Meta Gene 2 (2014) 514–524

Contents lists available at ScienceDirect

Meta Gene

Association of the histone-lysine N-methyltransferase MLL5 gene with coronary artery disease in Chinese Han people

Qinghua Yuan [a,b], Xiang Xie [a,b], Zhenyan Fu [a,b], Xiang Ma [a,b], Yining Yang [a,b], Ding Huang [a], Fen Liu [b], Chuanfang Dai [a,b], Yitong Ma [a,b,*]

[a] Department of Cardiology, First Affiliated Hospital of Xinjiang Medical University, Urumqi 830054, PR China
[b] Xinjiang Key Laboratory of Cardiovascular Disease Research, Urumqi 830054, PR China

ARTICLE INFO

Article history:
Received 20 March 2014
Revised 7 June 2014
Accepted 9 June 2014
Available online 31 July 2014

Keywords:
MLL5
Single nucleotide polymorphism
Coronary artery disease
Case–control study

ABSTRACT

Background: MLL5, a member of the histone-lysine N-methyltransferase family, has been implicated in the control of the cell cycle progression and survival. The aim of this study was to explore the relationship between the interaction of histone-lysine N-methyltransferase MLL5 gene polymorphism and CAD in a Chinese Han population.

Methods: Using a case–control study of Chinese CAD patients (n = 565) and healthy controls (n = 694), we investigated the MLL5 gene polymorphism by the use of polymerase chain reaction fragment length polymorphism (PCR–RFLP) analysis.

Results: For total, the distribution of SNP1 (rs12671368) and SNP2 (rs2192932) genotypes showed a significant difference between CAD and control participants ($P_1 = 0.03$, $P_2 = 0.02$). For total the distribution of SNP1 (rs12671368) and SNP2 (rs2192932) alleles in the dominant model (GG vs. AA + AG) and the recessive model (AA vs. AG + GG) showed a significant difference between CAD and control participants (for allele: $P_1 < 0.01$ and $P_2 = 0.05$, for dominant model: $P_1 > 0.05$ and $P_2 = 0.02$, for recessive model: $P_1 = 0.03$ and $P_2 = 0.78$, respectively). For total the significant difference of the distribution of SNP1 and SNP2 in the dominant model and recessive model was retained after adjusting for covariates (for dominant model: SNP1 OR: 1.68, 95% confidence interval [CI]: 1.08–2.64, P = 0.02; SNP2

* Corresponding author at: Department of Cardiology, First Affiliated Hospital of Xinjiang Medical University, Urumqi 830011, PR China. Tel.: +86 991 4366169; fax: +86 991 4364303, +86 991 4366169.
E-mail address: myt-xj@163.com (Y. Ma).

http://dx.doi.org/10.1016/j.mgene.2014.06.001

OR: 0.51, 95% CI: 0.36–0.72, P = 0.01; for recessive model: SNP1 OR: 1.84, 95% confidence interval [CI]: 1.28–2.64, P < 0.01; SNP2 OR: 0.65, 95% CI: 0.35–1.22, P = 0.18).

Conclusions: The GG genotype of rs12671368 and the AA genotype of rs2192932 in the MLL5 gene could be protective genetic markers of CAD.

Introduction

Coronary artery disease (CAD) is one of the major causes of mortality; the World Health Organization reports that currently 17.1 million deaths a year are attributed to these diseases, and the number is expected to rise (Cartwright et al., 2011). Many epidemiological studies show that smoking, obesity, diabetes, hypertension and hyperlipidemia are risk factors for CAD. However, in some individuals, coronary artery disease is not associated with any of these risk factors, suggesting that other genetic factors contribute to a predisposition to coronary atherosclerosis and its thrombotic complications (Qi et al., 2013). Many studies show that gene epigenetics changes have a close relationship with atherosclerosis such as that induction of elevated homocysteine levels can change human vascular smooth muscle cells' DNA methylation changes leading to the occurrence of atherosclerosis (Yideng et al., 2007). Histone-lysine N-methyltransferase MLL5 also called mixed lineage leukemia 5 (MLL5) is also known as lysine methyltransferases 2E (KMT2E). MLL5 belongs to the MLL family of methyltransferases that regulate gene expression during developmental processes. These enzymes catalyze the addition of methyl groups to the ε-amino moiety of lysine and are highly specific for lysine 4 of histone H3. The MLL family member MLL3 plays an important role in generating fat. MLL5 is a putative tumor suppressor encoding a SET and PHD domain protein homologous to *Drosophila* trithorax (Cheng et al., 2008; Deng et al., 2004; Emerling et al., 2002). and yeast SET3 (Kittler et al., 2007). Recent studies establish a role for MLL5 in hematopoietic stem cells (Heuser et al., 2009; Madan et al., 2009; Zhang et al., 2009), and three independent studies reported the genetic analysis of MLL5 deficiency in mice that suffer from mild growth retardation but do not develop spontaneous leukemia (Heuser et al., 2009; Madan et al., 2009; Zhang et al., 2009). A recent (Bjorkegren et al., 2014) study stated that MLL5 was the only key master regulator that was specific for the mature atherosclerosis regression network according to the hypergeometric test, as well as the specific master regulator of partial regression in mature lesions.

Previous studies have shown that overexpression or knockdown of MLL5 impeded cell-cycle progression and proposed that MLL5 may participate in the cell-cycle regulatory network at multiple stages of the cell cycle (Cheng et al., 2008; Deng et al., 2004). A study showed that MLL5 is a substrate of Cdc2 kinase, and phosphorylation of MLL5 is required for mitosis progression (Liu et al., 2010).

To the best of our knowledge, there has been no previous study on the association between human MLL5 and CAD. Therefore, our aim was to investigate this via a haplotype-based case–control study that used SNPs in conjunction with separate analyses of data with regard to sex.

Methods

Subjects diagnosed with CAD were recruited at the First Affiliated Hospital of Xinjiang Medical University, from 2006 to 2012. We enrolled 565 coronary artery disease (CAD) patients. All subjects who agreed to participate in the study were evaluated by way of a detailed questionnaire that provided information about coronary risk factors such as smoking habit, the presence of diabetes mellitus or hypertension. CAD was diagnosed by angiography, which was defined as the presence of at least one significant coronary artery stenosis of ≥50% luminal diameter on coronary angiography. The study also enrolled 694 Han people as controls. The subjects were from a Han population who lived in the Xinjiang Uygur Autonomous Region of China. These individuals did not have: a history of CAD; electrocardiographic signs of CAD. Demographic data about the presence of traditional coronary risk factors, including

hypertension, diabetes mellitus, smoking habit, alcohol consumption and serum cholesterol, were collected from all study participants. We used SPSS 19.0.

Hypertension was defined as having a systolic blood pressure above 140 mm Hg or/and diastolic blood pressure above 90 mm Hg or using any anti-hypertensive agent.

Dyslipidemia was diagnosed according to the current guidelines from the National Cholesterol Education Program (NCEP) Adult Treatment Panel (ATP) III: any abnormal status of TG, HDL-C and LDL-C (TG $\geq$ 1.70 mmol/L, HDL-C < 0.91 mmol/L, LDL-C $\geq$ 3.46 mmol/L). Hypercholesterolemia was defined as a documented total cholesterol value $\geq$200 mg/dL ($\geq$5.2 mmol/L) (Kesteloot et al., 1989).

In addition, individuals with fasting plasma glucose >7.0 mmol/L or with a history of diabetes or treatment with insulin were considered diabetic. Smoking was classified as smokers (including current and ex-smokers) or non-smokers. All patients with impaired renal function, malignancy, connective tissue disease, valvular disease or chronic inflammatory disease were excluded.

Ethical approval of the study protocol

Written informed consent was obtained from all participants. All participants explicitly provided permission for DNA analyses as well as collection of relevant clinical data. This study was approved by the Ethics Committee of the First Affiliated Hospital of Xinjiang Medical University (Urumqi, China). It was conducted according to the standards of the Declaration of Helsinki.

Genotyping

The human *MLL5* gene is located on chromosome 7q22.1, spans approximately 6.7 kbp, and contains 26 exons. There are 2157 SNPs listed in the National Center for Biotechnology Information SNP database Build 129 (http://www.ncbi.nlm.nih.gov/SNP). We screened the data for tag SNPs on the International HapMap Project website (http://www.hapmap.org/index.html.ja) using a cutoff level of $r^2 \geq 0.8$. For the minor allele frequencies (MAF), we used a cutoff level of $\geq$0.2. SNPs with relatively high minor allele frequencies have been shown to be very useful as genetic markers for genetic case–control studies. In accordance with these criteria, there are 4 tag SNPs and we selected SNP1 (rs12671368) and SNP2 (rs2192932) for this gene (Fig. 1). SNP2 was located in the promoter region, whereas SNP1 was located in the coding regions of MLL5-1 intron 5 and in the regions of MLL5-2 intron 6. Blood samples were collected from all participants, and genomic DNA was extracted from the peripheral blood leukocytes by phenol and chloroform extraction. Genotyping was performed using polymerase chain reaction fragment length polymorphism. Extraction using the whole blood genomic DNA extraction kit (Day Biochemical Co., LTD.) shall be carried out in accordance with the manual steps. Primers using Primer premier 5.0 software design the SNP1 sense primer: 5′AGAAGGTGTGATTGGGA3′; antisense primer: 5′TAATACTTGGCTTTA GGG3′ and the SNP2 sense primer: 5′GGAGAAAGGCAGATAAGAAAATAGGG3′; antisense primer: 5′ AGGTAGAGAGGGAGAAGTGACGATG3′, Genotyping for SNP1 and SNP2 in the present case–control study was done by PCR amplification of 340 bp and 540 bp, followed by restriction digestion with bst4cI and Eco130I (Ferments, Beijing, China).

Characteristics of study participants

The study cohort consists of 1259 subjects (709 men; 550 women). The clinical and metabolic characteristics of the study population are shown separately for men and women in Table 1.

MLL5 genotype and allele frequencies

All genotyped SNPs were in the Hardy–Weinberg equilibrium and commonly with MAF ($P \geq 0.2$). Table 2 shows detailed information for each SNP as well as the allele frequencies.

Genotyping of MLL5 SNPs

There are 2157 SNPs for the human MLL5 gene listed in the National Center for Biotechnology Information SNP database (http://www.ncbi.nlm.nih.gov/SNP). It has been observed that adjacent SNPs are often highly correlated. To reduce genotyping cost, many algorithms have been developed to select the smallest set of SNPs such that all the other SNPs can be inferred from them. We selected the tag SNPs of the

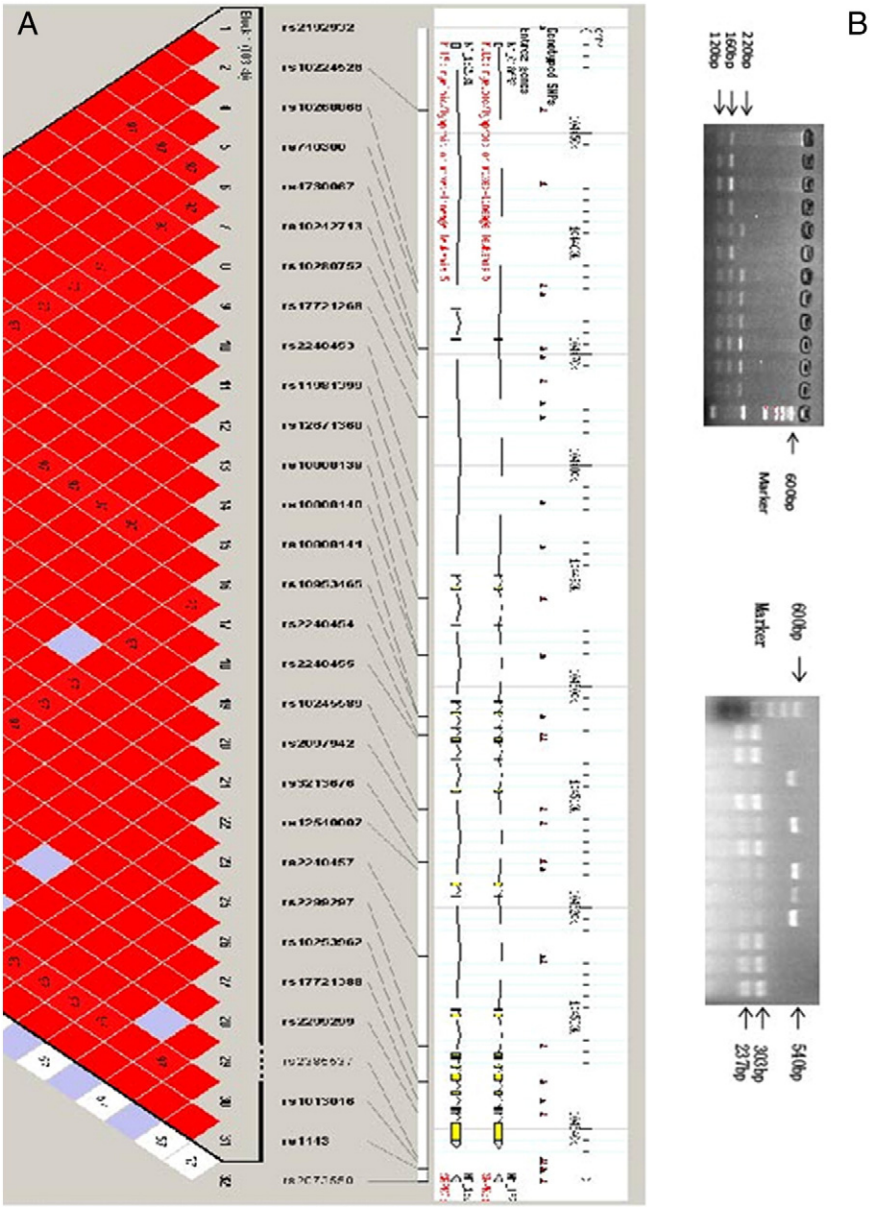

Fig. 1. A: tag SNPs' location in the MLL5 gene; B: (top) SNP1 (rs12671368) PCR products by bst4c1 enzyme digestion; (bottom) SNP2 (rs2192932) PCR products by Eco1301 enzyme digestion.

Table 1
Characteristics of study participants.

	Total			Men			Women		
	CAD patients	Control subjects	P value	CAD patients	Control subjects	P value	CAD patients	Control subjects	P value
Number (n)	565	694	0.77	415	294	0.35	150	400	0.56
Age, mean (SD)	59.93 ± 9.90	49.46 ± 9.32	0.00*	58.80 ± 10.13	48.71 ± 8.86	0.00*	63.04 ± 8.53	50.01 ± 9.62	<0.01*
BMI, mean (SD)	25.75 ± 3.45	24.58 ± 3.44	0.01*	26.06 ± 3.47	24.81 ± 3.12	0.15	24.87 ± 3.23	24.41 ± 3.65	0.29
BUN, mean (SD)	6.18 ± 3.17	4.91 ± 1.41	0.02*	6.38 ± 3.38	5.19 ± 1.42	0.81	5.61 ± 2.39	4.70 ± 1.36	0.08
Cr, mean (SD)	85.66 ± 60.73	76.28 ± 20.64	0.23	87.34 ± 65.15	88.39 ± 20.12	0.11	80.90 ± 45.80	67.36 ± 15.96	0.53
UA, mean (SD)	283.31 ± 134.06	293.70 ± 77.42	0.16	280.93 ± 137.60	331.02 ± 76.36	0.36	290.07 ± 123.70	266.36 ± 65.99	0.30
Glu, mean (SD)	6.46 ± 2.72	5.06 ± 0.67	<0.01*	6.51 ± 2.72	5.22 ± 0.73	<0.01*	6.33 ± 2.74	4.95 ± 0.66	<0.01*
TG, mean (SD)	5.31 ± 8.69	1.47 ± 1.08	<0.01*	5.80 ± 9.09	1.70 ± 1.32	0.35	3.90 ± 7.22	1.31 ± 0.82	<0.01*
TC, mean (SD)	5.14 ± 2.82	4.61 ± 0.10	<0.01*	5.26 ± 2.95	4.60 ± 0.99	0.06	4.79 ± 2.39	4.61 ± 1.00	0.12
HDL, mean (SD)	14.26 ± 35.96	1.29 ± 0.47	<0.01*	16.10 ± 37.71	1.27 ± 0.45	0.00*	7.68 ± 25.81	1.30 ± 0.49	<0.01*
LDL, mean (SD)	2.30 ± 1.69	2.91 ± 0.95	<0.01*	2.20 ± 0.99	2.86 ± 0.88	0.11	2.33 ± 1.04	2.94 ± 0.99	0.20
Hypertension (%)	231 (40.9%)	278 (40.1%)	0.77	173 (41.7%)	112 (38.1%)	0.35	58 (38.7%)	166 (41.5%)	0.56
DM (%)	103 (18.3%)	1 (0.1%)	<0.01*	76 (18.4%)	1 (0.3%)	0.00*	27 (18.1%)	0 (0%)	<0.01*
Smoke (%)	200 (35.4%)	204 (29.4%)	0.03*	200 (48.2%)	201 (68.4%)	0.00*	0 (0%)	3 (0.8%)	0.57
Alcohol consumption (%)	137 (24.2%)	130 (18.7%)	0.02*	136 (32.8%)	120 (40.8%)	0.03*	1 (0.7%)	10 (2.5%)	0.30

BMI, body mass index; BUN, blood urea nitrogen; Cr, creatinine; Glu, glucose; TG, triglyceride; TC, total cholesterol; HDL, high density lipoprotein; LDL, low density lipoprotein; DM, diabetes mellitus. Continuous variables were expressed as mean ± standard deviation. P value of continuous variables was calculated by independent T-test.
The P value of categorical variable was calculated by Fisher's exact test.
* $P < 0.05$.

Table 2
Genotype and allele distributions in patients with CAD and control subjects.

				Total			Men			Women		
				CAD n (%)	Control n (%)	P value	CAD n (%)	Control n (%)	P value	CAD n (%)	Control n (%)	P value
rs12671368 (SNP1)	Genotype	Dominant model	G/G	88 (15.6)	138 (19.9)	0.04*	67 (16.1)	58 (19.7)	0.34	21 (14.0)	80 (20.0)	0.04*
			A/A	213 (37.7)	220 (31.7)		149 (35.9)	93 (31.6)		64 (42.7)	127 (31.8)	
			G/A	264 (46.7)	336 (48.4)		199 (48.0)	143 (48.6)		65 (43.3)	193 (48.3)	
			GG	88 (15.6)	138 (19.9)	0.06	67 (16.1)	58 (19.7)	0.23	21 (14.0)	80 (20.0)	0.11
			AG + AA	477 (84.4)	556 (81.1)		348 (83.9)	236 (80.3)		129 (87.0)	320 (80.1)	
		Recessive model	AA	213 (37.7)	220 (31.7)	0.03*	149 (35.9)	93 (31.6)	0.26	64 (42.7)	127 (31.8)	0.02*
			AG + GG	352 (62.3)	474 (68.3)		266 (64.1)	201 (68.4)		86 (57.3)	273 (68.3)	
		Additive model	GA	264 (46.7)	336 (48.4)	0.57	199 (48.0)	143 (48.6)	0.88	65 (43.3)	193 (48.3)	0.34
			GG + AA	301 (53.2)	358 (51.6)		216 (52.0)	151 (51.3)		85 (56.7)	207 (51.8)	
	Allele		G	440 (38.9)	612 (44.1)	0.01*	333 (40.1)	259 (44.0)	0.14	107 (35.7)	353 (44.1)	0.01*
			A	690 (61.1)	776 (55.9)		497 (59.9)	329 (56.0)		193 (64.3)	447 (55.9)	
rs2192932 (SNP2)	Genotype	Dominant model	G/G	324 (57.3)	348 (50.1)	0.02*	234 (56.4)	151 (51.4)	0.41	90 (60.0)	197 (49.3)	0.01*
			A/A	37 (6.5)	65 (9.4)		32 (7.7)	24 (8.2)		5 (3.3)	41 (10.3)	
			G/A	204 (36.1)	281 (40.5)		149 (35.9)	119 (40.5)		55 (36.7)	162 (40.5)	
			GG	323 (57.3)	348 (50.1)	0.01*	233 (56.3)	151 (51.4)	0.19	90 (60.0)	197 (49.3)	0.03*
			AG + AA	241 (42.7)	346 (49.9)		181 (43.7)	143 (48.6)		60 (40)	203 (50.7)	
		Recessive model	AA	37 (6.5)	65 (9.4)	0.08	32 (7.7)	24 (8.2)	0.89	5 (3.3)	41 (10.3)	0.01*
			AG + GG	528 (93.5)	629 (90.6)		383 (92.3)	270 (91.8)		145 (96.7)	359 (89.8)	
		Additive model	GA	204 (36.1)	281 (40.5)	0.12	149 (35.9)	119 (40.5)	0.24	55 (36.7)	162 (40.5)	0.01*
			GG + AA	361 (63.9)	413 (59.5)		266 (64.1)	175 (59.5)		95 (63.3)	238 (59.5)	
	Allele		G	852 (75.4)	977 (70.4)	0.01*	617 (74.3)	421 (71.6)	0.27	235 (78.3)	556 (69.5)	<0.01*
			A	278 (24.6)	411 (29.6)		213 (25.7)	167 (28.4)		65 (21.7)	244 (30.5)	

CAD, coronary artery disease.
The P value of genotype was calculated by Fisher's exact test.
* $P < 0.05$.

Table 3
Multiple logistic regression analysis for CAD patients and control subjects of a Han population.

		Total			Men			Women		
		OR	95% CI	P	OR	95% CI	P	OR	95% CI	P
a										
rs12671368	Dominant model (GG vs. GA + AA)	1.69	1.08–2.64	0.02*	1.64	0.92–2.93	0.09	1.95	0.90–4.19	0.09
	Hypertension	0.96	0.68–1.36	0.82	1.11	0.71–1.73	0.54	0.81	0.44–1.48	0.50
	DM	343.28	32.77–3595.81	<0.01*	96.03	11.96–770.87	<0.01*	NA	NA	0.10
rs12671368	Recessive model (AA vs. GA + GG)	1.84	1.28–2.64	<0.01*	1.58	0.99–2.52	0.06	2.77	1.45–5.19	<0.01*
	Hypertension	0.90	0.63–1.28	0.56	1.06	0.68–1.66	0.64	0.72	0.39–1.32	0.29
	DM	332.85	33.82–3275.94	<0.01*	99.46	12.30–804.48	<0.01*	NA	NA	0.10
b										
rs2192932	Dominant model (GG vs. GA + AA)	0.51	0.36–0.72	0.01*	0.60	0.39–0.95	0.02*	0.31	0.16–0.57	<0.01*
	Hypertension	0.89	0.63–1.27	0.89	1.08	0.69–1.70	0.73	0.66	0.35–1.22	0.18
	DM	352.60	36.36–3419.33	<0.01*	103.19	12.66–841.04	<0.01*	NA	NA	0.10*
rs2192932	Recessive model (AA vs. GA + GG)	0.65	0.35–1.22	0.18	1.07	0.48–2.38	0.87	0.28	0.08–0.92	0.04*
	Hypertension	0.94	0.66–1.33	0.73	1.08	0.69–1.69	0.52	0.79	0.43–1.44	0.44
	DM	333.09	46.279–5569.708	<0.01*	93.02	11.68–741.04	<0.01*	NA	NA	0.10

Adjusting for age; BMI (body mass index); sex; hypertension; smoking habit; TG (triglyceride); cholesterol; HDL-C (high-density lipoprotein cholesterol); LDL-C (low density lipoprotein); DM (diabetes mellitus); alcohol consumption.
The P value of genotype was calculated by Fisher's exact test.
* $P < 0.05$.

MLL5 gene on the International HapMap Project website using phase I & II database (http://www.hapmap.org/). We obtained two tag SNPs (rs12671368 and rs2192932 for the MLL5 gene) spanning the coding region of 5.6 kb in MLL5-1 and 5.6 kb in MLL5-2 for Chinese Han subjects using MAF ≥ 0.2 and linkage disequilibrium patterns with $r^2 \geq 0.8$ as a cutoff. Because rs12671368 and rs2192932 are synonymous mutations, therefore, in the present study, two SNPs of the MLL5 gene were selected for genotype detection. Genomic DNA was extracted from peripheral blood leukocytes using a DNA extraction kit (Beijing Biotech Company Limited, Beijing, China). Genotyping was confirmed by polymerase chain reaction (PCR) restriction fragment length polymorphism (RFLP) analysis. The primers of these two SNPs were designed by using Primer Premier 5.0. Their synthesis was undertaken by Shanghai Genery Biological Technology Company Limited (Shanghai, China). The primer pair sequences, annealing temperatures, and restriction enzymes for the two SNPs have been described previously. Digestion of restriction enzymes was according to the manufacturer's instructions. To ensure that the results are verified, we used sequenced genomic DNAs as positive controls in our assays.

Statistical analyses

All continuous variables are expressed as the mean ± SD. Differences in continuous variables between CAD patients and control subjects were analyzed using the Mann–Whitney U test. Differences in categorical variables were analyzed using Fisher's exact test. The Hardy–Weinberg equilibrium was assessed by χ^2 analyses. Differences in the distributions of genotypes and alleles between CAD patients and control subjects were analyzed using Fisher's exact test. Based on the genotype data of the genetic variations, linkage disequilibrium (LD) analyses and a haplotype-based case–control study were carried out using the expectation maximization (EM) algorithm with the SNPAlyze software program ver 3.2 (Dynacom, Yokohama, Japan). $|D'|$ values >0.5 were used to assign SNP locations to one haplotype block. Tagged SNPs showing $r^2 < 0.5$ (which meant availability for the haplotype) were selected. In this haplotype-based case–control study, haplotypes with a frequency <0.01 were excluded. The frequency distribution of occurrence of the haplotypes was calculated by χ^2 analyses. In addition, logistic regression analyses were undertaken to assess the contribution of essential hypertension, DM, BMI, smoking habit and other factors entered into the model. $P < 0.05$ was considered significant. Statistical analyses were carried out using SPSS ver 12 (SPSS, Chicago, IL, USA).

Result

Characteristics of study participants

The clinical characteristics of the study participants from a Han population are shown in Table 1, and there was a significant difference in age, BMI and LDL-C between CAD patients and control subjects. In the Han population, for total, men, and women, the plasma concentration of Glu, HDL, and diabetes mellitus (DM) were significantly higher for patients with CAD than for control participants. For total and men, the plasma concentration of total cholesterol (TC) was significantly higher for patients with CAD than for control participants, but the factors smoking habit and alcohol consumption in total were significantly higher for patients with CAD, while for men these two factors were significantly higher for control participants. For total and women, the plasma concentration of triglyceride (TG), and blood urea nitrogen (BUN) were significantly higher for patients with CAD than for control participants.

Genotype and allele frequencies

The distribution of genotypes and alleles of SNP1 (rs12671368) and SNP2 (rs2192932) were shown in Table 2. For total, the distributions of SNP1 and SNP2 genotypes in the dominant model (GG vs. AG + AA) and recessive model (AA vs. AG + GG) show a significant difference between CAD and control participants ($P = 0.036$, $P_{1D} = 0.048$, $P_{1R} = 0.026$ and $P = 0.022$, $P_{2D} = 0.012$, $P_{2R} = 0.012$, respectively). For women the distribution of SNP1 and SNP2 genotypes in the dominant model, recessive model and additive model (AG vs. AA + GG) showed a significant difference between CAD and control

participants ($P = 0.040$, $P_{1D} = 0.110$, $P_{1R} = 0.021$, $P_{1A} = 0.338$; and $P = 0.011$, $P_{2D} = 0.028$, $P_{2R} = 0.009$, $P_{2A} = 0.009$).

For total and women, the distribution of SNP1 (rs12671368) G alleles is significantly lower for CAD patients compared to control participants (total: 61.1% vs. 55.9%; men: 59.9% vs. 56.0%; women: 64.3% vs. 55.9%).

For total and women, the distribution of a SNP2 (rs2192932) allele was significantly lower in CAD patients than in control participants ($P = 0.005$, $P = 0.004$, respectively) (total: 24.6% vs. 29.6%; men: 25.7% vs. 28.4%; women: 21.7% vs. 30.5%).

Multiple logistic regression analyses

After adjusting for age, sex, hypertension, diabetes, smoking habit, triglyceride, cholesterol, high-density lipoprotein cholesterol (HDL-C), low density lipoprotein (LDL-C) and the effects of other risk factors on coronary heart disease, multivariate logistic regression model analysis was performed. Table 3 showed the multiple logistic regression results, but BMI and other factors are not shown. For total and women, the significant difference of SNP1 (rs12671368) and SNP2 (rs2192932) was retained after adjusting for other cardiovascular risk factors, whether including the factor age or not. SNP1: for total participants — dominant model: OR: 1.69, 95% confidence interval [CI]: 1.08–2.64, $P_{1D} = 0.022$; recessive model: OR: 1.84, 95% confidence interval [CI]: 1.28–2.64, $P_{1R} = 0.001$ and for women — recessive model: OR: 2.77, 95% CI: 1.45–5.19, $P_{1R} = 0.001$. SNP2: for total — dominant model: OR: 0.51, 95% confidence interval [CI]: 0.36–0.72, $P_{2D} = 0.012$; recessive model: OR: 0.65, 95% confidence interval [CI]: 0.35–1.22, $P_{2R} = 0.182$, for women — dominant model: OR: 0.31, 95% CI: 0.16–0.57, $P_{2D} = 0.000$; recessive model: OR: 0.28, 95% CI: 0.08–0.92, $P_{2R} = 0.039$, and for men — dominant model: OR: 0.60, 95% CI: 0.39–0.95, $P_{2D} = 0.024$.

Two SNP haplotypes of MLL5

In the present study, we identified two novel polymorphisms, SNP1 and SNP2, and found that the GA* haplotype is associated with decreased risk for CAD in a Han population of Xinjiang, Western China. In the haplotype-based case–control analysis, by using SNP1 and SNP2, four haplotypes were established in the Han population. Distribution of the GA* haplotypes was significantly different between CAD patients and control subjects ($P < 0.001$). In addition, the frequency of the GA* haplotype was significantly lower for CAD patients than for control subjects ($P < 0.001$) and the GA* haplotype remained associated with decreased risk for CAD in the Han population (OR = 0.677, $P < 0.001$, 95% CI [0.558–0.822]).

Discussion

The observation of our study is that variation in the MLL5 gene is associated with CAD in Chinese Han subjects. The GG genotype of rs12671368 and/or AA genotype of rs2192932 was significantly lower in CAD patients compared with healthy subjects, and the GA* haplotype is associated with decreased risk for CAD in the Han population. To our knowledge, this is the first study to investigate a common allelic variant in the MLL5 gene and its association with CAD in Asian subjects.

Histone-lysine N-methyltransferase MLL5 also called mixed lineage leukemia 5 (MLL5) is located on human chromosome 7q22.1 (Emerling et al., 2002), a region exhibiting commonly recurring cytogenetic aberrations detected in myeloid malignancies. MLL5 belongs to the MLL family of methyltransferases that regulate gene expression during developmental processes. These enzymes catalyze the addition of methyl groups to the ε-amino moiety of lysine and are highly specific for lysine 4 of histone H3. MLL5 has no intrinsic histone lysine methyltransferase (HKMT) activity (Sebastian et al., 2009), however, once posttranslationally modified through glycosylation at Thr440 in the SET domain, the short isoform MLL5 is capable of generating mono- and dimethylated H3K4 marks (Fujiki et al., 2009). Importantly, a tri-methylated H3K4 mark at lysine 4 (H3K4me2/3) is thought to be the product of the methyltransferase activity of MLL5 (Zhou et al., 2013). Post translational modifications of histones are a key epigenetic mechanism used to regulate gene transcription, chromatin condensation, DNA damage sensing and repair. Knockout mice studies showed that murine MLL5 is required in normal hematopoiesis (Heuser et al.,

2009; Madan et al., 2009), as well as in maturation of spermatozoa (Yap et al., 2011). MLL5 is a putative tumor suppressor encoding a SET and PHD domain protein homologous to *Drosophila* trithorax, indirectly regulates H3K4 methylation, represses cyclin A2 expression and promotes myogenic differentiation (Sebastian et al., 2009). MLL5 is a mater factor in mediating regression of established atherosclerosis following lipid lowering interventions by regulating a significant number of relevant genes (Bjorkegren et al., 2014).

Epigenetic mechanisms which regulate chromosomal organization and gene expression include several levels of organization, all of which contribute to the activity of gene expression. The principle mechanisms of epigenetic changes in mammals are DNA methylation and modifications of histone tails which result in altered chromatin structure (Turunen et al., 2009). Although, many researchers and experts recognized a common notice that risk factors (lifestyle, diet and genetic background) do not explain all features of cardiovascular morbidity (Rose, 1964), environmental influence during early life could induce epigenetic variation which may affect metabolism with a lifelong contribution to the cardiovascular health (Alkemade et al., 2007; Barker, 2002). Atherosclerosis is a chronic disease of large and medium sized arteries which is characterized by accumulation of cholesterol in the arterial wall together with proliferation of arterial smooth muscle cells (SMCs) and accumulation of extracellular matrix components which lead to occlusion of blood vessels, myocardial infarction, peripheral vascular disease, amputations, aneurysms and stroke. The genetic heritability of cardiovascular (Yla-Herttuala et al., 1989) disease (CVD) can vary significantly, depending on sex and on the condition in question. Data from some studies suggest a wide range of anywhere between approximately 40% and 80% genetic contribution to CVD (Boomsma et al., 2002; Elder et al., 2009). It has been speculated that epigenetic changes might also contribute to atherogenesis since it involves polyclonal proliferation of SMCs and some dietary components, recognized as risk factors of atherosclerosis, can affect methylation machinery in the arterial cells (Hiltunen and Ylä-Herttuala, 2003).

In our study, we hypothesized that variability in the gene might affect the risk of CAD, genotyped two SNPs of the gene in a Han population, and assessed the association between the polymorphism of the MLL5 gene and CAD using case–control analysis.

Age-related global hypomethylation is the dominant process, but ERα hypermethylation is not an exception and gene-specific hypermethylation may have profound effects on the pathogenesis of atherosclerosis. Age-related promoter hypermethylation of c-myc, c-fos, IGF2, MYOD1, N33, HIC1, versican, PAX6, DBCCR1, E-cadherin and P15 has been reported (Issa, 2000).

Hypomethylation of genomic DNA is characteristic of aging (Wilson and Jones, 1983). Aging and atherosclerosis occur simultaneously and it is sometimes difficult to distinguish between these two processes. In our study age was significantly higher in the CAD group (59.93 ± 9.899 vs. 49.46 ± 9.320, $P = 0.000$) although logistic regressions also find that age is an independent impact factor; whether we include or remove the age factor the P value is still meaningful. So we consider that age can't be a confounding factor.

For the Han population, there was a significant difference in the genotypic distribution of SNP1 (rs12671368) and SNP2 (rs2192932) between CAD patients and control subjects, and the distribution of the dominant model of SNP1 (GG vs. AG + AA) was significantly higher in CAD patients than in control subjects. The distribution of the recessive model of SNP1 (AA vs. AG + GG) was significantly lower in CAD patients than in control subjects. The distribution of the dominant model of SNP2 (GG vs. AG + AA) was significantly lower in CAD patients than in control subjects. The distribution of the recessive model of SNP2 (AA vs. AG + GG) was significantly higher in CAD patients than in control subjects. All the results similarly apply to women.

We found that there was an association between rs12671368 and rs2192932 of the MLL5 gene and CAD only in the female subgroup. This may be attributed to sex hormones. Sex hormones such as estrogens protect against oxidative stress and are known to be vasoprotective (Barnabas et al., 2013). Studies have found that, in patients with coronary heart disease (CHD), ER-a highly methylated promoter regions, and homocysteine may promote the occurrence of methylation and coronary atherosclerosis (Huang et al., 2009). But in our study, this set of data is limited by the smaller sample size for females. More research data is needed to support this conclusion.

In conclusion, we found that rs12671368 and rs2192932 are two novel polymorphisms of the MLL5 gene associated with CAD in a Han Chinese population. The GG genotype of rs12671368, and AA genotype

of rs2192932, with the GA* haplotype could be a protector genetic marker of CAD in a Han Chinese population. This result may broaden the knowledge of genetic variants and disease-association studies.

References

Alkemade, F.E., Gittenberger-de Groot, A.C., et al., 2007. Intrauterine exposure to maternal atherosclerotic risk factors increases the susceptibility to atherosclerosis in adult life. Arterioscler. Thromb. Vasc. Biol. 27 (10), 2228–2235.

Barker, D.J., 2002. Fetal programming of coronary heart disease. Trends Endocrinol. Metab. 13 (9), 364–368.

Barnabas, O., Wang, H., Gao, X.M., 2013. Role of estrogen in angiogenesis in cardiovascular diseases. J. Geriatr. Cardiol. 10 (4), 377–382.

Bjorkegren, J.L.M., Hagg, S., et al., 2014. Plasma cholesterol-induced lesion networks activated before regression of early, mature, and advanced atherosclerosis. PLoS Genet. 10 (2), e1004201.

Boomsma, D., Busjahn, A., Peltonen, L., 2002. Classical twin studies and beyond. Nat. Rev. Genet. 3 (11), 872–882.

Cartwright, E.J., Oceandy, D., Austin, C., Neyses, L., 2011. Ca^{2+} signalling in cardiovascular disease: the role of the plasma membrane calcium pumps. Sci. China Life Sci. 54 (8), 691–698.

Cheng, F., Liu, J., Zhou, S.H., Wang, X.N., Chew, J.F., Deng, L.W., 2008. RNA interference against mixed lineage leukemia 5 resulted in cell cycle arrest. Int. J. Biochem. Cell Biol. 40, 2472–2481.

Deng, L.W., Chiu, I., Strominger, J.L., 2004. MLL 5 protein forms intranuclear foci, and overexpression inhibits cell cycle progression. Proc. Natl. Acad. Sci. U. S. A. 101, 757–762.

Elder, S.J., Lichtenstein, A.H., Pittas, A.G., et al., 2009. Genetic and environmental influences on factors associated with cardiovascular disease and the metabolic syndrome. J. Lipid Res. 50 (9), 1917–1926.

Emerling, B.M., Bonifas, J., et al., 2002. MLL5, a homolog of *Drosophila* trithorax located within a segment of chromosome band 7q22 implicated in myeloid leukemia. Oncogene 21 (31), 4849–4854.

Fujiki, R., Chikanishi, T., Hashiba, W., et al., 2009. GlcNAcylation of a histone methyltransferase in retinoic-acid-induced granulopoiesis. Nature 459 (7245), 455–459 (21).

Heuser, M., Yap, D.B., Leung, M., de Algara, T.R., Tafech, A., McKinney, S., et al., 2009. Loss of MLL5 results in pleiotropic hematopoietic defects, reduced neutrophil immune function, and extreme sensitivity to DNA demethylation. Blood 113, 1432–1443.

Hiltunen, M.O., Ylä-Herttuala, S., 2003. DNA methylation, smooth muscle cells, and atherogenesis. Arterioscler. Thromb. Vasc. Biol. 23 (10), 1750–1753.

Huang, Y.S., Zhi, Y.F., Wang, S.R., 2009. Hypermethylation of estrogen receptor-alpha gene in atheromatosis patients and its correlation with homocysteine. Pathophysiology 16 (4), 259–265.

Issa, J.P., 2000. CpG-island methylation in aging and cancer. Curr. Top. Microbiol. Immunol. 249, 101–118.

Kesteloot, H., Geboers, J., Joossens, J.V., et al., 1989. On the within population relationship between nutrition and serum lipids. The BIRNH study. Eur. Heart J. 10, 196.

Kittler, R., et al., 2007. Genome-scale RNAi profiling of cell division in human tissue culture cells. Nat. Cell Biol. 9, 1401–1412.

Liu, J., WangXN, Cheng F., LiouYC, Deng L.W., 2010. Phosphorylation of mixed lineage leukemia 5 by CDC2 affects its cellular distribution and is required for mitotic entry. J. Biol. Chem. 285, 20904–20914.

Madan, V., Madan, B., Brykczynska, U., Zilbermann, F., Hogeveen, K., Dohner, K., et al., 2009. Impaired function of primitive hematopoietic cells in mice lacking the Mixed-Lineage-Leukemia homolog MLL5. Blood 113, 1444–1454.

Qi, L., Qi, Q., Prudente, S., Mendonca, C., Andreozzi, F., di Pietro, N., et al., 2013. Association between a genetic variant related to glutamic acid metabolism and coronary heart disease in individuals with type 2 diabetes. JAMA 310 (8), 821–828.

Rose, G., 1964. Familial patterns in ischaemic heart disease. Br. J. Prev. Soc. Med. 18, 75–80.

Sebastian, S., Sreenivas, P., Sambasivan, R., et al., 2009. MLL5, a trithorax homolog, indirectly regulates H3K4 methylation, represses cyclin A2 expression, and promotes myogenic differentiation. Proc. Natl. Acad. Sci. U. S. A. 106 (12), 4719–4724 (24).

Turunen, M.P., Aavik, E., et al., 2009. Epigenetics and atherosclerosis. Biochim. Biophys. Acta 1790 (9), 886–891.

Wilson, V.L., Jones, P.A., 1983. DNA methylation decreases in aging but not in immortal cells. Science 220 (4601), 1055–1057.

Yap, D.B., Walker, D.C., Prentice, L.M., McKinney, S., Turashvili, G., et al., 2011. Mll5 is required for normal spermatogenesis. PLoS ONE 6, e27127.

Yideng, J., Jianzhong, Z., Ying, H., Juan, S., Jinge, Z., Shenglan, W., Xiaoqun, H., Shuren, W., 2007. Homocysteine-mediated expression of SAHH, DNMTs, MBD2, and DNA hypomethylation potential pathogenic mechanism in VSMCs. DNA Cell Biol. 26 (8), 603–611 (Aug).

Yla-Herttuala, S., Palinski, W., et al., 1989. Evidence for the presence of oxidatively modified low density lipoprotein in atherosclerotic lesions of rabbit and man. J. Clin. Invest. 84 (4), 1086–1095.

Zhang, Y., Wong, J., Klinger, M., Tran, M.T., Shannon, K.M., Killeen, N., 2009. MLL5 contributes to hematopoietic stem cell fitness and homeostasis. Blood 113, 1455–1463.

Zhou, P., Wang, Z. Yuan, Zhou, X., Liu, C., Wan, L., et al., 2013. Mixed lineage leukemia 5 (MLL5) regulates cell cycle progression and E2F1-responsive gene expression via association with host cell factor-1 (HCF-1). J. Biol. Chem. 288, 17532–17543.