a crucial epigenetic modification involved in diabetes mellitus [8], atherosclerosis [9] and cardiovascular diseases [10]. DNA methylation can influence the structure of the genome and regulate the process of transcription, which mediates alterations in the expression of specific genes [11, 12]. Candidate approach

<sup>†</sup>Chunlan Liu and Mengxia Li contributed equally to this work

\*Correspondence: sc100@126.com; rongxiyang@njmu.edu.cn

<sup>1</sup> Department of Epidemiology and Biostatistics, School of Public Health, Nanjing Medical University, 101 Longmian Avenue, Jiangning, Nanjing 211166, China

Full list of author information is available at the end of the article



studies have identified the association between stroke and altered blood-based DNA methylation in *LINE-1*, *ABCB1* and *CBS* genes, but mainly in case–control studies with small sample sizes [13–15]. Our previous study has also revealed that aberrant methylation of *ACTB* in peripheral blood was associated with the risk of stroke [16].

The HTRA serine protease 1 (encoded by *HTRA1* gene) is one of the human serine proteases that participate in diverse physiological processes including the maintenance of mitochondrial homeostasis, cell signaling and apoptosis [17]. Several studies have shown that heterozygous *HTRA1* mutations are closely related to cerebral small vessel disease (SVD), small-vessel ischemic stroke and lacunar stroke [18–23]. Our previous study also showed that *HTRA1* variants might confer the genetic susceptibility to stroke [24]. However, hereditary stroke attributable to *HTRA1* gene constitutes only a small proportion of all strokes. So far, no data are available about the association between blood-based *HTRA1* methylation and stroke, especially in the prospective studies.

In the present study, the association between peripheral blood-based *HTRA1* methylation and the risk of stroke in the Chinese population was evaluated in an independent case–control study consisting of 190 stroke patients and 190 age- and sex-matched controls and was

further verified in a prospective nested case–control study including 139 cases who developed stroke within 2 years after recruitment and 144 matched stroke-free controls from a cohort with a total of 11,151 subjects.

## Results

### Demographic characteristics of the subjects

The demographic and clinical characteristics of subjects in the case–control study and prospective nested case–control study are presented in Table 1. In the case–control study, the stroke cases had a higher proportion of drinking, higher prevalence of hypertension and diabetes, and higher levels of diastolic blood pressure (DBP), but lower levels of high-density lipoprotein cholesterol (HDL-C) than controls (all  $P < 0.05$ ). No significant difference existed in age, sex, smoking status, systolic blood pressure (SBP), total cholesterol (TC), triglycerides (TG), low-density lipoprotein cholesterol (LDL-C), and fasting plasma glucose between the two groups.

In the prospective nested case–control study, there were no significant differences in age, sex, SBP, DBP, smoking status, drinking status, history of hypertension, diabetes, levels of TC, TG, HDL-C, LDL-C and fasting plasma glucose between stroke cases and controls. The median time between blood draw and initial diagnosis of stroke in cases was 1.32 years.

**Table 1** Demographic and clinical characteristics of subjects

Characteristics	Group	Case–control study				Prospective nested case–control study			
		Controls (n = 190)	Stroke cases (n = 190)	$t/\chi^2$	P value	Controls (n = 144)	Stroke cases (n = 139)	$t/\chi^2$	P value
Age (mean $\pm$ SD, years)		65.89 $\pm$ 10.12	64.99 $\pm$ 9.85	0.884	0.377	67.77 $\pm$ 9.11	67.64 $\pm$ 9.51	0.117	0.907
Sex	Male	83 (43.7%)	79 (41.6%)	0.172	0.678	83 (57.6%)	81 (58.3%)	0.012	0.914
	Female	107 (56.3%)	111 (58.4%)			61 (42.4%)	58 (41.7%)		
SBP (mean $\pm$ SD, mmHg)		148.65 $\pm$ 19.76	152.44 $\pm$ 22.27	1.748	0.081	147.74 $\pm$ 20.19	147.73 $\pm$ 22.28	0.005	0.996
DBP (mean $\pm$ SD, mmHg)		80.46 $\pm$ 11.75	85.62 $\pm$ 14.42	3.814	1.60E–04	81.50 $\pm$ 9.75	81.16 $\pm$ 11.86	0.262	0.794
Smoking (n, %)	Yes	53 (28.0%)	47 (24.7%)	0.533	0.465	36 (25.0%)	35 (25.2%)	0.001	0.972
	No	136 (72.0%)	143 (75.3%)			108 (75.0%)	104 (74.8%)		
Drinking (n, %)	Yes	69 (36.5%)	30 (15.8%)	21.075	4.00E–06	47 (32.6%)	47 (33.8%)	0.044	0.834
	No	120 (63.5%)	160 (84.2%)			97 (67.4%)	92 (66.2%)		
History of hypertension (n, %)	Yes	155 (82.0%)	125 (65.8%)	12.919	0.003	87 (60.4%)	96 (69.1%)	2.315	0.128
	No	34 (18.0%)	65 (34.2%)			57 (39.6%)	43 (30.9%)		
History of diabetes (n, %)	Yes	71 (37.6%)	51 (26.8%)	4.992	0.025	34 (23.6%)	42 (30.2%)	1.571	0.210
	No	118 (62.4%)	139 (73.2%)			110 (76.4%)	97 (69.8%)		
TC (mean $\pm$ SD, mmol/L)		4.84 $\pm$ 0.82	4.97 $\pm$ 4.42	0.392	0.695	5.11 $\pm$ 0.91	5.20 $\pm$ 0.93	0.869	0.386
TG (mean $\pm$ SD, mmol/L)		1.64 $\pm$ 1.29	1.80 $\pm$ 1.45	1.113	0.267	1.56 $\pm$ 1.00	1.56 $\pm$ 0.88	0.014	0.989
HDL-C (mean $\pm$ SD, mmol/L)		1.49 $\pm$ 0.39	1.27 $\pm$ 0.38	5.453	9.13E–08	1.54 $\pm$ 0.46	1.55 $\pm$ 0.42	0.179	0.858
LDL-C (mean $\pm$ SD, mmol/L)		2.65 $\pm$ 0.69	2.76 $\pm$ 2.00	0.744	0.457	2.88 $\pm$ 0.80	2.96 $\pm$ 0.75	0.848	0.397
Fasting glucose (mean $\pm$ SD, mmol/L)		6.77 $\pm$ 2.01	6.62 $\pm$ 7.26	0.254	0.800	6.61 $\pm$ 2.26	6.64 $\pm$ 2.24	0.126	0.900

SBP systolic blood pressure, DBP diastolic blood pressure, TC total cholesterol, TG triglycerides, HDL-C high-density lipoprotein cholesterol, LDL-C low-density lipoprotein cholesterol

### Association between *HTRA1* methylation and stroke in the case–control study and the prospective nested case–control study

To evaluate the association between *HTRA1* methylation and the risk of stroke in the case–control study and the prospective nested case–control study, two amplicons, HTRA1-A and HTRA1-B, including 17 measurable CpG sites were semiquantitatively determined by Agena MALDI-TOF mass spectrometry.

In the case–control study (190 cases vs. 190 controls), HTRA1\_A\_CpG\_6 showed significantly lower methylation levels in stroke cases than in the controls [median = 0.10 (interquartile range, IQR = 0.06–0.13) and 0.12 (IQR = 0.10–0.15) for cases and controls, respectively, OR per +10% methylation = 0.54, 95% CI 0.34–0.86,  $P = 0.009$  by logistic regression adjusted for age, sex, smoking, drinking, hypertension, diabetes, TC, TG, HDL-C and LDL-C, Table 2]. The other CpG sites in the HTRA1 amplicons showed no significant associations with stroke (Table 2).

In the prospective nested case–control study, comparing 144 controls and all 139 cases who developed stroke within 2 years, the methylation levels of three measurable CpG loci (CpG\_3.4, CpG\_8.9, and CpG\_13.14.15) in the

HTRA1-B amplicon were significantly higher in cases than in controls (ORs per +10% methylation from 1.33 to 1.42,  $P \leq 0.034$  for all by logistic regression adjusted for age, sex, smoking, drinking, hypertension, diabetes, TC, TG, HDL-C and LDL-C, Fig. 1A, Additional file 1: Table S1).

Since the onset time of stroke may have an effect on the methylation levels of *HTRA1*, we further performed analyses comparing controls and stroke cases stratified by time of stroke onset in the prospective nested case–control study. Interestingly, the differences between the aforementioned three CpG loci became more pronounced in stroke cases with onset time  $\leq 1.5$  years (ORs per +10% methylation from 1.43 to 1.58,  $P \leq 0.016$  for all by logistic regression adjusted for age, sex, smoking, drinking, hypertension, diabetes, TC, TG, HDL-C and LDL-C, Fig. 1B, Additional file 1: Table S2) and in cases with onset time  $\leq 1.32$  years (ORs per +10% methylation from 1.43 to 1.59,  $P \leq 0.025$  for all by logistic regression adjusted for age, sex, smoking, drinking, hypertension, diabetes, TC, TG, HDL-C and LDL-C, Fig. 1C, Additional file 1: Table S3), especially in stroke cases with onset time  $\leq 1$  year (ORs per +10% methylation from 1.67 to 1.88,  $P \leq 0.022$  for all by logistic regression adjusted for

**Table 2** Methylation difference of *HTRA1* between 190 stroke cases and 190 controls in the case–control study

CpG sites	Controls	Stroke cases	Crude OR (95% CI)	P value	OR (95% CI)*	P value*	OR (95% CI)*	P value*
	Median (IQR)	Median (IQR)	Per + 10% methylation		Per + 10% methylation		Per + 10% methylation	
HTRA1_A_CpG_1	0.09 (0.05–0.12)	0.10 (0.07–0.13)	1.27 (0.92–1.75)	0.146	1.27 (0.92–1.76)	0.140	1.22 (0.85–1.77)	0.284
HTRA1_A_CpG_2	0.65 (0.56–0.73)	0.64 (0.58–0.74)	1.01 (0.92–1.12)	0.803	1.01 (0.92–1.12)	0.821	0.99 (0.88–1.11)	0.860
HTRA1_A_CpG_3	0.29 (0.22–0.35)	0.29 (0.24–0.35)	0.96 (0.78–1.17)	0.662	0.95 (0.78–1.16)	0.636	0.92 (0.73–1.15)	0.466
HTRA1_A_CpG_4	0.12 (0.09–0.15)	0.13 (0.11–0.15)	1.48 (0.96–2.30)	0.078	1.47 (0.95–2.28)	0.087	1.29 (0.79–2.11)	0.310
HTRA1_A_CpG_5	0.37 (0.31–0.45)	0.38 (0.30–0.45)	0.99 (0.86–1.14)	0.891	0.99 (0.86–1.14)	0.903	0.99 (0.85–1.16)	0.913
HTRA1_A_CpG_6	0.12 (0.10–0.15)	0.10 (0.06–0.13)	0.50 (0.32–0.78)	0.002	0.51 (0.33–0.79)	0.003	0.54 (0.34–0.86)	<b>0.009</b>
HTRA1_A_CpG_7	0.36 (0.28–0.45)	0.37 (0.30–0.45)	1.08 (0.92–1.27)	0.343	1.07 (0.92–1.26)	0.378	1.09 (0.91–1.30)	0.369
HTRA1_B_CpG_1	0.16 (0.09–0.41)	0.18 (0.11–0.44)	1.04 (0.95–1.13)	0.384	1.04 (0.96–1.14)	0.333	1.05 (0.95–1.16)	0.308
HTRA1_B_CpG_2	0.09 (0.05–0.12)	0.10 (0.07–0.14)	1.45 (1.02–2.06)	0.039	1.45 (1.02–2.06)	0.040	1.25 (0.86–1.82)	0.243
HTRA1_B_CpG_3.4	0.28 (0.24–0.33)	0.29 (0.25–0.34)	1.18 (0.90–1.55)	0.236	1.19 (0.90–1.57)	0.221	1.11 (0.81–1.52)	0.513
HTRA1_B_CpG_5	0.26 (0.20–0.34)	0.28 (0.22–0.33)	1.09 (0.90–1.33)	0.378	1.10 (0.90–1.34)	0.367	1.06 (0.85–1.32)	0.612
HTRA1_B_CpG_7	0.47 (0.36–0.54)	0.47 (0.37–0.54)	1.06 (0.92–1.22)	0.456	1.06 (0.92–1.22)	0.448	1.01 (0.85–1.19)	0.957
HTRA1_B_CpG_8.9	0.28 (0.24–0.33)	0.29 (0.25–0.34)	1.18 (0.90–1.55)	0.236	1.19 (0.90–1.57)	0.221	1.11 (0.81–1.52)	0.513
HTRA1_B_CpG_10	0.61 (0.53–0.71)	0.63 (0.55–0.70)	1.06 (0.92–1.23)	0.445	1.06 (0.92–1.23)	0.423	1.03 (0.87–1.22)	0.722
HTRA1_B_CpG_11.12	0.72 (0.65–0.80)	0.73 (0.66–0.80)	1.09 (0.92–1.30)	0.315	1.10 (0.92–1.31)	0.300	1.07 (0.88–1.31)	0.504
HTRA1_B_CpG_13.14.15	0.49 (0.41–0.57)	0.51 (0.45–0.57)	1.15 (0.97–1.36)	0.115	1.15 (0.97–1.36)	0.109	1.12 (0.92–1.36)	0.261
HTRA1_B_CpG_16	0.53 (0.45–0.63)	0.53 (0.45–0.68)	1.03 (0.91–1.17)	0.636	1.03 (0.91–1.17)	0.658	0.99 (0.86–1.15)	0.911

Bold values indicated  $P < 0.05$

\* Logistic regression, adjusted for age and sex;

\* Logistic regression, adjusted for age, sex, smoking, drinking, hypertension, diabetes, TC, TG, HDL-C and LDL-C

age, sex, smoking, drinking, hypertension, diabetes, TC, TG, HDL-C and LDL-C, Fig. 1D, Additional file 1: Table S4). Additionally, CpG\_5, CpG\_7, CpG\_10 and CpG\_11.12 in the HTRA1-B amplicon showed borderline associations with stroke in cases with shorter onset time (Fig. 1A–D, Additional file 1: Tables S1–S4). Moreover, stroke cases with onset time  $\leq 1.5$  years showed significantly lower HTRA1\_A\_CpG\_6 methylation levels than controls (OR per +10% methylation = 0.39, 95% CI 0.17–0.90,  $P=0.027$  by logistic regression adjusted for age, sex, smoking, drinking, hypertension, diabetes, TC, TG, HDL-C and LDL-C, Fig. 1B, Additional file 1: Table S2). Similarly, the differences became more pronounced in stroke cases with onset time  $\leq 1.32$  years (OR per +10% methylation = 0.10, 95% CI 0.03–0.42,  $P=0.002$  by logistic regression adjusted for age, sex, smoking, drinking, hypertension, diabetes, TC, TG, HDL-C and LDL-C, Fig. 1C, Additional file 1: Table S3), and in cases with onset time  $\leq 1$  year (OR per +10% methylation = 0.12, 95% CI 0.02–0.70,  $P=0.018$  by logistic regression adjusted for age, sex, smoking, drinking, hypertension, diabetes, TC, TG, HDL-C and LDL-C, Fig. 1D, Additional file 1: Table S4). HTRA1\_A\_CpG\_7 and HTRA1\_A\_CpG\_5 also showed significant association with stroke in cases with onset time  $\leq 1.5$  years and  $\leq 1$  year, respectively ( $P=0.046$  and  $0.007$  by logistic regression adjusted for age, sex, smoking, drinking, hypertension, diabetes, TC, TG, HDL-C and LDL-C, Fig. 1B and D Additional file 1: Table S2, Additional file 1: Table S4). Overall, the OR per +10% methylation showed clear trends of decrease in HTRA1\_A\_CpG\_6 in stroke cases with onset time  $< 2$  years,  $\leq 1.5$  years,  $\leq 1.32$  years and  $\leq 1$  year (Fig. 1E). The OR per +10% methylation showed clear trends of increase in CpG\_3.4, CpG\_5, CpG\_7, CpG\_8.9, CpG\_10, CpG\_11.12 and CpG\_13.14.15 in the HTRA1-B amplicon in stroke cases with onset time  $< 2$  years,  $\leq 1.5$  years,  $\leq 1.32$  years and  $\leq 1$  year (Fig. 1F).

#### Association between blood-based methylation of *HTRA1* and stroke in the older population

Age was reported to play an essential role in the pattern of DNA methylation [25]; we therefore evaluated the association between *HTRA1* methylation in the blood and stroke stratified by two major age groups ( $< 70$  and  $\geq 70$  years). In the case–control study, stroke

cases showed significantly lower methylation levels of HTRA1\_A\_CpG\_3 and HTRA1\_A\_CpG\_6 than controls among participants  $\geq 70$  years old (ORs per +10% methylation = 0.63 and 0.29,  $P=0.022$  and  $0.008$ , respectively, by logistic regression adjusted for age, sex, smoking, drinking, hypertension, diabetes, TC, TG, HDL-C and LDL-C, Table 3). However, no association between *HTRA1* methylation and stroke was observed among subjects  $< 70$  years old (Table 3).

In the prospective nested case–control study, the methylation levels of CpG\_3.4, CpG\_5, CpG\_8.9 and CpG\_13.14.15 in the HTRA1\_B amplicon were significantly higher in stroke cases than in controls among participants  $\geq 70$  years old (ORs per +10% methylation from 1.70 to 2.36,  $P \leq 0.021$  for all by logistic regression adjusting for age, sex, smoking, drinking, hypertension, diabetes, TC, TG, HDL-C and LDL-C, Table 4). This association was not observed in subjects  $< 70$  years old in the prospective nested case–control study as well (Table 4).

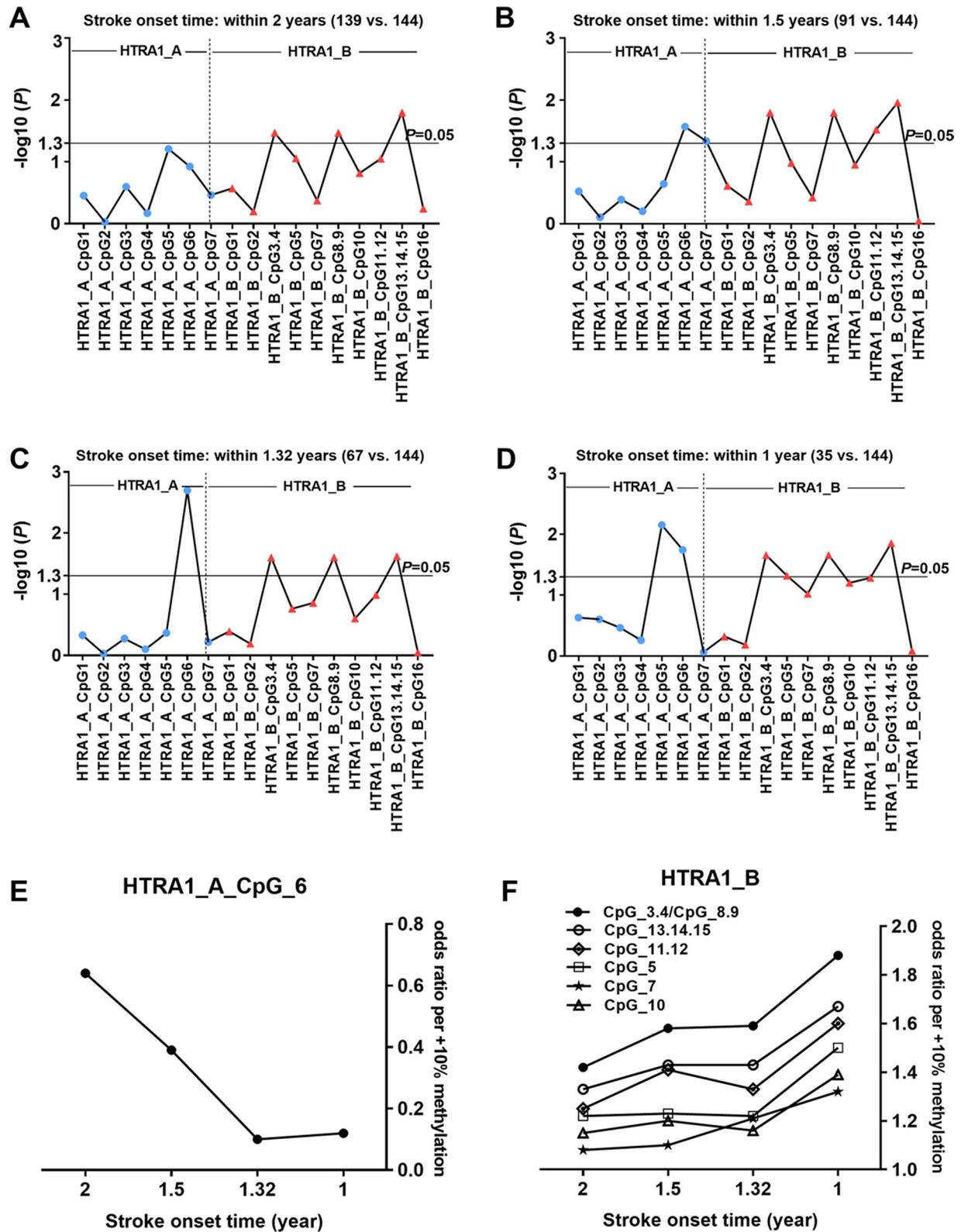
#### Correlation between *HTRA1* methylation and stroke-related characteristics

To explore the relationship between the blood-based *HTRA1* methylation and stroke-related characteristics, the subjects including both stroke cases and controls with available data were interpreted. In the case–control study, weak correlations were observed between the methylation levels of a few CpG sites in the HTRA1 amplicons and hypertension, HDL-C, current smoking status, TC as well as TG levels (Additional file 1: Table S5). None of the 17 CpG loci showed any correlation with drinking, diabetes and levels of LDL-C (Additional file 1: Table S5).

In the prospective nested case–control study, the methylation level of HTRA1\_A\_CpG\_2 was positively correlated with current drinking status (Spearman rho = 0.199,  $P=0.001$ , Additional file 1: Table S6), while the methylation levels of CpG\_3, CpG\_4 and CpG\_5 in HTRA1\_A amplicon were inversely correlated with hypertension (Spearman rho =  $-0.121$ ,  $-0.263$  and  $-0.198$ , respectively, Additional file 1: Table S6). No correlations were observed between individual CpG sites and smoking, diabetes, levels of TC, TG, HDL-C and LDL-C (Additional file 1: Table S6).

(See figure on next page.)

**Fig. 1** Onset-time-dependent association between *HTRA1* methylation and stroke in the prospective nested case–control study. The analysis was performed for stroke cases with onset time **A**  $< 2$  years, **B**  $\leq 1.5$  years, **C**  $\leq 1.32$  years, and **D**  $\leq 1$  year. The  $P$  values of all the 17 measurable CpG loci in the HTRA1-A and HTRA1-B amplicons were calculated by logistic regression adjusting for age, sex, smoking, drinking, hypertension, diabetes, TC, TG, HDL-C and LDL-C, and all the  $P$  values were transformed by  $-\log_{10}$ . The dotted lines indicate the thresholds of different  $P$  values of 0.05. **E** and **F** The OR per +10% methylation of HTRA1\_A\_CpG\_6 and CpG\_3.4, CpG\_5, CpG\_7, CpG\_8.9, CpG\_10, CpG\_11.12 as well as CpG\_13.14.15 in the HTRA1-B amplicon in stroke cases with onset time  $< 2$  years,  $\leq 1.5$  years,  $\leq 1.32$  years and  $\leq 1$  year based on logistic regression analysis adjusting for age, sex, smoking, drinking, hypertension, diabetes, TC, TG, HDL-C and LDL-C



**Fig. 1** (See legend on previous page.)

**Table 3** Methylation difference of *HTRA1* between stroke cases and controls stratified by age in the case-control study

CpG sites	Controls Median (IQR)	Stroke cases Median (IQR)	Crude OR (95% CI) Per + 10% methylation	P value	OR (95% CI)* Per + 10% methylation	P value*	OR (95% CI)* Per + 10% methylation	P value*
<i>Parti-</i> <i>pants &lt; 70 years old</i> (116 stroke cases vs. 125 controls)								
HTRA1_A_CpG_1	0.09 (0.05–0.13)	0.11 (0.08–0.13)	1.23 (0.84–1.81)	0.293	1.23 (0.84–1.81)	0.293	1.20 (0.76–1.89)	0.448
HTRA1_A_CpG_2	0.65 (0.57–0.79)	0.64 (0.56–0.74)	0.98 (0.86–1.11)	0.707	0.98 (0.86–1.11)	0.696	0.96 (0.83–1.10)	0.510
HTRA1_A_CpG_3	0.28 (0.22–0.35)	0.30 (0.25–0.35)	1.10 (0.85–1.41)	0.484	1.10 (0.85–1.41)	0.489	1.13 (0.85–1.51)	0.409
HTRA1_A_CpG_4	0.12 (0.09–0.15)	0.13 (0.11–0.15)	1.31 (0.74–2.30)	0.354	1.30 (0.74–2.29)	0.369	1.02 (0.54–1.94)	0.951
HTRA1_A_CpG_5	0.37 (0.31–0.43)	0.38 (0.30–0.43)	0.96 (0.81–1.15)	0.679	0.96 (0.81–1.15)	0.680	0.97 (0.79–1.18)	0.731
HTRA1_A_CpG_6	0.12 (0.10–0.14)	0.10 (0.06–0.14)	0.66 (0.40–1.08)	0.097	0.66 (0.40–1.09)	0.106	0.69 (0.41–1.16)	0.163
HTRA1_A_CpG_7	0.36 (0.28–0.45)	0.37 (0.31–0.45)	1.04 (0.86–1.27)	0.685	1.04 (0.85–1.27)	0.713	1.06 (0.85–1.34)	0.602
HTRA1_B_CpG_1	0.16 (0.09–0.40)	0.18 (0.11–0.50)	1.09 (0.98–1.21)	0.122	1.09 (0.98–1.22)	0.123	1.11 (0.98–1.25)	0.107
HTRA1_B_CpG_2	0.09 (0.05–0.12)	0.11 (0.07–0.15)	1.66 (0.93–3.03)	0.102	1.65 (0.92–3.01)	0.103	1.80 (0.96–3.23)	0.062
HTRA1_B_CpG_3,4	0.28 (0.23–0.32)	0.29 (0.25–0.34)	1.41 (0.97–2.03)	0.069	1.42 (0.98–2.05)	0.063	1.29 (0.85–1.96)	0.230
HTRA1_B_CpG_5	0.25 (0.19–0.33)	0.28 (0.22–0.33)	1.19 (0.92–1.54)	0.190	1.20 (0.93–1.56)	0.169	1.18 (0.88–1.58)	0.274
HTRA1_B_CpG_7	0.47 (0.35–0.54)	0.46 (0.38–0.53)	1.07 (0.88–1.29)	0.520	1.06 (0.88–1.29)	0.535	0.99 (0.79–1.23)	0.906
HTRA1_B_CpG_8,9	0.28 (0.23–0.32)	0.29 (0.25–0.34)	1.41 (0.97–2.03)	0.069	1.42 (0.98–2.05)	0.063	1.29 (0.85–1.96)	0.230
HTRA1_B_CpG_10	0.61 (0.53–0.71)	0.64 (0.55–0.70)	1.14 (0.94–1.38)	0.180	1.14 (0.94–1.38)	0.173	1.13 (0.92–1.40)	0.249
HTRA1_B_CpG_11,12	0.72 (0.64–0.79)	0.73 (0.67–0.80)	1.13 (0.89–1.43)	0.330	1.12 (0.89–1.42)	0.343	1.11 (0.85–1.45)	0.451
HTRA1_B_CpG_13,14,15	0.48 (0.40–0.56)	0.50 (0.45–0.57)	1.24 (0.99–1.55)	0.066	1.24 (0.99–1.55)	0.068	1.19 (0.92–1.55)	0.186
HTRA1_B_CpG_16	0.51 (0.43–0.62)	0.53 (0.44–0.66)	1.08 (0.92–1.27)	0.367	1.08 (0.92–1.28)	0.356	1.03 (0.85–1.24)	0.800
<i>Parti-</i> <i>pants ≥ 70 years old</i> (74 stroke cases vs. 65 controls)								
HTRA1_A_CpG_1	0.09 (0.06–0.12)	0.10 (0.07–0.14)	1.38 (0.77–2.45)	0.278	1.37 (0.77–2.44)	0.282	1.43 (0.73–2.82)	0.300
HTRA1_A_CpG_2	0.63 (0.56–0.71)	0.65 (0.59–0.78)	1.08 (0.92–1.27)	0.335	1.09 (0.92–1.28)	0.340	1.05 (0.85–1.30)	0.631
HTRA1_A_CpG_3	0.32 (0.25–0.38)	0.28 (0.22–0.36)	0.76 (0.54–1.05)	0.097	0.74 (0.53–1.04)	0.087	0.63 (0.42–0.94)	<b>0.022</b>
HTRA1_A_CpG_4	0.12 (0.09–0.15)	0.12 (0.11–0.15)	1.79 (0.87–3.71)	0.116	1.80 (0.87–3.74)	0.113	1.92 (0.78–4.70)	0.156
HTRA1_A_CpG_5	0.37 (0.3–0.46)	0.38 (0.29–0.49)	1.04 (0.82–1.32)	0.750	1.04 (0.82–1.32)	0.752	1.01 (0.76–1.33)	0.967
HTRA1_A_CpG_6	0.12 (0.10–0.15)	0.10 (0.07–0.13)	0.27 (0.12–0.63)	0.003	0.27 (0.11–0.62)	0.002	0.29 (0.12–0.73)	<b>0.008</b>
HTRA1_A_CpG_7	0.36 (0.27–0.46)	0.38 (0.30–0.48)	1.16 (0.88–1.53)	0.282	1.15 (0.88–1.52)	0.310	1.12 (0.81–1.53)	0.497
HTRA1_B_CpG_1	0.15 (0.09–0.42)	0.17 (0.09–0.42)	0.96 (0.83–1.10)	0.557	0.95 (0.83–1.10)	0.529	0.94 (0.79–1.13)	0.511
HTRA1_B_CpG_2	0.08 (0.05–0.14)	0.09 (0.07–0.12)	0.92 (0.56–1.52)	0.756	0.92 (0.56–1.52)	0.749	0.70 (0.39–1.28)	0.251
HTRA1_B_CpG_3,4	0.30 (0.25–0.34)	0.29 (0.23–0.34)	0.92 (0.60–1.41)	0.701	0.93 (0.60–1.44)	0.737	0.85 (0.50–1.43)	0.534
HTRA1_B_CpG_5	0.27 (0.22–0.36)	0.28 (0.21–0.34)	0.96 (0.71–1.30)	0.779	0.97 (0.71–1.32)	0.822	0.87 (0.60–1.26)	0.455
HTRA1_B_CpG_7	0.48 (0.38–0.55)	0.49 (0.34–0.56)	1.04 (0.84–1.29)	0.731	1.05 (0.84–1.32)	0.650	1.00 (0.75–1.33)	0.999
HTRA1_B_CpG_8,9	0.30 (0.25–0.34)	0.29 (0.23–0.34)	0.92 (0.60–1.41)	0.701	0.93 (0.60–1.44)	0.737	0.85 (0.50–1.43)	0.534
HTRA1_B_CpG_10	0.63 (0.55–0.73)	0.62 (0.53–0.72)	0.94 (0.74–1.19)	0.615	0.94 (0.74–1.20)	0.619	0.78 (0.57–1.07)	0.117
HTRA1_B_CpG_11,12	0.72 (0.66–0.82)	0.72 (0.65–0.82)	1.06 (0.81–1.37)	0.685	1.06 (0.81–1.39)	0.669	0.96 (0.67–1.37)	0.815
HTRA1_B_CpG_13,14,15	0.51 (0.42–0.62)	0.51 (0.44–0.57)	1.02 (0.79–1.32)	0.863	1.03 (0.79–1.34)	0.819	0.98 (0.70–1.38)	0.914
HTRA1_B_CpG_16	0.55 (0.47–0.67)	0.53 (0.45–0.69)	0.95 (0.78–1.17)	0.642	0.96 (0.78–1.17)	0.659	0.92 (0.72–1.17)	0.496

Bold values indicated  $P < 0.05$ 

\* logistic regression, adjusted for age and sex

\* logistic regression, adjusted for age, sex, smoking, drinking, hypertension, diabetes, TC, TG, HDL-C and LDL-C

**Table 4** Methylation difference of *HTRA1* between stroke cases and controls stratified by age in the prospective nested case–control study

CpG sites	Controls Median (IQR)	Stroke cases Median (IQR)	Crude OR (95% CI) Per + 10% methylation	P value	OR (95% CI)* Per + 10% methylation	P value*	OR (95% CI)* Per + 10% methylation	P value*
<i>Participants &lt; 70 years old (88 stroke cases vs. 90 controls)</i>								
HTRA1_A_CpG_1	0.07 (0.05–0.10)	0.07 (0.05–0.09)	0.72 (0.33–1.61)	0.427	0.72 (0.33–1.61)	0.428	0.67 (0.30–1.54)	0.348
HTRA1_A_CpG_2	0.78 (0.74–0.84)	0.77 (0.73–0.82)	1.09 (0.85–1.40)	0.504	1.08 (0.84–1.39)	0.542	1.15 (0.88–1.50)	0.316
HTRA1_A_CpG_3	0.44 (0.41–0.51)	0.46 (0.41–0.52)	1.02 (0.75–1.39)	0.899	1.05 (0.76–1.44)	0.772	1.05 (0.75–1.46)	0.780
HTRA1_A_CpG_4	0.19 (0.16–0.21)	0.20 (0.17–0.22)	1.39 (0.65–2.99)	0.400	1.38 (0.63–2.99)	0.421	1.54 (0.69–3.43)	0.295
HTRA1_A_CpG_5	0.41 (0.35–0.49)	0.41 (0.35–0.46)	0.89 (0.67–1.17)	0.394	0.88 (0.66–1.16)	0.365	0.87 (0.65–1.16)	0.339
HTRA1_A_CpG_6	0.09 (0.08–0.11)	0.08 (0.07–0.10)	0.59 (0.30–1.18)	0.135	0.57 (0.28–1.14)	0.110	0.55 (0.27–1.12)	0.101
HTRA1_A_CpG_7	0.39 (0.34–0.45)	0.40 (0.35–0.45)	1.20 (0.82–1.76)	0.343	1.23 (0.84–1.81)	0.297	1.22 (0.82–1.81)	0.320
HTRA1_B_CpG_1	0.19 (0.11–0.29)	0.18 (0.11–0.30)	1.11 (0.91–1.36)	0.299	1.11 (0.91–1.36)	0.322	1.14 (0.93–1.41)	0.208
HTRA1_B_CpG_2	0.12 (0.09–0.16)	0.13 (0.08–0.17)	1.17 (0.69–2.01)	0.560	1.19 (0.69–2.04)	0.534	1.19 (0.69–2.07)	0.528
HTRA1_B_CpG_3,4	0.31 (0.26–0.35)	0.33 (0.27–0.37)	1.23 (0.85–1.79)	0.278	1.21 (0.83–1.77)	0.321	1.22 (0.83–1.79)	0.312
HTRA1_B_CpG_5	0.34 (0.28–0.40)	0.34 (0.26–0.43)	1.06 (0.80–1.41)	0.688	1.05 (0.79–1.39)	0.751	1.04 (0.78–1.40)	0.781
HTRA1_B_CpG_7	0.47 (0.39–0.53)	0.46 (0.38–0.58)	1.04 (0.83–1.32)	0.720	1.03 (0.81–1.30)	0.826	1.03 (0.80–1.31)	0.833
HTRA1_B_CpG_8,9	0.31 (0.26–0.35)	0.33 (0.27–0.37)	1.23 (0.85–1.79)	0.278	1.21 (0.83–1.77)	0.321	1.22 (0.83–1.79)	0.312
HTRA1_B_CpG_10	0.61 (0.55–0.69)	0.66 (0.56–0.75)	1.17 (0.91–1.51)	0.233	1.17 (0.90–1.51)	0.236	1.18 (0.90–1.54)	0.224
HTRA1_B_CpG_11,12	0.75 (0.69–0.79)	0.75 (0.70–0.82)	1.14 (0.83–1.58)	0.428	1.12 (0.81–1.55)	0.509	1.11 (0.79–1.55)	0.558
HTRA1_B_CpG_13,14,15	0.54 (0.46–0.61)	0.56 (0.47–0.61)	1.17 (0.88–1.54)	0.281	1.15 (0.87–1.53)	0.314	1.16 (0.87–1.54)	0.317
HTRA1_B_CpG_16	0.48 (0.42–0.58)	0.49 (0.37–0.57)	0.92 (0.76–1.12)	0.393	0.91 (0.75–1.11)	0.351	0.90 (0.74–1.10)	0.321
<i>Participants ≥ 70 years old (51 stroke cases vs. 53 controls)</i>								
HTRA1_A_CpG_1	0.07 (0.06–0.08)	0.07 (0.05–0.09)	0.73 (0.25–2.13)	0.563	0.66 (0.22–1.98)	0.457	0.67 (0.20–2.22)	0.511
HTRA1_A_CpG_2	0.76 (0.71–0.82)	0.76 (0.71–0.81)	0.86 (0.65–1.14)	0.303	0.87 (0.65–1.16)	0.330	0.86 (0.60–1.23)	0.418
HTRA1_A_CpG_3	0.44 (0.38–0.47)	0.46 (0.42–0.54)	1.78 (1.08–2.93)	0.024	1.77 (1.06–2.93)	0.028	1.69 (0.97–2.93)	0.063
HTRA1_A_CpG_4	0.20 (0.17–0.22)	0.19 (0.17–0.21)	1.02 (0.50–2.09)	0.962	1.01 (0.49–2.10)	0.975	0.90 (0.40–2.02)	0.803
HTRA1_A_CpG_5	0.44 (0.38–0.51)	0.40 (0.34–0.45)	0.71 (0.50–1.01)	0.054	0.72 (0.50–1.02)	0.065	0.62 (0.40–1.01)	0.053
HTRA1_A_CpG_6	0.09 (0.08–0.10)	0.08 (0.07–0.10)	0.94 (0.35–2.48)	0.895	1.02 (0.38–2.74)	0.973	0.99 (0.33–2.96)	0.988
HTRA1_A_CpG_7	0.39 (0.35–0.42)	0.40 (0.34–0.48)	1.10 (0.69–1.75)	0.688	1.06 (0.66–1.71)	0.806	0.91 (0.54–1.52)	0.711
HTRA1_B_CpG_1	0.17 (0.12–0.25)	0.18 (0.10–0.29)	1.08 (0.85–1.36)	0.530	1.06 (0.84–1.35)	0.621	0.98 (0.75–1.28)	0.900
HTRA1_B_CpG_2	0.13 (0.09–0.17)	0.11 (0.08–0.16)	0.94 (0.52–1.69)	0.832	0.98 (0.54–1.77)	0.936	1.18 (0.60–2.31)	0.636
HTRA1_B_CpG_3,4	0.30 (0.25–0.35)	0.32 (0.27–0.37)	1.72 (0.97–3.06)	0.064	1.73 (0.97–3.08)	0.064	2.36 (1.18–4.73)	<b>0.015</b>
HTRA1_B_CpG_5	0.31 (0.26–0.39)	0.38 (0.28–0.47)	1.50 (1.03–2.19)	0.033	1.59 (1.07–2.37)	0.023	1.70 (1.08–2.67)	<b>0.021</b>
HTRA1_B_CpG_7	0.49 (0.40–0.53)	0.48 (0.41–0.57)	1.10 (0.82–1.48)	0.540	1.09 (0.81–1.47)	0.583	1.09 (0.78–1.52)	0.614
HTRA1_B_CpG_8,9	0.30 (0.25–0.35)	0.32 (0.27–0.37)	1.72 (0.97–3.06)	0.064	1.73 (0.97–3.08)	0.064	2.36 (1.18–4.73)	<b>0.015</b>
HTRA1_B_CpG_10	0.64 (0.54–0.69)	0.68 (0.53–0.76)	1.14 (0.87–1.49)	0.333	1.14 (0.87–1.49)	0.338	1.23 (0.91–1.65)	0.178
HTRA1_B_CpG_11,12	0.73 (0.68–0.78)	0.77 (0.68–0.83)	1.54 (1.01–2.35)	0.043	1.49 (0.98–2.28)	0.064	1.35 (0.85–2.15)	0.203
HTRA1_B_CpG_13,14,15	0.53 (0.47–0.57)	0.57 (0.50–0.67)	1.76 (1.16–2.69)	0.008	1.78 (1.15–2.76)	0.009	1.76 (1.09–2.86)	<b>0.021</b>
HTRA1_B_CpG_16	0.47 (0.39–0.55)	0.49 (0.37–0.60)	1.02 (0.79–1.30)	0.901	1.03 (0.80–1.32)	0.837	1.04 (0.79–1.37)	0.763

Bold values indicated  $P < 0.05$ 

\* logistic regression, adjusted for age and sex

\* logistic regression, adjusted for age, sex, smoking, drinking, hypertension, diabetes, TC, TG, HDL-C and LDL-C

### Expression of *HTRA1* in peripheral blood leukocytes

To understand the correlation between the methylation and expression of *HTRA1* gene, RNA was isolated from peripheral blood leukocytes of 96 stroke cases and 96 controls in the case–control study. The median of relative expression levels of *HTRA1* in the leukocytes of stroke cases was 1.68-fold lower than in the controls (*HTRA1* relative expression, median and IQR of cases: 0.56 (0.25–1.42), median and IQR of controls: 0.94 (0.36–2.52),  $P=0.023$ , Fig. 2A).

Next, we plan to measure the expression of *HTRA1* in peripheral blood leukocytes in a prospective study. However, no RNA materials are available from the nested case–control study mentioned above as well as in the materials and methods. Thus, we analyzed blood RNA from 48 stroke cases who developed stroke within 1 year after the enrollment and 48 matched stroke-free controls from another prospective study in Jurong (Additional file 1: Table S7). The median of relative expression levels of *HTRA1* in the leukocytes of stroke cases was 1.24-fold lower than in the controls (*HTRA1* relative expression, median and IQR of cases: 0.76 (0.49–1.34), median and IQR of controls: 0.94 (0.64–1.46), Fig. 2B), although the  $P$  value did not reach significance ( $P=0.054$ ) probably due to the limited sample size.

### *HTRA1* methylation as a marker for the early detection of stroke

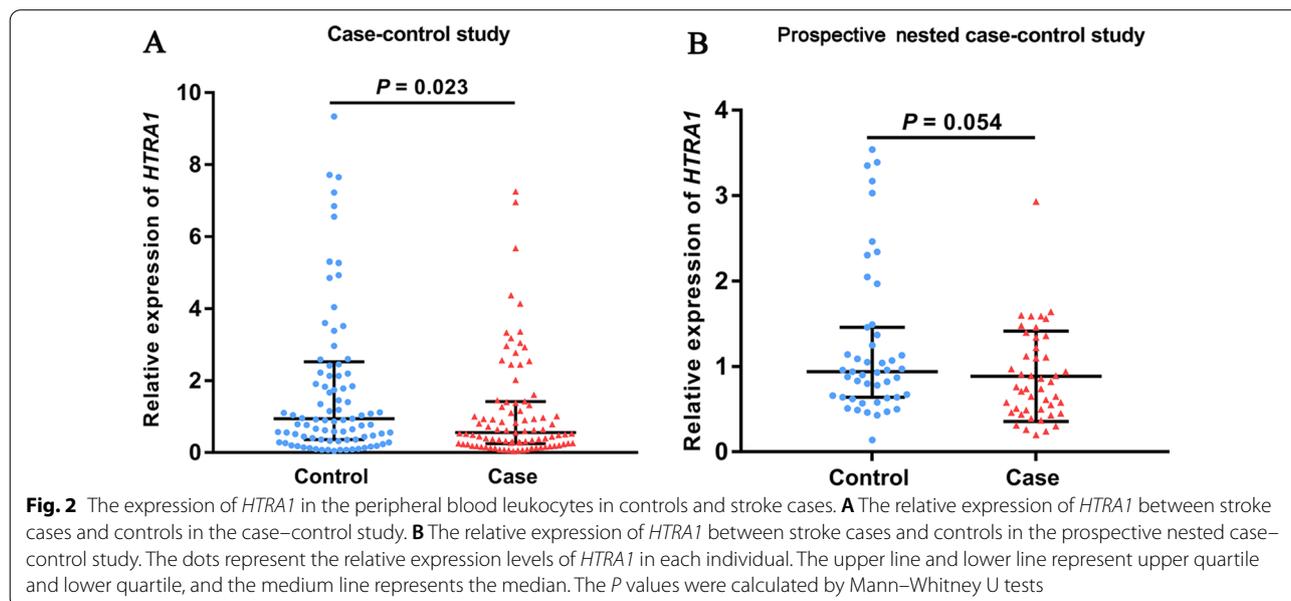
To estimate the potential clinical utility of *HTRA1* methylation as a marker for the early detection of stroke, ROC curve analyses were performed based on the methylation data generated from the prospective nested case–control

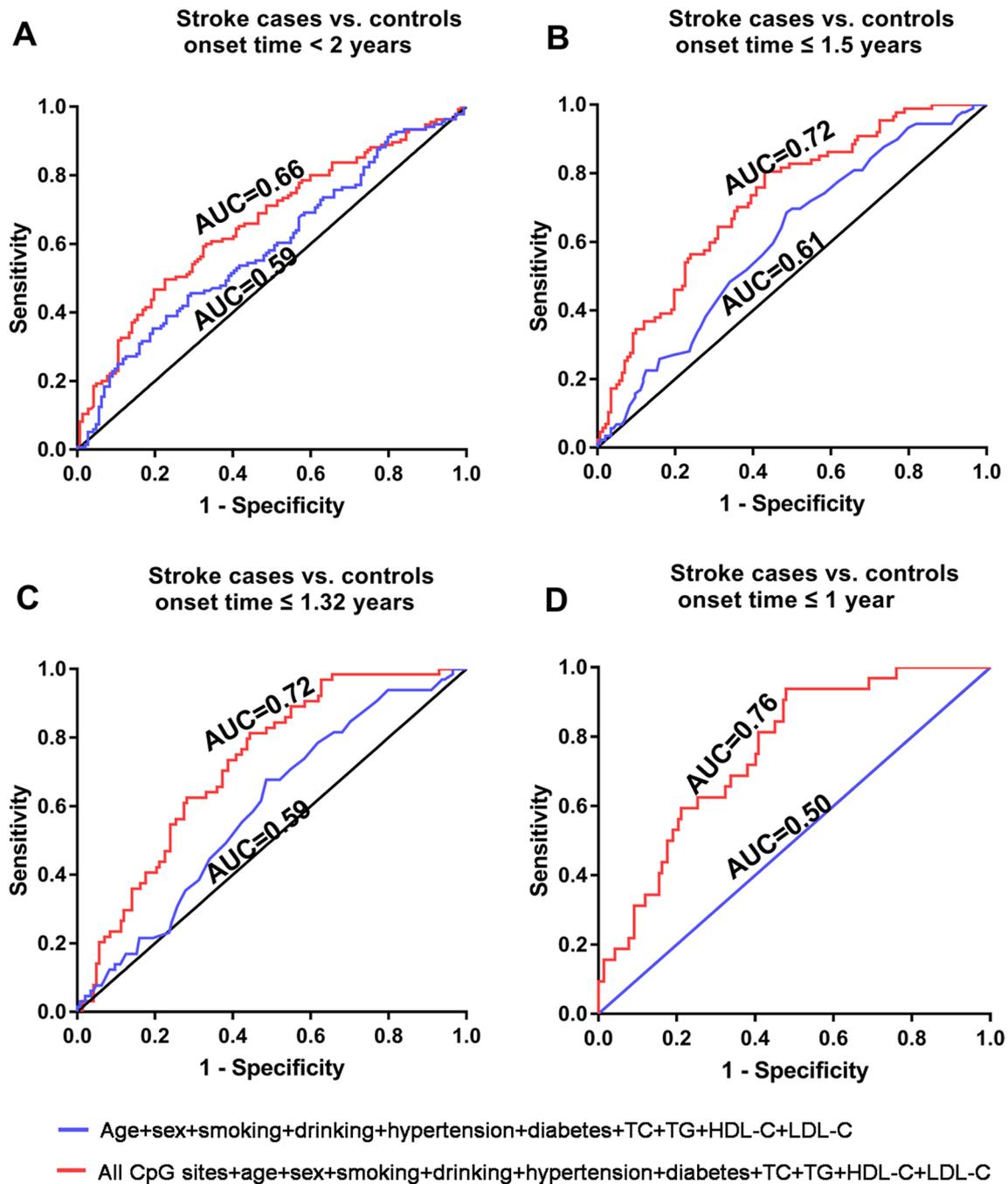
study using all the analyzed CpG sites and stroke-related variables including age, sex, smoking, drinking, hypertension, diabetes, TC, TG, HDL-C and LDL-C by logistic regression. The combination of all the analyzed CpG sites and stroke-related variables showed better discriminatory power for differentiating stroke cases at any time of onset than stroke-related variables alone (Fig. 3A–D). For stroke cases with shorter onset time, the *HTRA1* methylation and stroke-related variables showed increasing discriminatory power for differentiating stroke cases from healthy controls (area under the curve (AUC)=0.66, 0.72 and 0.72 for stroke cases with onset time < 2 years,  $\leq 1.5$  years and  $\leq 1.32$  years, respectively, Fig. 3A–C). Notably, the combination of *HTRA1* methylation and stroke-related variables had robust power to distinguish stroke cases with onset time  $\leq 1$  year from controls (AUC=0.76, 95% CI 0.67–0.84, Fig. 3D).

### Discussion

In our study, the stroke-related altered *HTRA1* methylation in blood cells was observed in both hospital-based case–control study and prospective nested case–control study. The prospective data further disclosed that the changed *HTRA1* methylation was detectable 2 years pre-clinical of stroke and became more pronounced 1.5 years and 1 year before the clinical determination of stroke. These findings provided pronounced evidence that differential DNA methylation signatures in the peripheral blood could be a potential biomarker for the risk evaluation and preclinical detection of stroke.

*HTRA1*, a primarily secreted serine protease, degrades a number of substrates that involve in a variety of disease





**Fig. 3** Methylation level of *HTRA1* in the peripheral blood DNA as a marker for the early detection of stroke. ROC curve analyses for the discriminatory power of *HTRA1* methylation to distinguish stroke cases with onset time **A** < 2 years, **B**  $\leq 1.5$  years, **C**  $\leq 1.32$  years, and **D**  $\leq 1$  year from controls. The ROC analyses were performed by logistic regression using all the analyzed CpG sites and stroke-related variables including age, sex, smoking, drinking, hypertension, diabetes, TC, TG, HDL-C and LDL-C

processes such as neurodegeneration [26], age-related macular degeneration [27], and arthritis [28]. The loss of *HTRA1* function has been reported to cause cerebral autosomal recessive arteriopathy with subcortical

infarcts and leukoencephalopathy (CARASIL), an inherited form of cerebral SVD characterized by an early-onset stroke in young adults [29]. To date, little is yet known about the epigenetic impact of *HTRA1* on the initiation

and progression of vascular disorders including stroke. This study was novel in reporting a significant association between aberrant blood-based *HTRAI* and increased risk of stroke, and this association was more pronounced when the baseline was closer to the onset time of stroke. The prospective results suggested that the methylation level of *HTRAI* in blood DNA might be a preclinical marker for stroke. Moreover, significantly decreased expression of *HTRAI* was observed in the stroke cases, as well as in the preclinical stroke cases, indicating that the increased *HTRAI* methylation in blood cells might be biologically meaningful. Future prospective multicenter studies are warranted to verify these findings.

Many studies disclosed that the DNA methylation in human peripheral leukocytes could vary with age [30, 31]. In our case-control study and prospective study, a significant association between methylation of *HTRAI* and stroke was mainly observed in the older population, which was in accordance with the fact that the elderly men have a higher risk of stroke [32], and further indicated that *HTRAI* methylation might be an age-dependent factor for the risk of stroke. Alternatively, the methylation pattern might be affected by lifestyle such as alcohol drinking [33] and other stroke-related factors like hypertension [34]. In our prospective nested case-control study, we observed that the methylation level of *HTRAI*\_A\_CpG\_2 was positively correlated with current drinking status, while the methylation levels of CpG\_3, CpG\_4 and CpG\_5 in the *HTRAI*\_A amplicon were inversely correlated with hypertension. The specific mechanisms of DNA methylation and alcohol drinking as well as hypertension remained unclear. Further studies with larger sample sizes are needed to investigate the complex interactions between lifestyle, environmental factors, and DNA methylation throughout the progression of stroke.

Despite the prevalence of stroke, risk prediction and early detection remain challenging. Early detection of stroke is vital for providing timely treatment through thrombolysis, improving clinical outcomes and reducing resulting complications, and leading to better survival results and less disease burden [35, 36]. At present, magnetic resonance imaging (MRI) and computed tomography (CT) provide efficient tools for stroke diagnosis, but are only efficient at the onset of stroke. In this study, the blood-based *HTRAI* methylation can objectively predict stroke a year in advance with an area under the curve of 0.76. We should admit that the *HTRAI* methylation alone is not sufficient as a screening tool, the combination of multiple biomarkers including blood-based DNA methylation profiles and others as well as stroke-related environmental factors might be a solution for the screening and prediction of stroke in preclinic.

The long-term follow-up of stroke incidence and prospectively collected samples with careful uniform processing and storage in this study could exclude the possibility of observed differences between cases and controls owing to potential treatment or processing effects. Although with a 2-year follow-up in a cohort of 11,151 subjects, only 139 stroke cases were identified after the enrollment. This observed *HTRAI* methylation-related risk of stroke was still based on a relatively small sample size. The association between *HTRAI* methylation and stroke in the case-control study was not completely consistent with the prospective nested case-control study when each CpG was considered. This may also be attributed to the limited sample size. Further prospective studies with larger sample sizes are needed for validation and may also provide additional insights.

## Conclusions

In conclusion, this study revealed a significant correlation between altered *HTRAI* methylation and stroke in preclinic, thus suggesting the potential of blood-based DNA methylation for the risk evaluation and even prevention of stroke. Further validation studies are warranted to ensure the generalizability of our findings and evaluate the potential usage of *HTRAI* methylation as a clinical biomarker for stroke.

## Methods

### Study population

#### *Hospital-based case-control study*

The case-control study included 190 stroke patients and 190 controls matched for age and sex. The patients (mean age of  $64.99 \pm 9.85$  years) were recruited from the Affiliated Hospital of Xuzhou Medical University, Affiliated Jiangning Hospital of Nanjing Medical University and Jurong People's Hospital from January 2019 to August 2020. The clinical diagnosis of stroke was based on the clinical examination and confirmed by MRI or CT. The diagnostic information and clinical examination data were obtained from hospital medical records. The demographic information including age, sex, nationality, status of smoking and status of drinking and medical history of patients were collected using standardized questionnaires. The controls (mean age of  $65.89 \pm 10.12$  years) were recruited from a cohort study conducted from November 2018 to January 2019 in Jurong City, Jiangsu Province, and those who remained stroke-free during the 2-year follow-up were selected. Blood samples were collected after 8 h from the last meal or after an overnight fast to measure the levels of fasting plasma glucose, TC, TG, HDL-C, and LDL-C. The individuals were considered as diabetes cases in the presence of fasting plasma glucose  $\geq 7.0$  mmol/L or a self-reported diabetes history

according to the diagnostic criteria for diabetes suggested by the American Diabetes Association. Additionally, blood samples of all the cases and controls were processed following the same protocol and stored at  $-80^{\circ}\text{C}$  after drawing.

#### **Prospective nested case–control study**

Subjects for this nested case–control study were selected from an ongoing population-based cohort study conducted in Jurong City, Jiangsu Province, which has been described in our previous study [16]. Briefly, a total of 11,151 subjects aged 18 years or older were recruited in this cohort from October to November of 2015, and all the subjects were reported to be stroke-free at baseline. Detailed information on demography and status of smoking and status of drinking as well as family history were collected via questionnaires, along with the collection of blood samples and anthropometric measurements at enrollment. The incidence of stroke was identified by records from the local hospitals and health service centers in the community. All the 139 subjects who developed a stroke within 2 years after the enrollment were included in this study as cases (mean age of  $67.64 \pm 9.51$  years), and 150 participants who remained stroke-free during the 2.71-year follow-up time were selected as controls. After excluding 6 low-quality clotted blood samples, 144 controls (mean age of  $67.77 \pm 9.11$  years) who well matched the stroke patients by age and sex were included.

#### **DNA processing and bisulfite conversion**

The Genomic DNA Extraction Kit (Rebece, Nanjing, China) was used to extract human peripheral blood cells' genomic DNA as described previously [37]. Then the extracted DNA was further bisulfate-converted by EZ-96 DNA Methylation Gold Kit (Zymo Research, Orange County, USA) according to the manufacturer's protocol. After bisulfite treatment, all non-methylated cytosine (C) bases were converted to uracil (U), whereas all methylated C bases were kept unchanged.

#### **MALDI-TOF mass spectrometry**

DNA methylation of *HTRA1* in the peripheral blood was quantitatively measured using MALDI-TOF mass spectrometry (Agena Bioscience, San Diego, California, USA) as described by Yang et al. [38, 39]. Briefly, the bisulfite-treated genomic DNA was amplified by bisulfite-specific primers. The primers and sequences of polymerase chain reaction (PCR) amplicons (*HTRA1-A* and *HTRA1-B*) are presented in Additional file 1: Fig. S1. The *HTRA1-A* amplicon in the gene promoter region is located at 664 bp upstream of the nearby CpG island. The *HTRA1-B* amplicon in the first intron is located at 120 bp downstream of the nearby CpG island. Alternatively, these two

amplicons are located at the so-called CpG island shores where differential methylation occurs frequently [40, 41]. There are no SNPs located at the primer regions or overlapped with any of the CpG sites in the amplicon. The PCR products were processed following shrimp alkaline phosphatase cleanup, T cleavage and Resin Clean steps. The final products were transferred to a SpectroCHIP G384 by a Nanodispenser RS1000 apparatus (Agena, USA). The chips were detected by MassARRAY spectrometry. The quantitative methylation levels of each CpG site or aggregate of multiple CpG sites were collected by EpiTYPER v1.2 software. For each batch of MassARRAY analyses, stroke case-and-control pairs were randomly arranged, and the samples were treated and analyzed in parallel in all the processes.

#### **Quantitative real-time PCR**

Total RNA was extracted from peripheral blood leukocytes using the QIAamp RNA Blood Mini Kit (Qiagen, Catalog no. 52304, Germany) and then reverse-transcribed into cDNA using the PrimeScript™ RT Reagent Kit (Takara, RR047A, Japan). Quantitative real-time PCR was performed to determine the expressions of *HTRA1* gene and the housekeeping gene glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*) as an endogenous control using  $2 \times$  SYBR Green qPCR Master Mix (Bimake, B21202, USA). The relative expression of *HTRA1* for each sample was normalized by *GAPDH* according to the  $2^{-\Delta\Delta Ct}$  method.

#### **Statistical analysis**

All statistical analyses were conducted in SPSS version 25.0 (SPSS Inc., Chicago, USA). The methylation levels of each CpG site and gene expression data with non-Gaussian distribution were expressed as median (interquartile range). Differences between cases and controls were analyzed with Mann–Whitney U tests. The association between the level of *HTRA1* methylation and stroke was assessed using logistic regression models with covariates-adjusted odds ratios (ORs) and 95% confidence intervals (CIs). The correlation between *HTRA1* methylation and stroke-related characteristics was evaluated by Spearman's rank correlation coefficients. Receiver operating characteristic (ROC) curve analysis was performed to assess the discriminatory power of methylation levels. A two-tailed  $P < 0.05$  was considered statistically significant.

#### **Abbreviations**

AUC: Area under the curve; CARASIL: Cerebral autosomal recessive arteriopathy with subcortical infarcts and leukoencephalopathy; CIs: Confidence intervals; CpG: Cytosine–phosphate–guanine; CT: Computed tomography; DBP: Diastolic blood pressure; HDL-C: High-density lipoprotein cholesterol; *HTRA1*: HTRA serine protease 1; LDL-C: Low-density lipoprotein cholesterol; MRI: Magnetic resonance imaging; ORs: Odds ratios; PCR: Polymerase chain

reaction; ROC: Receiver operating characteristic; SBP: Systolic blood pressure; SNP: Single nucleotide polymorphism; SVD: Small vessel disease; TC: Total cholesterol; TG: Triglycerides; TGF- $\beta$ : Transforming growth factor- $\beta$ .

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13148-022-01418-0>.

**Additional file 1. Table S1.** Methylation difference of *HTRA1* between 144 controls and 139 stroke cases with onset time < 2 years in the prospective nested case-control study. **Table S2.** Methylation difference of *HTRA1* between 144 controls and 91 stroke cases with onset time  $\leq$  1.5 years in the prospective nested case-control study. **Table S3.** Methylation difference of *HTRA1* between 144 controls and 67 stroke cases with onset time  $\leq$  1.32 years in the prospective nested case-control study. **Table S4.** Methylation difference of *HTRA1* between 144 controls and 35 stroke cases with onset time  $\leq$  1 year in the prospective nested case-control study. **Table S5.** Correlation between *HTRA1* methylation and stroke-related characteristics in the case-control study. **Table S6.** Correlation between *HTRA1* methylation and stroke-related characteristics in the prospective nested case-control study. **Table S7.** Description of the subjects for RNA analysis. **Fig. S1** Amplicon and primer design for MassARRAY methylation analysis.

### Acknowledgements

We would like to thank Yu Liu, Min Liu and Haifeng Xu for their contribution to the study. We would also like to express our gratitude to all of the participants of the contributing studies to each of these datasets.

### Author contributions

CL and ML analyzed the data and drafted the article. QY and YF collected the data. CS and RY performed the conception or design of the work, data interpretation and critical revision of the article. All authors reviewed the manuscript. All authors read and approved the final manuscript.

### Funding

This work was funded by the National Natural Science Foundation of China (Grant No. 81573232) and by the research funding from Nanjing Medical University. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

### Availability of data and materials

The datasets reported in the current study are available from the corresponding author upon reasonable request.

### Declarations

#### Ethics approval and consent to participate

The studies were approved by the Ethics Committee of Nanjing Medical University. All the recruited participants were Chinese Han and provided written informed consent.

#### Consent for publication

All authors give consent for the publication of the manuscript.

#### Competing interests

Rongxi Yang, Chong Shen, Chunlan Liu and Qiming Yin are inventors of a provisional patent application relating to the subject matter of this manuscript and therefore declare potential competing interests.

#### Author details

<sup>1</sup>Department of Epidemiology and Biostatistics, School of Public Health, Nanjing Medical University, 101 Longmian Avenue, Jiangning, Nanjing 211166, China. <sup>2</sup>Division of Clinical Epidemiology, Affiliated Geriatric Hospital of Nanjing Medical University, Nanjing 211166, China.

Received: 12 April 2022 Accepted: 22 December 2022  
Published online: 29 December 2022

### References

- Kim J, Thayabaranathan T, Donnan GA, Howard G, Howard VJ, Rothwell PM, Feigin V, Norrving B, Owolabi M, Pandian J, Liu L, Cadilhac DA, Thrift AG. Global stroke statistics 2019. *Int J Stroke*. 2020;15(8):819–38.
- Thayabaranathan T, Kim J, Cadilhac D, Thrift AG, Donnan G, Howard G, Howard VJ, Rothwell PM, Feigin VL, Norrving B, Owolabi M, Pandian JD, Liu L, Olaiya MT. Global stroke statistics 2022. *Int J Stroke*. 2022. <https://doi.org/10.1177/17474930221123175>.
- Ma Q, Li R, Wang L, Yin P, Wang Y, Yan C, Ren Y, Qian Z, Vaughn MG, McMillin SE, Hay SI, Naghavi M, Cai M, Wang C, Zhang Z, Zhou M, Lin H, Yang Y. Temporal trend and attributable risk factors of stroke burden in China, 1990–2019: an analysis for the Global Burden of Disease Study 2019. *Lancet Public Health*. 2021;6(12):e897–906.
- Benjamin EJ, Blaha MJ, Chiuve SE, Cushman M, Das SR, Deo R, de Ferranti SD, Floyd J, Fornage M, Gillespie C, Isasi CR, Jimenez MC, Jordan LC, Judd SE, Lackland D, Lichtman JH, Lisabeth L, Liu S, Longenecker CT, Mackey RH, Matsushita K, Mozaffarian D, Mussolino ME, Nasir K, Neumar RW, Palaniappan L, Pandey DK, Thiagarajan RR, Reeves MJ, Ritchey M, Rodriguez CJ, Roth GA, Rosamond WD, Sasson C, Towfighi A, Tsao CW, Turner MB, Virani SS, Voeks JH, Willey JZ, Wilkins JT, Wu JH, Alger HM, Wong SS, Muntner P, American Heart Association Statistics Committee, Stroke Statistics Subcommittee. Heart disease and stroke statistics-2017 update: a report from the American Heart Association. *Circulation*. 2017;135(10):e146–603.
- Bevan S, Traylor M, Adib-Samii P, Malik R, Paul NL, Jackson C, Farrall M, Rothwell PM, Sudlow C, Dichgans M, Markus HS. Genetic heritability of ischemic stroke and the contribution of previously reported candidate gene and genomewide associations. *Stroke*. 2012;43(12):3161–7.
- Malik R, Dichgans M. Challenges and opportunities in stroke genetics. *Cardiovasc Res*. 2018;114(9):1226–40.
- Felling RJ, Song H. Epigenetic mechanisms of neuroplasticity and the implications for stroke recovery. *Exp Neurol*. 2015;268:37–45.
- Zhou Z, Sun B, Li X, Zhu C. DNA methylation landscapes in the pathogenesis of type 2 diabetes mellitus. *Nutr Metab (Lond)*. 2018;15:47.
- Zaina S, Heyn H, Carmona FJ, Varol N, Sayols S, Condom E, Ramirez-Ruz J, Gomez A, Goncalves I, Moran S, Esteller M. DNA methylation map of human atherosclerosis. *Circ Cardiovasc Genet*. 2014;7(5):692–700.
- Liu G, Bin P, Wang T, Ren W, Zhong J, Liang J, Hu CA, Zeng Z, Yin Y. DNA methylation and the potential role of methyl-containing nutrients in cardiovascular diseases. *Oxid Med Cell Longev*. 2017;2017:1670815.
- Nicoglou A, Merlin F. Epigenetics: a way to bridge the gap between biological fields. *Stud Hist Philos Biol Biomed Sci*. 2017;66:73–82.
- Bird A. Perceptions of epigenetics. *Nature*. 2007;447(7143):396–8.
- Yang J, Zhou JS, Zhao YX, Yang ZH, Zhao HD, Zhang YD, Zou JJ. ABCB1 hypomethylation is associated with decreased antiplatelet effects of clopidogrel in Chinese ischemic stroke patients. *Pharmazie*. 2015;70(2):97–102.
- Wang C, Xu G, Wen Q, Peng X, Chen H, Zhang J, Xu S, Zhang C, Zhang M, Ma J, Hui Z, Wu G, Ma M. CBS promoter hypermethylation increases the risk of hypertension and stroke. *Clinics (Sao Paulo)*. 2019;74:e630.
- Zhou S, Zhang Y, Wang L, Zhang Z, Cai B, Liu K, Zhang H, Dai M, Sun L, Xu X, Cai H, Liu X, Lu G, Xu G. CDKN2B methylation is associated with carotid artery calcification in ischemic stroke patients. *J Transl Med*. 2016;14(1):333.
- Liu C, Yin Q, Li M, Fan Y, Shen C, Yang R. ACTB methylation in blood as a potential marker for the pre-clinical detection of stroke: a prospective nested case-control study. *Front Neurosci*. 2021;15:644943.
- Zurawa-Janicka D, Wentz T, Jarzab M, Skorko-Glonek J, Glaza P, Gieldon A, Ciarkowski J, Lipinska B. Structural insights into the activation mechanisms of human HtrA serine proteases. *Arch Biochem Biophys*. 2017;621:6–23.
- Wardlaw JM, Smith C, Dichgans M. Mechanisms of sporadic cerebral small vessel disease: insights from neuroimaging. *Lancet Neurol*. 2013;12(5):483–97.
- Verdura E, Herve D, Scharrer E, Amador Mdel M, Guyant-Marechal L, Philippi A, Corlobe A, Bergametti F, Gazal S, Prieto-Morin C, Beaufort N, Le Bail B, Viakhireva I, Dichgans M, Chabriat H, Haffner C, Tournier-Lasserre E.

- Heterozygous HTRA1 mutations are associated with autosomal dominant cerebral small vessel disease. *Brain*. 2015;138(Pt 8):2347–58.
20. Nozaki H, Kato T, Nihonmatsu M, Saito Y, Mizuta I, Noda T, Koike R, Miyazaki K, Kaito M, Ito S, Makino M, Koyama A, Shiga A, Uemura M, Sekine Y, Murakami A, Moritani S, Hara K, Yokoseki A, Kuwano R, Endo N, Momotsu T, Yoshida M, Nishizawa M, Mizuno T, Onodera O. Distinct molecular mechanisms of HTRA1 mutants in manifesting heterozygotes with CARASIL. *Neurology*. 2016;86(21):1964–74.
  21. Lee YC, Chung CP, Chao NC, Fuh JL, Chang FC, Soong BW, Liao YC. Characterization of heterozygous HTRA1 mutations in Taiwanese patients with cerebral small vessel disease. *Stroke*. 2018;49(7):1593–601.
  22. Rannikmae K, Sivakumaran V, Millar H, Malik R, Anderson CD, Chong M, Dave T, Falcone GJ, Fernandez-Cadenas I, Jimenez-Conde J, Lindgren A, Montaner J, O'Donnell M, Pare G, Radmanesh F, Rost NS, Slowik A, Soderholm M, Traylor M, Pulit SL, Seshadri S, Worrall BB, Woo D, Markus HS, Mitchell BD, Dichgans M, Rosand J, Sudlow CLM, Stroke Genetics Network (SiGN), METASTROKE Collaboration, International Stroke Genetics Consortium (ISGC). COL4A2 is associated with lacunar ischemic stroke and deep ICH: meta-analyses among 21,500 cases and 40,600 controls. *Neurology*. 2017;89(17):1829–39.
  23. Mishra A, Chauhan G, Violleau MH, Vojinovic D, Jian X, Bis JC, Li S, Saba Y, Grenier-Boley B, Yang Q, Bartz TM, Hofer E, Soumare A, Peng F, Duperron MG, Foglio M, Mosley TH, Schmidt R, Psaty BM, Launer LJ, Boerwinkle E, Zhu Y, Mazoyer B, Lathrop M, Bellenguez C, Van Duijn CM, Ikram MA, Schmidt H, Longstreth WT, Fornage M, Seshadri S, Joutel A, Tzourio C, Debette S. Association of variants in HTRA1 and NOTCH3 with MRI-defined extremes of cerebral small vessel disease in older subjects. *Brain*. 2019;142(4):1009–23.
  24. Tian Y, Tang W, Yang S, Zhao Y, Chen Y, Zhao X, Liu C, Chen X, Shen C. HTRA1 variants and the interaction with smoking confer the genetic susceptibility to ischemic stroke. *Int J Med Sci*. 2021;18(8):1840–7.
  25. Horvath S, Zhang Y, Langfelder P, Kahn RS, Boks MP, van Eijk K, van den Berg LH, Ophoff RA. Aging effects on DNA methylation modules in human brain and blood tissue. *Genome Biol*. 2012;13(10):R97.
  26. Tennstaedt A, Popsel S, Truebestein L, Hauske P, Brockmann A, Schmidt N, Irle I, Sacca B, Niemeyer CM, Brandt R, Ksiezak-Reding H, Tirniceriu AL, Egensperger R, Baldi A, Dehmelt L, Kaiser M, Huber R, Clausen T, Ehrmann M. Human high temperature requirement serine protease A1 (HTRA1) degrades tau protein aggregates. *J Biol Chem*. 2012;287(25):20931–41.
  27. Yang Z, Camp NJ, Sun H, Tong Z, Gibbs D, Cameron DJ, Chen H, Zhao Y, Pearson E, Li X, Chien J, Dewan A, Harmon J, Bernstein PS, Shridhar V, Zabriskie NA, Hoh J, Howes K, Zhang K. A variant of the HTRA1 gene increases susceptibility to age-related macular degeneration. *Science*. 2006;314(5801):992–3.
  28. Tiaden AN, Richards PJ. The emerging roles of HTRA1 in musculoskeletal disease. *Am J Pathol*. 2013;182(5):1482–8.
  29. Hara K, Shiga A, Fukutake T, Nozaki H, Miyashita A, Yokoseki A, Kawata H, Koyama A, Arima K, Takahashi T, Ikeda M, Shiota H, Tamura M, Shimoe Y, Hirayama M, Arisato T, Yanagawa S, Tanaka A, Nakano I, Ikeda S, Yoshida Y, Yamamoto T, Ikeuchi T, Kuwano R, Nishizawa M, Tsuji S, Onodera O. Association of HTRA1 mutations and familial ischemic cerebral small-vessel disease. *N Engl J Med*. 2009;360(17):1729–39.
  30. Jung SE, Shin KJ, Lee HY. DNA methylation-based age prediction from various tissues and body fluids. *BMB Rep*. 2017;50(11):546–53.
  31. Chen X, Zhao C, Zhao Z, Wang H, Fang Z. Specific glioma prognostic subtype distinctions based on DNA methylation patterns. *Front Genet*. 2019;10:786.
  32. Chao TF, Lip GY, Liu CJ, Tuan TC, Chen SJ, Wang KL, Lin YJ, Chang SL, Lo LW, Hu YF, Chen TJ, Chiang CE, Chen SA. Validation of a modified CHA2DS2-VASc score for stroke risk stratification in Asian patients with atrial fibrillation: a nationwide cohort study. *Stroke*. 2016;47(10):2462–9.
  33. Liu C, Marioni RE, Hedman AK, Pfeiffer L, Tsai PC, Reynolds LM, Just AC, Duan Q, Boer CG, Tanaka T, Elks CE, Aslibekyan S, Brody JA, Kuhnel B, Herder C, Almlil LM, Zhi D, Wang Y, Huan T, Yao C, Mendelson MM, Joehanes R, Liang L, Love SA, Guan W, Shah S, McRae AF, Kretschmer A, Prokisch H, Strauch K, Peters A, Visscher PM, Wray NR, Guo X, Wiggins KL, Smith AK, Binder EB, Ressler KJ, Irvin MR, Absher DM, Hernandez D, Ferrucci L, Bandinelli S, Lohman K, Ding J, Trevisi L, Gustafsson S, Sandling JH, Stolk L, Uitterlinden AG, Yet I, Castillo-Fernandez JE, Spector TD, Schwartz JD, Vokonas P, Lind L, Li Y, Fornage M, Arnett DK, Wareham NJ, Sotoodehnia N, Ong KK, van Meurs JBJ, Conneely KN, Baccarelli AA, Deary IJ, Bell JT, North KE, Liu Y, Waldenberger M, London SJ, Ingelsson E, Levy D. A DNA methylation biomarker of alcohol consumption. *Mol Psychiatry*. 2018;23(2):422–33.
  34. Han L, Liu Y, Duan S, Perry B, Li W, He Y. DNA methylation and hypertension: emerging evidence and challenges. *Brief Funct Genom*. 2016;15(6):460–9.
  35. Ducas RA, Labos C, Allen D, Golian M, Jeyaraman M, Lys J, Mann A, Copstein L, Vokey S, Rabbani R, Zarychanski R, Abou-Setta AM, Menkis AH. Association of pre-hospital ECG administration with clinical outcomes in ST-segment myocardial infarction: a systematic review and meta-analysis. *Can J Cardiol*. 2016;32(12):1531–41.
  36. Saver JL. Time is brain-quantified. *Stroke*. 2006;37(1):263–6.
  37. Yin Q, Yang X, Li L, Xu T, Zhou W, Gu W, Ma F, Yang R. The association between breast cancer and blood-based methylation of S100P and HYAL2 in the Chinese population. *Front Genet*. 2020;11:977.
  38. Yang R, Pfutze K, Zucknick M, Sutter C, Wappenschmidt B, Marme F, Qu B, Cuk K, Engel C, Schott S, Schneeweiss A, Brenner H, Claus R, Plass C, Bugert P, Hoth M, Sohn C, Schmutzler R, Bartram CR, Burwinkel B. DNA methylation array analyses identified breast cancer-associated HYAL2 methylation in peripheral blood. *Int J Cancer*. 2015;136(8):1845–55.
  39. Yang R, Stocker S, Schott S, Heil J, Marme F, Cuk K, Chen B, Golatta M, Zhou Y, Sutter C, Wappenschmidt B, Schmutzler R, Bugert P, Qu B, Bartram CR, Sohn C, Schneeweiss A, Burwinkel B. The association between breast cancer and S100P methylation in peripheral blood by multicenter case-control studies. *Carcinogenesis*. 2017;38(3):312–20.
  40. Doi A, Park IH, Wen B, Murakami P, Aryee MJ, Irizarry R, Herb B, Ladd-Acosta C, Rho J, Loewer S, Miller J, Schlaeger T, Daley GQ, Feinberg AP. Differential methylation of tissue- and cancer-specific CpG island shores distinguishes human induced pluripotent stem cells, embryonic stem cells and fibroblasts. *Nat Genet*. 2009;41(12):1350–3.
  41. Irizarry RA, Ladd-Acosta C, Wen B, Wu Z, Montano C, Onyango P, Cui H, Gabo K, Rongione M, Webster M, Ji H, Potash J, Sabuncyan S, Feinberg AP. The human colon cancer methylome shows similar hypo- and hyper-methylation at conserved tissue-specific CpG island shores. *Nat Genet*. 2009;41(2):178–86.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

