Plasminogen Activator Inhibitor-1 Gene
Selection of Tagging Single Nucleotide Polymorphisms and Association With Coronary Heart Disease

Shaoyong Su, Shufeng Chen, Jiangong Zhao, Jianfeng Huang, Xiaoling Wang, Runsheng Chen, Dongfeng Gu

Objective—To explore the effect of plasminogen activator inhibitor-1 (*PAI-1*) gene variations on the risk of coronary heart disease (CHD) in Chinese Han population.

Methods and Results—We screened all exons and the promoter region of *PAI-1* gene in 48 patients and identified 17 polymorphisms. Five tagging single nucleotide polymorphisms were selected and genotyped in 816 patients with CHD and 937 controls. In the total sample, no main effects of the loci or haplotypes reached statistical significance after adjusting environmental covariates. However, a strongly significant gene–smoking interaction was observed. Among nonsmokers, 2 polymorphisms located at promoter region (rs2227631 and rs1799889) showed significant association with CHD. The cases had higher frequency of rs2227631 A allele and rs1799889 4G allele than the controls (0.42 versus 0.33, $P=0.001$; 0.60 versus 0.52, $P=0.002$). Haplotype analyses confirmed the effects of the *PAI-1* gene–smoking interaction on CHD risk. Compared with the most common haplotype G-5G-A-A-T (35.1%), the haplotype A-4G-A-A-C (32.7%) significantly increased the risk of CHD with adjusted odds ratio of 1.51 (95% CI, 1.12 to 2.05; $P=0.008$) in nonsmokers.

Conclusion—This study identified a significant interaction between *PAI-1* gene and smoking status. Both single locus and haplotype analyses indicated that rs2227631 A allele and rs1799889 4G allele increased the risk of CHD among nonsmokers in Chinese. **(*Arterioscler Thromb Vasc Biol*. 2006;26:948-954.)**

Key Words: coronary heart disease ■ plasminogen activator inhibitor-1 gene ■ tagging SNP ■ haplotype-based association study ■ case-control

Plasminogen activator inhibitor type 1 (PAI-1) is the primary inhibitor of both tissue- and urinary-type plasminogen activators. A reduced plasma fibrinolytic activity, mainly attributable to increased plasma levels of PAI-1, has been associated with coronary heart disease (CHD)[1] and recurrent myocardial infarction.[2] Experimental induction of endothelial injury can stimulate PAI-1 expression and facilitate thrombosis.[3] Genetically determined variability in PAI-1 expression has been suggested as a risk factor for coronary atherogenesis and thrombosis.[4]

The human *PAI-1* gene is mapped on chromosome 7q21.3-q22, and several polymorphisms within the *PAI-1* gene have been described.[5] A single base (guanosine) insertion/deletion polymorphism (4G/5G) located in the promoter region seems to be functionally important.[6] The homozygous or heterozygous carriage of 4G allele had been associated with higher PAI-1 levels and increased risk of CHD,[4,7–10] although the relationship was not confirmed in other studies.[5,11–13] This inconsistency is not unexpected for a multifactorial disease like CHD. One possible reason is that the effects of the polymorphisms on CHD risk may vary according to the presence or absence of other cardiovascular risk factors that affect PAI-1 concentrations (eg, age, gender, smoking, and obesity).[9,14–19] Another reason is that the above association studies are limited to a single single nucleotide polymorphism (SNP), which not only fails to include all the potential risk-conferring variations in the *PAI-1* gene but also fails to explore all the common haplotype variations of this gene and their potential effects on CHD risk. Neale and Sham argued that this kind of association analysis was potentially problematic in the context of replication because a replication study might not provide supportive or negative evidence if only the associated allele from the initial study was examined. They suggested a gene-based approach in which all variants including single SNP and haplotype variants within a candidate gene are considered jointly.[20] The increasing knowledge of how the pattern of linkage disequilibrium (LD) varies across human genome has enabled the design of selecting a

Original received September 9, 2005; final version accepted December 15, 2005.

From the Division of Population Genetics and Prevention (S.S., S.C., J.Z., J.H., X.W., D.G.), Fu Wai Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China; National Human Genome Center at Beijing (S.S., D.G.), China; and Institute of Biophysics (R.C.), Chinese Academy of Sciences, Beijing, China.

Correspondence to Dongfeng Gu, MD, MS, Professor and Chair, Division of Population genetics and Prevention, Cardiovascular Institute and Fu Wai Hospital, No. 167 Beilishi Rd, Beijing, 100037, PR China. E-mail gudf@yahoo.com

Arterioscler Thromb Vasc Biol. **is available at http://www.atvbaha.org** **DOI: 10.1161/01.ATV.0000204731.17646.f2**

minimized number of SNPs (tagging SNPs [tSNPs]) to capture most information of all variants.[21,22] Furthermore, recent developments of indirect approaches have made it feasible to use the tSNPs to predict the association between those remaining SNPs and the trait.[23]

In the present study, a multistep case-control study was designed for evaluating the contribution of single SNP and haplotype variations in *PAI-1* gene to CHD in a large sample of Chinese Han population. We first resequenced the *PAI-1* gene to identify all putative functional common polymorphisms within a subsample. Then, tSNPs were selected by optimizing the ability of smaller subsets of markers to predict both common haplotypes and remaining polymorphisms. These tSNPs were further genotyped in the main study. Finally, SNP-based and haplotype-based association analyses were performed to test the possible effect of *PAI-1* gene on CHD, as well as the interaction between these variants and environmental factors, including age, sex, body mass index (BMI), and smoking. The potential association between CHD and those remaining SNPs was also predicted.

Methods

Subjects

A total of 816 patients with CHD were recruited from hospitalized patients of Fu Wai Hospital and Cardiovascular Institute (Beijing, China) between October 1997 and September 2001. Individuals who survived an acute myocardial infarction or documented by coronary angiography a ≥70% stenosis in a major epicardial artery were eligible. Subjects with congenital heart disease, cardiomyopathy, valvular disease, and renal or hepatic disease were excluded from the study. A total of 937 controls were randomly selected from subjects participating in a community-based survey of cardiovascular risk factors in Beijing. The control subjects were judged to be free of CHD by history, clinical examination, electrocardiography, and Rose questionnaire.

Details of medical history were obtained from all participants by standardized questionnaire. Blood pressure, height, weight, and waistline were measured by trained nurses according to standardized protocols. Cigarette smokers were defined as persons who smoked ≥100 cigarettes in their lifetime. The protocol was approved by the local bioethical committee, and informed consent was obtained from each participant.

Identification of Polymorphisms and Genotyping

5′ flanking region up to 1.5 kb upstream from transcription-initiation sites, all 9 exons and exon/intron boundaries, as well as 3′UTR of *PAI-1* gene were screened by direct sequencing in 48 randomly selected patients. Five tSNPs were selected (see section below) and genotyped in all 1753 subjects. rs2227631, rs2227639, rs2227660, and rs11178 were genotyped using standard polymerase chain reaction (PCR)/restriction fragment length polymorphism protocols. rs1799889 was genotyped by allele-specific PCR amplification. The primers and related restriction endonuclease can be obtained by request.

Selection of tSNPs and Prediction

Two approaches were used to identify optimal subset of markers. The first one is developed by Stram et al,[22] which chooses a subset of tSNPs by optimizing the predictability of common haplotypes. It considers a measure of association (R^2_H) between the true number of copies of haplotypes and the predicted number of copies of haplotypes that each individual has, where the prediction is based on the knowledge of tSNPs. The program *tagsnps* was used to choose such tSNPs with the following parameters: common haplotypes were defined as the estimated frequency was >5%, and sets of tSNPs resolving the common haplotypes were selected at a R^2_H threshold of 0.9, as suggested by Stram et al.[22]

The second approach is developed by Chapman et al,[23] which chooses an optimal set of tSNPs in such a way that the allele frequencies of those SNPs not selected as tSNPs can be predicted well. A series of prediction equations are calculated and the predictive efficiency is assessed in terms of R^2_L, which measures the proportion of variance of each remaining SNP "explained" by regression on the tSNP alleles (locus-based scoring). These regression equations can also be used to predict the association of those ungenotyped SNPs with the trait by using the LD information between the ungenotyped SNPs and tSNPs in the scan sample of 48 patients. The package *htSNP2* was used to choose such a tSNPs set that could predict remaining SNPs with minimum R^2_L of 0.8, as suggested by Chapman et al.[23] The package *genassoc* was used to predict the association of those ungenotyped SNPs with the trait. The main difference between the approaches of Stram et al[22] and Chapman et al[23] is that the former is based on prediction of extended haplotypes from the marker haplotypes, whereas the latter is based on prediction of single SNP loci.

Association Analyses

The main purpose of our analyses was to test the associations between tSNPs and haplotype variations in the *PAI-1* gene with CHD. We further investigated whether the effect of *PAI-1* gene on CHD was dependent on age, sex, BMI, and smoking status by testing interactions of individual tSNP and haplotypes with these variables.

Analyses were done separately for each of the tSNPs and followed up by haplotype analyses. For individual SNP analyses, we first tested the 2-*df* codominant model using a χ^2 test, and in the presence of a significant association, a dominant, a recessive, and an additive model were further tested to find the best mode of inheritance. Logistic regression was followed to adjust for covariates including age, sex, BMI, smoking status, history of hypertension and diabetes, high-density lipoprotein cholesterol (HDL-C), and low-density lipoprotein cholesterol (LDL-C) levels.

To test the association of statistically inferred haplotypes with CHD, we used the *Haplo. score* approach as outlined by Schaid et al.[24] The method models an individual's phenotype as a function of each inferred haplotype, weighted by their estimated probability, to account for haplotype ambiguity. To explore the interaction between haplotypes and other factors, the *Haplo. glm*[25] approach was performed. Both *Haplo. score* and *Haplo. glm* were implemented in the *Haplo. stats* software. A diplotype analysis was then followed by using a weighted logistic regression, with the weights being the probability for each possible haplotype pair combination for an individual as estimated by *Haplo. score*. Only the haplotypes and diplotypes with frequency >5% were considered for the haplotype and diplotype analyses, respectively.

Descriptive statistical analyses were performed with STATA.[26] The pattern of pairwise LD between the SNPs was measured by D' and r^2 calculated by the software GOLD.[27]

Results

Table 1 shows the distribution of the clinical and biological characteristics of the subjects. Age and the percentage of men were greater among subjects with CHD than among controls. Compared with the controls, the case group had more patients with hypertension and diabetes, as well as more smokers. The case group also had significantly higher BMI, systolic blood pressure, LDL-C, serum triglyceride levels, and fasting glucose levels, and lower HDL-C levels than the controls. Diastolic blood pressure was significantly lower in cases than in controls, which could be attributable to medication use in the patients after they were diagnosed.

A total of 17 polymorphisms in the *PAI-1* gene were detected. Position information and mutation types on all SNPs are presented in Table 2, along with the minor allele

TABLE 1. Comparison of Characteristics Between Cases and Controls

	Cases (n=816)	Controls (n=937)	*P* Value
Age, y	54.50±8.89	52.26±10.35	<0.0001
Gender, male (%)	78.43	74.17	0.037
BMI, kg/m^2	26.52±3.22	24.79±3.29	<0.0001
Waist, cm	91.79±8.88	85.17±9.92	<0.0001
SBP, mm Hg	131.40±20.68	127.17±18.14	<0.0001
DBP, mm Hg	76.37±11.12	80.09±9.85	<0.0001
HDL-C, mg/dL	42.13±9.63	49.51±11.80	<0.0001
LDL-C, mg/dL	127.79±39.69	122.67±32.78	0.0032
TC, mg/dL	202.33±43.72	198.61±37.77	0.0562
TG, mg/dL	160.72±108.23	128.67±85.91	<0.0001
GLU, mg/dL	108.29±37.28	99.65±28.68	<0.0001
Hypertension, yes (%)	63.48	32.76	<0.0001
Diabetes, yes (%)	25.98	8.64	<0.0001
Smoker, yes (%)	62.01	57.10	0.037
Drinker, yes (%)	46.69	42.49	0.078

Means±SD values for continuous variables.

SBP indicates systolic blood pressure; DBP, diastolic blood pressure; TC, total cholesterol; TG, triglycerides; GLU, glucose.

frequencies. There were 6 SNPs newly identified in our data, with relatively low frequencies (Table 2). SNPs with minor allele frequency >5% were included in the after analyses.

TABLE 2. Summary Data on 17 Polymorphisms Discovered in 48 Subjects

Polymorphisms[a]	Region[b]	Position[c]	Alleles[d]	Minor Allele Frequency[e] (%)
rs2227631	5′ flanking	−844	G/A	34.4
-801 T>A*	5′ flanking	−801	T/A	3.1
rs1799889[f]	5′ flanking	−675	4G/5G	44.4
rs2227639	Intron 1	+524	A/T	5.3
rs6092[g]	Exon 2	+1333	G/A	2.5
rs6090[h]	Exon 2	+1339	G/A	2.5
+3879 C>T*	Intron 3	+3879	C/T	7.6
rs2227660	Intron 3	+3891	A/G	7.3
rs2227668	Intron 3	+4471	G/A	7.3
rs2227684	Intron 4	+6547	A/G	42.7
rs2070682	Intron 5	+6883	C/T	42.7
+8574 A>G*	Intron 7	+8574	A/G	1.2
rs11178	3′ UTR	+10700	T/C	41.7
+10779 T>C*	3′ UTR	+10779	T/C	1.1
+10817 A>C*	3′ UTR	+10817	A/C	1.1
+10845 C>T*	3′ UTR	+10845	C/T	1.1
rs7242	3′ UTR	+11061	T/G	42.7

*The SNPs newly identified in the present study.

[a]rs No. was shown if present in the dbSNP database; [b]5′ flanking means the upstream from the first transcribed nucleotide; [c]the base immediately preceding the start of transcription numbered as ′−1′; [d]with major allele given first and minor allele given second; [e]using 48 Chinese Han patients; [f]reported previously as ′4G/5G polymorphism′; [g]A nonsynonymous mutation at codon 15 (ALA/THR); [h]a nonsynonymous mutation at codon 17 (VAL/ILE).

Pairwise LD coefficients of the 10 common SNPs were displayed in supplemental Table I and supplemental Figure I (available online at http://atvb.ahajournals.org). There were almost perfect LDs among rs2227684, rs2070682, rs11178, and rs7242, as well as between rs2227660 and rs2227668. Significant association between rs2227631 and rs1799889 was also observed. The inferred haplotypes of the 10 SNPs were shown in supplemental Table II (available online at http://atvb.ahajournals.org). Only 5 common haplotypes (frequency >5%) were observed, comprising 95% of the total.

A total of 5 tSNPs (rs2227631, rs1799889, rs2227639, rs2227660, and rs11178) were selected by *tagsnps* and *htSNP2* programs. With this optimal set chosen, both common haplotypes and unmeasured loci can be accurately predicted (supplemental Table II). The minimum values of R^2_H and R^2_L were 0.94 for haplotype C and 0.90 for +3879 C>T, respectively. These 5 tSNPs were further genotyped in all subjects.

Table 3 shows the genotype frequencies of the 5 tSNPs in the whole cohort. No significant difference was found between cases and controls for any polymorphisms under the 2-*df* codominant model in either univariate or multivariate analysis. The following interaction models indicated that there were significant interactions between smoking and rs2227631, as well as rs1799889. We then separated the study population into smokers and nonsmokers (Table 4). Only among nonsmokers (characteristics of cases and controls in nonsmokers were summarized in supplemental Table III, available online at http://atvb.ahajournals.org) were the frequencies of rs2227631 A allele and rs1799889 4G allele significantly higher in cases than in controls (0.42 versus 0.33, $P=0.001$; 0.60 versus 0.52, $P=0.002$). With the rs2227631 GG genotype used as the reference, the adjusted odds ratios (ORs) for CHD associated with genotypes GA and AA were 1.60 (95% CI, 1.09 to 2.33; $P=0.016$) and 1.82 (95% CI, 1.08 to 3.07; $P=0.026$), respectively. Compared with the rs1799889 5G allele carriers, the adjusted OR was 1.59 (95% CI, 1.10 to 2.30; $P=0.014$) for 4G4G homozygotes. None of the 5 polymorphisms showed significant interactions with age, BMI, or gender.

The approach we used to select an optimal set of tSNPs[23] can also be used to predict which of those SNPs not typed in the full cohort might also show association with the trait. The Figure shows the fine mapping results (including the ungenotyped SNPs) based on prediction of single SNPs from the combination of the 5 tSNPs. Associations of both tSNPs and predicted values for ungenotyped SNPs with CHD in the whole sample as well as in the nonsmokers under the additive genetic model are indicated by $-\log_{10}$ (*P* value). In the whole samples, the global test was nonsignificant ($P=0.10$), with a significant association between rs1799889 and CHD ($P=0.023$). However, in the nonsmokers, the global test was significant ($P=0.007$) with rs2227631 ($P=0.002$), rs1799889 ($P=0.003$), and rs11178 ($P=0.036$) responsible for this effect. Furthermore, additional significant associations are predicted with SNPs rs2227684, rs2070682, and rs7242 ($Ps=0.029$).

As shown in Table 5, the global *P* value for haplotypes in all the subjects was not significant ($P=0.224$). However, a

TABLE 3. Genotype Distributions of *PAI-1* Gene tSNPs in Cases and Controls

		Case (n=816)		Control (n=937)			
		Count	Frequency (%)	Count	Frequency (%)	*P* Value	*P* Value* (adjusted)
rs2227631	GG	304	37.62	393	42.3		
	GA	381	47.15	406	43.7		
	AA	123	15.22	139	13.99	0.139	0.096
rs1799889	4G4G	272	33.5	275	29.54		
	4G5G	390	48.03	446	47.91		
	5G5G	150	18.47	210	22.56	0.059	0.153
rs2227639	AA	735	91.19	862	92.39		
	AT	68	8.44	69	7.4		
	TT	3	0.37	2	0.21	0.626	0.498
rs2227660	AA	733	90.27	854	91.83		
	AG	73	8.99	71	7.63		
	GG	6	0.74	5	0.54	0.525	0.892
rs11178	TT	211	25.99	280	30.24		
	TC	413	50.86	436	47.08		
	CC	188	23.15	210	22.68	0.130	0.272

*Age, sex, BMI, HDL-C, LDL-C, hypertension, diabetes, and smoking were adjusted.

significant interaction between haplotypes and smoking status was observed. In nonsmokers, the haplotype global score was significant ($P=0.015$), with higher frequency of Hap2 (A-4G-A-A-C) in cases than in controls. With the most common haplotype G-5G-A-A-T (Hap1) used as the reference, the adjusted OR for CHD associated with Hap2 was 1.51 (95% CI, 1.22 to 2.05; $P=0.008$). The diplotype analyses showed the same results in nonsmokers ($P=0.006$). Compared with the homozygote of Hap1, the ORs for the heterozygote of Hap1 and Hap2, and the homozygote of Hap2 were 1.76 (95% CI, 1.04 to 2.97; $P=0.034$) and 2.20 (95% CI, 1.19 to 4.09; $P=0.012$), respectively (Table 5).

Discussion

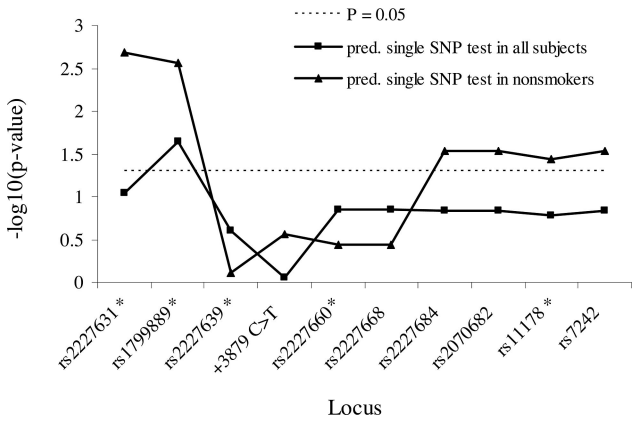
Among the 10 common SNPs, rs2227631 (-844G>A), rs1799889 (-675 4G/5G polymorphism), and rs7242 (+11053T>G) have been reported previously[5,28] (names of the polymorphisms used in their publications are in parentheses). The frequencies of rs1799889 5G allele (0.44) and rs7242 G allele (0.43) were similar, whereas the frequency of rs2227631 A allele (0.34) was relatively lower in our Chinese Han population compared with the reported allele frequencies (0.5 for rs1799889 5G, 0.4 for rs7242 G, and 0.55 for rs2227631 A) in whites.[5]

The possible association between the rs1799889 and the risk of cardiovascular disease has been studied extensively, leading to controversial results.[4,5,8–14] In the present study, we found a significant interaction between this polymorphism and smoking status, with the 4G allele increasing the CHD risk in nonsmokers, but not in smokers or in whole samples, which was consistent with some previous reports.[29,30] Hoekstra et al[29] did not observe association of the 4G allele with the higher prevalence of atherosclerosis in white smoking males. In a study in consecutive patients with coronary stent placement, after 6-month angiographic follow-up, Ortlepp et al found that nonsmoking 4G/4G carriers showed a significant greater late lumen loss compared with nonsmoking 4G/5G or 5G/5G carriers, whereas smoking 5G/5G carriers had the highest late loss of all smoking patients.[30] These studies indicated that the 4G allele and 5G allele might play different roles in the atherosclerosis process according to the absence or presence of cigarette smoking. Another possible

TABLE 4. Adjusted OR for CHD Associated With rs2227631 and rs1799889, According to Smoking Status

		Nonsmokers (n=712)		Smokers (n=1041)	
		OR (95% CI)	*P* Value[a]	OR (95% CI)	*P* Value[a]
rs2227631	GG[b]	—	—	—	—
	GA	1.60 (1.09–2.33)	0.016	1.13 (0.83–1.53)	0.436
	AA	1.82 (1.08–3.07)	0.026	0.84 (0.54–1.30)	0.438
rs1799889	5G5G[c]	—	—	—	—
	4G5G	1.30 (0.82–2.08)	0.269	1.18 (0.82–1.70)	0.373
	4G4G	1.92 (1.16–3.16)	0.011	1.08 (0.72–1.60)	0.711

[a]Adjusted for age, sex, BMI, HDL-C, LDL-C, hypertension, diabetes; [b]GG genotype was used as the reference; [c]5G5G genotype was used as the reference.



Prediction of all association from tSNPs. rs2227631, rs1799889, rs2227639, rs2227660, and rs11178 were tSNPs.

explanation is that the impairment of fibrinolytic mechanisms in smokers might occur as a consequence of smoking that was known to increase the level of PAI-1,[31] thus possibly masking the influence of individual variants of this polymorphism in our population. In contrast, Gardemann et al[19] observed that the association between the 4G allele and coronary stenosis was more pronounced in current and former smokers than in the whole population in whites. However, both cases and controls in their study were patients who underwent coronary angiography for diagnostic purposes and were categorized as cases or controls at a criterion of coronary stenosis ≥50% or <50%, whereas cases in our study were survivors of an acute myocardial infarction or having a coronary stenosis ≥70%, and controls were randomly selected from a Chinese population–based sample, which might cause the discrepancy.

In addition to the rs1799889 polymorphism, another SNP located in promoter region of *PAI-1* gene, rs2227631, is also potentially implicated in the regulation of the PAI-1 gene. However, to date, very few data are available on the relationships between this SNP and CHD. Only 1 study was performed, without observing any impact of the rs2227631 on the risk and extent of CHD.[32] Our study is the first report that a significant association between this polymorphism and CHD was found in nonsmokers with cases having a higher frequency of rs2227631 A allele.

The tSNPs were also used to predict associations of the ungenotyped SNPs with the trait. Among nonsmokers, besides rs2227631 and rs1799889, rs11178 T allele also showed significant association with increased CHD risk. Given the fact that rs11178 was genotyped and we did not observe the

TABLE 5. Association Between *PAI-1* Gene Haplotype, Diplotype, and CHD

		Frequency				
Variables	Haplotype[a]	All	Case	Control	OR (95% CI)	*P* Value[b]
In all subjects						
Hap1*	G-5G-A-A-T	0.352	0.335	0.367	—	—
Hap2	A-4G-A-A-C	0.329	0.347	0.313	1.17 (0.98–1.42)	NS
Hap3	G-4G-A-A-C	0.095	0.096	0.095	1.07 (0.80–1.42)	NS
Hap4	G-4G-A-A-T	0.093	0.095	0.092	1.03 (0.77–1.39)	NS
Global test[c]						0.224
In nonsmokers						
Hap1*	G-5G-A-A-T	0.351	0.317	0.377	—	—
Hap2†	A-4G-A-A-C	0.327	0.376	0.289	1.51 (1.12–2.05)	0.008
Hap3	G-4G-A-A-C	0.106	0.102	0.109	1.06 (0.68–1.67)	NS
Hap4	G-4G-A-A-T	0.092	0.093	0.091	1.06 (0.66–1.71)	NS
Global test[c]						0.015
In nonsmokers	Diplotype					
Hap1-Hap1**	G-5G-A-A-T G-5G-A-A-T	0.133	0.112	0.149	—	—
Hap1-Hap2†	G-5G-A-A-T A-4G-A-A-C	0.228	0.263	0.200	1.76 (1.04–2.97)	0.034
Hap2-Hap2†	A-4G-A-A-C A-4G-A-A-C	0.107	0.137	0.083	2.20 (1.19–4.09)	0.012
Hap1-Hap3	G-5G-A-A-T G-4G-A-A-C	0.072	0.047	0.092	0.63 (0.29–1.35)	NS
Hap2-Hap3	A-4G-A-A-C G-4G-A-A-C	0.069	0.075	0.065	1.53 (0.76–3.09)	NS
Hap1-Hap4	G-5G-A-A-T G-4G-A-A-T	0.061	0.046	0.072	0.85 (0.40–1.83)	NS
Hap2-Hap4	A-4G-A-A-C G-4G-A-A-T	0.055	0.040	0.065	0.78 (0.34–1.78)	NS
Global test[d]						0.006

[a]Loci are arranged in the order rs2227631-rs1799889-rs2227639-rs2227660-rs11178; [b]environmental covariates were adjusted; [c]the global score statistic was test with 4 *df*, adjusted for environmental covariates; [d]the global test with 7 *df*.

*Haplotype G-5G-A-A-T (Hap1) was chosen to be the baseline haplotype; **diplotype Hap1-Hap1 was chosen to be the baseline diplotype; †significant variables.

interaction between rs11178 and smoking status in the whole cohort, the modest correlation between rs11178 and CHD in nonsmokers might be attributable to the strong LD of this SNP with rs2227631 and rs1799889 (supplemental Table I). Furthermore, additional significant associations are predicted with SNPs rs2227684, rs2070682, and rs7242 in nonsmokers, which was expected because of the almost perfect LD between these 3 SNPs and rs11178. Considering the fact that we included all the potential functional SNPs and the 2 SNPs (rs2227631 and rs1799889) that located in the promoter region showed the most significant associations with CHD and have been observed to have impact on PAI-1 gene expression,[4,6,28] this indicates that these 2 SNPs or the other unknown SNPs locating in the 5′ flanking regions (out of our scanning region) might be the functional loci. Further functional studies are necessary to explore whether some specific haplotype pattern constructed by these 2 SNPs affects PAI-1 gene expression. In addition, validation of SNPs in the 5′ flanking region and explore whether they are in strong LD with these 2 SNPs are also warranted. One limitation of the present study is that there was no sodium-citrated blood collected at baseline. Therefore, we were not able to measure PAI-1 level or activity. However, plasma PAI-1 might not be a good marker of fibrinolytic potential in etiologic studies because it is sensitive to diurnal variation and is highly dependent on numerous confounding factors such as lipids, insulin, sex hormones, and inflammatory response.[33,34]

Similar to the single locus analyses, the exhaustively haplotype analyses also found significant interaction between haplotypes and smoking status. Only in nonsmokers was the main impact of haplotypes on CHD observed. Compared with the baseline haplotype G-5G-A-A-T (Hap1, haplotype A-4G-A-A-C (Hap2) was found to significantly increase the risk of CHD. The differentiating characteristics of Hap2 with Hap1 were the rs2227631 A allele and rs1799889 4G allele, which was consistent with the results of single SNP analyses. The rs11178 was another discriminator between Hap1 and Hap2, which might be attributable to the strong LD of this SNP with rs2227631 and rs1799889 as stated above.

In conclusion, the present association study identified a significant interaction between *PAI-1* gene and smoking status. Both single locus and haplotype analyses indicated that rs2227631 A allele and rs1799889 4G allele significantly increased risk of CHD among nonsmokers in Chinese Han population.

Acknowledgments

This work was funded by grants 2002BA711A05, 2002BA711A08, and 2002BA711A10 of the National Tenth Five-Year Plan Key Programs from Ministry of Science and Technology of the People's Republic of China and H020220030130 Biomedical Project from the Council of Science and Technology, Beijing.

References

1. Wiman B, Hamsten A. Impaired fibrinolysis and risk of thromboembolism. *Prog Cardiovasc Dis*. 1991;34:179–192.
2. Hamsten A, de Faire U, Walldius G, Dahlen G, Szamosi A, Landou C, Blomback M, Wiman B. Plasminogen activator inhibitor in plasma: risk factor for recurrent myocardial infarction. *Lancet*. 1987;2:3–9.
3. Fujii S, Sawa H, Saffitz JE, Lucore CL, Sobel BE. Induction of endothelial cell expression of the plasminogen activator inhibitor type 1 gene by thrombosis in vivo. *Circulation*. 1992;86:2000–2010.
4. Eriksson P, Kallin B, van 't Hooft FM, Bavenholm P, Hamsten A. Allele-specific increase in basal transcription of the plasminogen-activator inhibitor 1 gene is associated with myocardial infarction. *Proc Natl Acad Sci U S A*. 1995;92:1851–1855.
5. Henry M, Chomiki N, Scarabin PY, Alessi MC, Peiretti F, Arveiler D, Ferrieres J, Evans A, Amouyel P, Poirier O, Cambien F, Juhan-Vague I. Five frequent polymorphisms of the PAI-1 gene: lack of association between genotypes, PAI activity, and triglyceride levels in a healthy population. *Arterioscler Thromb Vasc Biol*. 1997;17:851–858.
6. Dawson SJ, Wiman B, Hamsten A, Green F, Humphries S, Henney AM. The two allele sequences of a common polymorphism in the promoter of the plasminogen activator inhibitor-1 (PAI-1) gene respond differently to interleukin-1 in HepG2 cells. *J Biol Chem*. 1993;268:10739–10745.
7. Dawson S, Hamsten A, Wiman B, Henney A, Humphries S. Genetic variation at the plasminogen activator inhibitor-1 locus is associated with altered levels of plasma plasminogen activator inhibitor-1 activity. *Arterioscler Thromb*. 1991;11:183–190.
8. Ossei-Gerning N, Mansfield MW, Stickland MH, Wilson IJ, Grant PJ. Plasminogen activator inhibitor-1 promoter 4G/5G genotype and plasma levels in relation to a history of myocardial infarction in patients characterized by coronary angiography. *Arterioscler Thromb Vasc Biol*. 1997;17:33–37.
9. Mansfield MW, Stickland MH, Grant PJ. Plasminogen activator inhibitor-1 (PAI-1) promoter polymorphism and coronary artery disease in non-insulin-dependent diabetes. *Thromb Haemost*. 1995;74:1032–1034.
10. Leander K, Wiman B, Hallqvist J, Sten-Linder M, de Faire U. PAI-1 level and the PAI-1 4G/5G polymorphism in relation to risk of non-fatal myocardial infarction: results from the Stockholm Heart Epidemiology Program (SHEEP). *Thromb Haemost*. 2003;89:1064–1071.
11. Ye S, Green FR, Scarabin PY, Nicaud V, Bara L, Dawson SJ, Humphries SE, Evans A, Luc G, Cambou JP, Arveiler D, Bara L, Cambien F. The 4G/5G genetic polymorphism in the promoter of the plasminogen activator inhibitor-1 (PAI-1) gene is associated with differences in plasma PAI-1 activity but not with risk of myocardial infarction in the ECTIM study. Etude CasTemoins de I'nfarctus du Mycocarde. *Thromb Haemost*. 1995;74:837–841.
12. Ridker PM, Hennekens CH, Lindpaintner K, Stampfer MJ, Miletich JP. Arterial and venous thrombosis is not associated with the 4G/5G polymorphism in the promoter of the plasminogen activator inhibitor gene in a large cohort of US men. *Circulation*. 1997;95:59–62.
13. Catto AJ, Carter AM, Stickland M, Bamford JM, Davies JA, Grant PJ. Plasminogen activator inhibitor-1 (PAI-1) 4G/5G promoter polymorphism and levels in subjects with cerebrovascular disease. *Thromb Haemost*. 1997;77:730–734.
14. Panahloo A, Mohamed-Ali V, Lane A, Green F, Humphries SE, Yudkin JS. Determinants of plasminogen activator inhibitor 1 activity in treated NIDDM and its relation to a polymorphism in the plasminogen activator inhibitor 1 gene. *Diabetes*. 1995;44:37–42.
15. Juhan-Vague I, Alessi MC. PAI-1, obesity, insulin resistance and risk of cardiovascular events. *Thromb Haemost*. 1997;78:656–660.
16. Iacoviello L, Burzotta F, Di Castelnuovo A, Zito F, Marchioli R, Donati MB. The 4G/5G polymorphism of PAI-1 promoter gene and the risk of myocardial infarction: a meta-analysis. *Thromb Haemost*. 1998;80:1029–1030.
17. Margaglione M, Cappucci G, d'Addedda M, Colaizzo D, Giuliani N, Vecchione G, Mascolo G, Grandone E, Di Minno G. PAI-1 plasma levels in a general population without clinical evidence of atherosclerosis: relation to environmental and genetic determinants. *Arterioscler Thromb Vasc Biol*. 1998;18:562–567.
18. Burzotta F, Di Castelnuovo A, Amore C, D'Orazio A, Di Bitondo R, Donati MB, Iacoviello L. 4G/5G promoter PAI-1 gene polymorphism is associated with plasmatic PAI-1 activity in Italians: a model of gene-environment interaction. *Thromb Haemost*. 1998;79:354–358.
19. Gardemann A, Lohre J, Katz N, Tillmanns H, Hehrlein FW, Haberbosch W. The 4G4G genotype of the plasminogen activator inhibitor 4G/5G gene polymorphism is associated with coronary atherosclerosis in patients at high risk for this disease. *Thromb Haemost*. 1999;82:1121–1126.
20. Neale BM, Sham PC. The future of association studies: gene-based analysis and replication. *Am J Hum Genet*. 2004;75:353–362.
21. Johnson GC, Esposito L, Barratt BJ, Smith AN, Heward J, Di Genova G, Ueda H, Cordell HJ, Eaves IA, Dudbridge F, Twells RC, Payne F, Hughes W, Nutland S, Stevens H, Carr P, Tuomilehto-Wolf E,

Tuomilehto J, Gough SC, Clayton DG, Todd JA. Haplotype tagging for the identification of common disease genes. *Nat Genet*. 2001;29:233–237.
22. Stram DO, Haiman CA, Hirschhorn JN, Altshuler D, Kolonel LN, Henderson BE, Pike MC. Choosing haplotype-tagging SNPS based on unphased genotype data using a preliminary sample of unrelated subjects with an example from the Multiethnic Cohort Study. *Hum Hered*. 2003; 55:27–36.
23. Chapman JM, Cooper JD, Todd JA, Clayton DG. Detecting disease associations due to linkage disequilibrium using haplotype tags: a class of tests and the determinants of statistical power. *Hum Hered*. 2003;56: 18–31.
24. Schaid DJ, Rowland CM, Tines DE, Jacobson RM, Poland GA. Score tests for association between traits and haplotypes when linkage phase is ambiguous. *Am J Hum Genet*. 2002;70:425–434.
25. Lake SL, Lyon H, Tantisira K, Silverman EK, Weiss ST, Laird NM, Schaid DJ. Estimation and tests of haplotype-environment interaction when linkage phase is ambiguous. *Hum Hered*. 2003;55:56–65.
26. Stata. Stata Corp: Stata statistical software: release 8.0. Stata Corporation, College Station, Tex, 2003.
27. Abecasis GR, Cookson WO. GOLD–graphical overview of linkage disequilibrium. *Bioinformatics*. 2000;16:182–183.
28. Grubic N, Stegnar M, Peternel P, Kaider A, Binder BR. A novel G/A and the 4G/5G polymorphism within the promoter of the plasminogen activator inhibitor-1 gene in patients with deep vein thrombosis. *Thromb Res*. 1996;84:431–443.
29. Hoekstra T, Geleijnse JM, de Waart F, Nederhand R, Kluft C, Kok FJ, Schouten EG. The 4G/5G-polymorphism in the PAI-1 gene is not associated with markers of atherosclerosis in male smokers. *Thromb Res*. 2002;107:115–119.
30. Ortlepp JR, Hoffmann R, Killian A, Lauscher J, Merkelbach-Brese S, Hanrath P. The 4G/5G promotor polymorphism of the plasminogen activator inhibitor-1 gene and late lumen loss after coronary stent placement in smoking and nonsmoking patients. *Clin Cardiol*. 2001; 24:585–591.
31. Simpson AJ, Gray RS, Moore NR, Booth NA. The effects of chronic smoking on the fibrinolytic potential of plasma and platelets. *Br J Haematol*. 1997;97:208–213.
32. Haselbauer A, Haberbosch W, Tillmanns H, Gardemann A. The impact of the PAI-1 A((-844))G promoter polymorphism on the risk and extent of coronary heart disease. *Thromb Haemost*. 2002;88:697–698.
33. Juhan-Vague I, Pyke SD, Alessi MC, Jespersen J, Haverkate F, Thompson SG. Fibrinolytic factors and the risk of myocardial infarction or sudden death in patients with angina pectoris. ECAT Study Group. European Concerted Action on Thrombosis and Disabilities. *Circulation*. 1996;94:2057–2063.
34. Bashir SA, Duffy SW, Qizilbash N. Repeat measurement of case-control data: corrections for measurement error in a study of ischemic stroke and haemostatic factors. *Int J Epidemiol*. 1997;26:64–70.