Angiotensin II type I receptor gene and myocardial infarction: tagging SNPs and haplotype based association study. The Beijing atherosclerosis study

Shaoyong Su[a,b], Jianhong Chen[a,b], Jiangong Zhao[a], Jianfeng Huang[a], Xiaoling Wang[c], Runsheng Chen[d], Dongfeng Gu[a,b]

Objectives **The present study aimed to assess the effect of haplotype variation in angiotensin II type I receptor (*AGTR1*) gene on the risk of myocardial infarction (MI) in Chinese males.**

Methods **We used 48 patients to identify the putative functional polymorphisms in *AGTR1* gene by direct sequencing. The program tagSNPs was used to identify an optimal set of tagging single nucleotide polymorphisms (SNPs). These selected SNPs were then genotyped in 419 male patients with MI and 400 age-matched male controls. The program haplo.stats was used to investigate the relationship between the haplotypes and MI.**

Results **Sixteen polymorphisms in *AGTR1* gene were identified. Based on the linkage disequilibrium pattern among these SNPs, six polymorphisms, SNP1, SNP6–SNP7 and SNP13–SNP15, were selected as haplotype tagging SNPs and further genotyped. Single SNP analyses indicated that the SNP1, SNP6 and SNP13 were significantly associated with MI, adjusted for covariates. Haplotype-based association analyses identified the frequency of haplotype AGATAA was lower in cases than in controls ($P = 0.006$). In comparison, three haplotypes (AAATAA, TAGCAA and AAACAG) were found to significantly increase the risk of MI with adjusted odds ratio equal to 1.33, 1.75 and 2.64, respectively ($P = 0.029$, 0.026 and 0.015).**

Conclusions **Our study suggests that common genetic variations in the *AGTR1* gene may affect the risk of MI in Chinese males, and that there might be several functional variants in *AGTR1* gene and the combined effect of these variants seemed to have a larger effect on the risk of MI in Chinese males. *Pharmacogenetics* 14:673–681 © 2004 Lippincott Williams & Wilkins**

Pharmacogenetics 2004, **14**:673–681

Keywords: myocardial infarction, angiotensin II type I receptor gene, haplotype tagging SNP, haplotype-based association test, case–control

[a]The Division of Population Genetics and Prevention, Fu Wai Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, 167 Beilishi Road, Beijing 100037, China, [b]National Human Genome Center at Beijing, North Yongchang Rd 3-707, Beijing, 100176, China, [c]Georgia Prevention Institute, Medical College of Georgia, Augusta, GA, USA, [d]Institute of Biophysics, Chinese Academy of Sciences, Datun Road 15, Beijing, 100101, China.

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.

Correspondence and requests for reprints to Dongfeng Gu, MD, MS, Professor and Director, Division of Population Genetics and Prevention, Cardiovascular Institute and Fu Wai Hospital, No. 167 Beilishi Rd, Beijing, 100037, PR China. Tel: +86 (10)68331752; fax: +86 (10)88363812; e-mail: gudf@yahoo.com

Received 4 May 2004
Accepted 22 July 1004

Introduction

The renin–angiotensin system (RAS) plays a critical physiological role in the cardiovascular system. In adult humans, most of the known effects of angiotensin II, the major active component of RAS, are mediated by angiotensin II type I receptor (AGTR1). Several groups [1–3] have demonstrated that AGTR1 is present predominantly in vascular smooth muscle cells and cardiomyocytes. The physiological role of this receptor makes *AGTR1* gene as a candidate gene for hypertension and coronary heart disease.

The human *AGTR1* gene is mapped to chromosome 3q21–q25, and consists of five exons, spanning more than 55 kb of genomic DNA. The first four exons encode 5′ untranslated sequences, while the fifth exon spans 2 kb and contains the entire coding region. A SNP located in the 3′ UTR of *AGTR1* gene, +1166A>C, has been characterized and investigated in relation to essential hypertension [4–6], aortic stiffness [7,8], and myocardial infarction (MI) [9–12], with conflicting results obtained. As the +1166A>C polymorphism does not seem to be functional, it was postulated that it might be in linkage disequilibrium (LD) with some unidentified functional variant(s). Recently, a total of nine polymorphisms were identified in the promoter region of the *AGTR1* gene, most of which were in complete (or nearly complete) LD [11,13–15]. The −810T>A polymorphism, or other polymorphisms that are in LD with it, was observed to

0960-314X

be associated with the risk of MI in Caucasian population [11], while Jin *et al.* [15] failed to detect the association between this polymorphism and coronary heart disease in Chinese population.

Of note, these studies only evaluated the association between individual SNPs and disease. Recent studies, both simulation and empirical, have demonstrated that the haplotype-based association test may be inherently more powerful than individual SNP analysis, since the method incorporates LD information from multiple markers [16,17]. Haplotypes are also useful when disease arises from the interaction of multiple cis-acting susceptibility variants. New statistical methods of testing the association between inferred haplotypes and a wide variety of traits have been recently developed. These flexible methods allow adjustment for non-genetic covariates, which is critical when analyzing complex traits. The next issue is, are all common SNPs necessary to construct haplotypes in large-scale association studies? Given the limitation in throughput and the cost of current genotyping technologies, it seems not practical. Fortunately, the increasing knowledge of how the pattern of LD varies across human genome has enabled the design of selecting a minimized number of SNPs (haplotype tagging SNPs, htSNPs) to capture most of the haplotypic diversity and several approaches have been suggested for identifying these optimal htSNPs [18–22]. As for *AGTR1* gene, the only study on LD pattern of these known polymorphisms was conducted by Zhu *et al.* [23,24]. However, they genotyped only five SNPs, three of which located in promoter region were completely concordant and thus redundant. The haplotypes consisting of these five SNPs revealed only the minimum variation in *AGTR1* gene and the patterns of LD in this gene remained unclear.

In the present study, a multi-step case–control study was designed for evaluating the contribution of haplotype variations in *AGTR1* gene to MI in Chinese Han population. We first scanned the *AGTR1* gene to identify all putative functional polymorphisms within a sub-sample. Then, htSNPs were selected by optimizing the predictability of the common haplotypes and further genotyped in the entire study population. Finally, haplotype-based association test was used to test the possible effect of *AGTR1* gene on MI. To the best of our knowledge, this is the most comprehensive study so far to describe the pattern of LD within *AGTR1* gene and also the first study to explore the effect of this gene on coronary heart disease by using htSNPs and haplotype based analysis.

Methods

Subjects

A total of 419 males who survived an acute MI were eligible for this study. They were enrolled from Fu Wai Hospital & Cardiovascular Institute (Beijing, China) between October 1997 and September 2001. Diagnoses of all cases in this study follow strict diagnostic rules based on signs, symptoms, electrocardiograms and cardiac enzymes. Subjects with congenital heart disease, cardiomyopathy, valvular disease, and renal or hepatic disease were excluded. Age-matched (± 2 years) male controls ($n = 400$) were randomly selected from subjects participating in a community-based survey of cardiovascular risk factors in Beijing. This community-based study included 1600 subjects aged 35 to 74 years, and were sampling by gender (50% men and 50% women) and by age distribution based on 1990 China census data. The control subjects were judged to be free of coronary heart disease by history, clinical examination, electrocardiography, and Rose questionnaire. All subjects were Chinese Han nationality.

Details of medical history were obtained from all participants by standardized questionnaire, together with information of drug intake, cigarette smoking, and alcohol consumption. Blood pressure, height, weight and waistline were measured by trained physicians or nurses according to standardized protocols. Informed written consent was obtained from each subject.

Identification of polymorphisms and genotyping

The 5′ flanking region up to 1.2 kb upstream from transcription-initiation sites, all five exons and exon/intron boundaries, as well as the 3′ untranslated region (UTR) of *AGTR1* gene were screened by direct sequencing in 48 randomly selected patients. Fluorescent dye-terminator cycle sequencing was performed, and products were analyzed with an Applied Biosystems 3700 capillary sequencer.

Six htSNPs were selected (see below section) and genotyped in all 819 subjects. SNP1 and SNP13 were genotyped using standard PCR–restriction fragment length polymorphism (RFLP) protocols, with the following primers: SNP1, forward 5′-TGTAGGCTTTG TCCATTTTTT-3′, reverse 5′-ATTGCACTGATTTG ATCTTT-3′; SNP13, forward 5′-CAAAGTCACCTGC ATCATCA-3′, reverse 5′-AGGAAACAGGAAACCCA GTAT-3′. SNP6, SNP7, SNP14 and SNP15 were genotyped using a mismatch-primer/RFLP method. For SNP6 assay, a forward primer with an intentional mismatch (5′-CAGCAGTTTTCTTTAATGTGTCA CC-3′, with mismatch underlined) was used with an unmodified reverse primer. The modified primer creates an *Hpa*II restriction site in the presence of the G allele, which is then detected by restriction enzyme digestion. For other three assays, the forward primers were: SNP7, 5′-CCCAATGTCTCTTACCCAAATTT GAT-3′, *Bcl*I, A allele; SNP14, 5′-CAGCACTTCA CTACCAAATAGGC-3′, *Hae*III, C allele; SNP15, 5′-AAAAGTATATATTCTACACACATAT-3′, *Nde*I, G

allele. Conditions of amplification and primers used in this study can be available from the authors upon request.

Linkage disequilibrium and haplotype structure

The pattern of pairwise LD between the SNPs was measured by Lewontin's disequilibrium coefficient D′. Haplotype frequencies were estimated using the 'partition–ligation' E-M algorithm, as implemented in the tagSNPs program [20]. The relationships between inferred haplotypes were investigated using the Reduced-Median (RM) network algorithm within NETWORK 4.0 software [25]. To reduce reticulation, only haplotypes with estimated frequency >1% were represented.

Selection of htSNPs

The program tagSNPs was used to identify an optimal set of htSNPs that optimize the predictability of common haplotypes by use of the statistic R^2_h, the squared correlation between estimated and true haplotype dosage (i.e., $E[\delta_h(H_i)|G_i]$ vs. $\delta_h(H_i)$). $\delta_h(H_i)$ is defined as the number of copies of haplotype h in pair of haplotypes H_i and $E[\delta_h(H_i)|G_i]$ is the predicted haplotype dosage given the observed genotypes G_i for individual i. We ran the program with the following parameters: common haplotypes were defined as 'the minimal set of haplotypes that covers 80% of existing haplotypes', and sets of htSNPs resolving the common haplotypes were selected at a R^2_h threshold of 0.7.

Association analyses

The main purpose of our analyses was to test the association between htSNPs and haplotype variation in the *AGTR1* gene with MI. Analyses were done separately for each of the htSNPs and followed up by haplotype analyses. Well-known coronary risk factors, including age, body mass index (BMI), smoking, alcohol use, history of hypertension, triglycerides (TG), HDL cholesterol (HDL-C) and glucose (GLU) were used as covariates.

To test the association of statistically inferred haplotypes with MI, we used the Haplo.score approach as outlined by Schaid *et al.* [26]. The method assigns the probability for each haplotype pair in each individual and then directly models an individual's phenotype as a function of each inferred haplotype pair, weighted by their estimated probability, to account for haplotype ambiguity. This program has the advantage that adjustment for covariates and computation of simulation *P*-values for global and each haplotype can be performed. The number of simulations for empirical *P*-values was set as 1000. Since haplo.score does not provide the magnitude of the effect of each haplotype, haplo.glm was performed to calculate adjusted odds ratios (ORs) and 95% CIs for each haplotype. This approach is based on a generalized linear model, and computes the regression of a trait on haplotypes and other covariates [27]. Both haplo.score and haplo.glm were implemented in the haplo.stats software developed using the R language.

Descriptive statistical analyses were performed with SAS software (SAS Institute Inc., Cary, North Carolina, USA). Allele frequencies were calculated by allele counting, and deviations of the observed genotype frequencies from Hardy–Weinberg equilibrium were identified by χ^2 test.

Results

Table 1 shows the distribution of the clinical and biological characteristics of the subjects. Compared with the controls, the case group had more patients with hypertension. The case group also had significantly higher BMI, serum triglyceride levels, and fasting glucose levels, and lower HDL cholesterol levels than the controls. Diastolic blood pressure was significantly lower in cases than in controls, which could be

Table 1 Comparison of characteristics between cases and controls

	Values in		
	Cases ($n = 419$)	Control ($n = 400$)	*P* value
Average age (years)	54.27 ± 9.74	54.23 ± 9.41	0.9496
BMI (kg/m^2)	26.61 ± 3.11	24.56 ± 3.15	<0.0001
SBP (mm Hg)	127.42 ± 19.84	128.66 ± 18.99	0.3643
DBP (mm Hg)	76.83 ± 11.18	80.91 ± 10.30	<0.0001
HDL-C (mg/dl)	40.52 ± 9.10	48.34 ± 11.12	<0.0001
LDL-C (mg/dl)	125.21 ± 38.86	124.5 ± 33.65	0.7790
TC (mg/dl)	198.91 ± 40.94	197.57 ± 36.49	0.6191
TG (mg/dl)	167.5 ± 142.19	123.96 ± 76.19	<0.0001
GLU (mg/dl)	106.32 ± 34.36	97.28 ± 23.57	<0.0001
Hypertension (yes/no)	198/219	137/263	0.0001
Smokers (yes/no)	321/98	286/114	0.0951
Drinkers (yes/no)	239/180	208/191	0.1585

Means ± SD values for continuous variables.

due to the result of medication in the patients after they were diagnosed.

We detected a total of 16 polymorphisms in *AGTR1* gene. Four polymorphisms were identified in the 5′ flanking region, at position −1106, −778, −681 and −478 upstream from the first transcribed nucleotide (SNP1–SNP4). No polymorphism was detected in the first three exons, but a T/C transition was identified in exon 4 (SNP11: rs1800766). A synonymous base substitution in exon 5 was identified (SNP13: rs5182). SNP14 (rs5186), located in the 3′ UTR, was previously described as +1166A>C. Position information and mutation types on all these SNPs are presented in Table 2, along with the minor allele frequencies. No significant deviation from Hardy–Weinberg equilibrium was found for any polymorphism, after Bonferroni corrections were applied.

Pairwise LD coefficients (D′) of the 16 SNPs and corresponding *P* values were displayed in Table 3. The *P*-values were adjusted by the number of SNP pairs in the gene using a Bonferroni correction, because pairwise LD tests within a gene were not independent. The shaded area indicates that significant LD exists within the *AGTR1* gene after adjusting for multiple tests ($P < 0.05/120 = 0.0004$). It was noted that all the polymorphisms in the 5′ flanking region (SNP1–SNP4) were in complete (or nearly complete) association with one another. Table 4 shows the inferred haplotypes (frequency > 0.01) of the 16 SNPs in the 48 subjects. Only two truly common haplotypes were observed comprising 60% of the total. Four additional haplotypes had frequencies around 5% whereas the remaining haplotypes were rare (<3%). Figure 1 shows a reduced median network for the haplotypes in Table 4 and illustrated several points in regard to tagging-SNPs efficiently. Each haplotype was represented by a circle whose area represented the frequency of that haplotype in the sample. SNPs separating haplotypes were indicated on relevant branches. There were long branches separating some of the major haplotypes (A and C–E); these branches featured several SNPs that were in complete LD because they fell on the same branch, for example, branch SNP1–SNP2–SNP3, branch SNP11–SNP13 and branch SNP4–SNP15–SNP5. This immediately suggested that some level of tagging should be effective (to reduce redundant typing on these branches). One discrepant SNP (SNP6: rs2276736), a single SNP that on its own separates two high-frequency haplotypes (A and B), was also evident (see [21], for details).

A total of six htSNPs (SNP1, SNP6, SNP7 and SNP13–15) were selected, through use of tagSNPs program, at a R^2_h threshold of 0.7 (Table 4). Note that the selection of htSNPs was consisted with the features of the haplotype network (Fig. 1). For long branches 1–2–3, 7–11 and 4–15–5, only SNP1, SNP7 and SNP15 were selected as tags. The discrepant SNP6, was also included in the tagSNPs set. These six htSNPs were genotyped in all subjects.

The genotype distributions of the *AGTR1* gene htSNPs are shown in Table 5. Single SNP analyses indicated that the SNP1, SNP6 and SNP13 were significantly associated with MI, adjusted for covariates. The cases had significantly higher frequencies of SNP1 T allele, SNP6 A allele and SNP13 C allele than were seen in the controls. Compared with the SNP1 AA, SNP6 GG or SNP13 TT homozygotes, the adjusted odds ratio (OR) was 1.50 (95% CI 1.04–2.15) for SNP1 T allele

Table 2 Summary data on 16 polymorphisms discovered in 48 subjects

Marker name (dbSNP name[a])	Region[b]	Position[c]	Alleles[d]	Minor allele Frequency[e] (%)
SNP1 (rs275650)	5′ flanking	−1106	A/T	20.2
SNP2 (rs275651)	5′ flanking	−778	T/A	20.2
SNP3 (rs275652)	5′ flanking	−681	T/G	20.2
SNP4 (rs1492078)	5′ flanking	−478	C/T	26.2
SNP5	Intron 1	9920	G/A	2.4
SNP6 (rs2276736)	Intron 2	10208	A/G	34.5
SNP7 (rs388915)	Intron 2	32091	A/G	19
SNP8	Intron 2	32203	A/G	2.4
SNP9	Intron 3	32399	A/G	1.2
SNP10	Intron 3	41880	G/T	1.2
SNP11 (rs1800766)	Exon 4	41977	T/C	17.9
SNP12	Intron 4	43042	G/T	3.6
SNP13 (rs5182)[f]	Exon 5	43730	T/C	25
SNP14 (rs5186)[g]	3′ UTR	44323	A/C	3.6
SNP15 (rs380400)	3′ UTR	45035	A/G	8.3
SNP16	3′ flanking	45126	T/C	1.2

[a]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]A silent mutation at codon 191 (Leucine/Leucine); [g]Reported previously as +1166A>C.

Table 3 Pairwise linkage disequilibrium coefficients between the 16 polymorphisms identified in 48 individuals

	SNP1	SNP2	SNP3	SNP4	SNP5	SNP6	SNP7	SNP8	SNP9	SNP10	SNP11	SNP12	SNP13	SNP14	SNP15	SNP16
SNP1		0.000	0.000	0.000	0.487	0.003	0.000	0.003	0.039	0.039	0.0004	0.699	0.052	0.382	0.171	0.618
SNP2	1		0.000	0.000	0.487	0.003	0.000	0.003	0.039	0.039	0.0004	0.699	0.052	0.382	0.171	0.618
SNP3	1	1		0.000	0.487	0.003	0.000	0.003	0.039	0.039	0.0004	0.699	0.052	0.382	0.171	0.618
SNP4	1	1	1		0.015	0.041	0.000	0.015	0.086	0.086	0.0002	0.301	0.0001	0.301	0.666	0.543
SNP5	1	1	1	1		0.071	0.476	0.835	0.835	0.883	0.413	0.699	0.013	0.0004	0.000	0.878
SNP6	0.747	0.747	0.747	0.477	1		0.0003	0.258	0.262	0.205	0.0005	0.454	0.013	0.020	0.186	0.447
SNP7	0.486	0.486	0.486	0.608	0.197	1		0.476	0.476	0.641	0.000	0.406	0.000	0.002	0.162	0.003
SNP8	1	1	1	1	1	1	0.197		0.835	0.883	0.413	0.788	1.000	0.788	0.651	0.878
SNP9	1	1	1	1	1	1	1	1		0.883	0.413	0.788	0.013	0.788	0.074	0.878
SNP10	1	1	1	1	1	1	1	1	1		0.653	0.850	0.562	0.850	0.750	0.941
SNP11	0.436	0.436	0.436	0.566	0.227	1	1	0.227	1	1		0.424	0.000	0.0001	0.179	0.649
SNP12	0.488	0.488	0.488	0.996	1	0.352	1	1	1	1	1		0.309	0.741	0.577	0.850
SNP13	0.227	0.227	0.227	0.472	1	0.51	0.908	0	1	1	1	1		0.002	0.000	0.562
SNP14	1	1	1	0.996	0.468	1	1	1	1	1	1	1	1		0.285	0.850
SNP15	1	1	1	0.117	1	0.317	1	1	1	1	1	1	1	0.207		0.751
SNP16	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	

*|D′| below diagonal and *P* value above diagonal; **Shading indicates statistical significance ($P < 0.0004$) after adjusting for 120 multiple tests.

carriers, 1.67 (95% CI 1.10–2.59) for SNP6 A allele carriers or 1.47 (95% CI 1.08–2.02) for SNP13 C allele carriers, respectively.

After we ignored the rare haplotypes (with estimated frequency < 0.03), there remained 8 haplotypes to evaluate (shown in Table 6). The haplotype-based test was performed through use of the haplo.stats program. For the haplo.score analysis, the adjusted global score statistic was 13.06, and with 8 df, the *P* value from the χ^2 distribution was 0.11; this *P* value was identical to the empirical *P* value ($P = 0.124$) based on 1000 simulation repetitions. Although the global score was negative, the empirical *P* value for the max (z_k^2) statistic, evaluating whether only a few haplotypes were strongly associated with a trait, was 0.06. Consistent with this statistic, we found the haplotype-specific score value of haplotype AGATAA was −2.73, with a *P* value of 0.0064, which was in close agreement with the empirical *P* value of 0.006 based on simulation (data not shown). The frequency of this haplotype (AGATAA) was significantly lower in cases than in controls (0.248 vs. 0.290).

To find possible risk haplotypes and evaluate the effect of each haplotype, haplo.glm model was further performed, in which haplotype AGATAA (Hap1), the protective one, was chosen as the baseline. The results of the model fits are shown in Table 6. Three patterns of haplotype were found to significantly increase the risk of MI, this was AAATAA (Hap6), TAGCAA (Hap7) and AAACAG (Hap8). Compared with the Hap1, the adjusted odds ratio was 1.33 (95% CI 0.99–1.79, $P = 0.029$) for Hap6, 1.75 (95% CI 1.00–3.08, $P = 0.026$) for Hap7, and 2.64 (95% CI 1.10–6.33, $P = 0.015$) for Hap8, respectively. This result suggested that individuals with these three haplotypes were at a significantly higher risk for the disease, in comparison with the haplotype AGATAA. For example, the relative risk of having MI is estimated as 6.97 for individuals who are homozygous for the AAACAG haplotype, as compared to those who are homozygous for the AGATAA haplotype.

Discussion

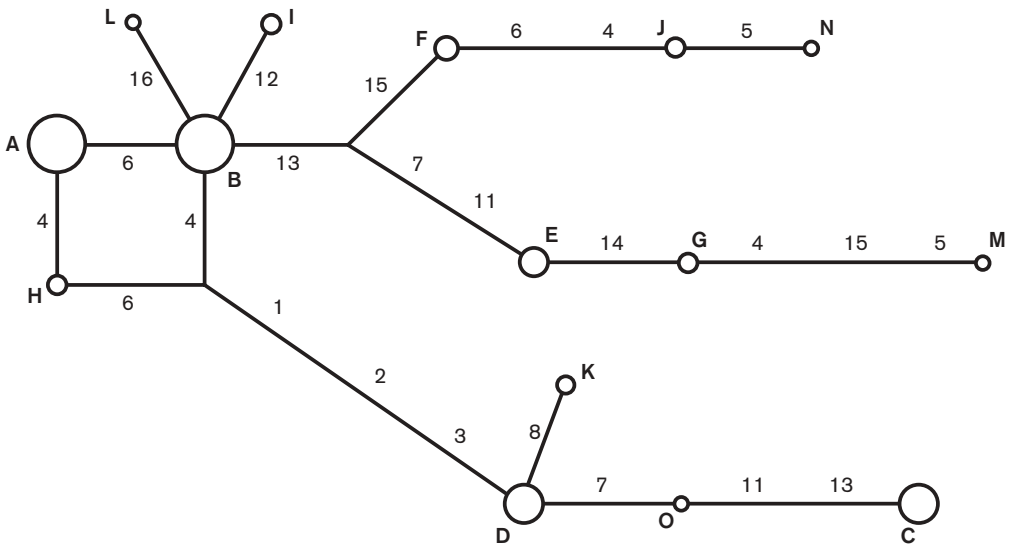
The recent developments of htSNP selection and haplotype-based analysis suggest that currently the most powerful genetic association studies would be (a) to screen the candidate gene for identifying the polymorphisms within a sub-sample randomly chosen from the study subjects and characterize their haplotype structures; (b) to identify the htSNPs sufficient representing the most common haplotypes and further genotype them in all subjects; and (c) to test the association between this gene and trait by using haplotype-based association tests. In the present study,

Table 4 Haplotype structure of *AGTR1* in Chinese Han population

ID[a]	Haplotype[b]	Frequency[c] (%)	Cumulative frequency (%)	R_h^{2d}
A	ATTCGGAAAGTGTAAT	31.2	31.2	0.85
B	ATTCGAAAAGTGTAAT	29.4	60.6	0.80
C	TAGTGAGAAGCGCAAT	7.0	67.6	0.73
D	TAGTGAAAAGTGTAAT	5.5	73.1	0.70
E	ATTCGAGAAGCGCAAT	4.3	77.4	0.91
F	ATTCGAAAAGTGCAGT	3.3	80.7	0.70
G	ATTCGAGAAGCGCCAT	2.5	83.2	–
H	ATTTGGAAAGTGTAAT	2.3	85.5	–
I	ATTCGAAAAGTTTAAT	2.1	87.6	–
J	ATTTGGAAAGTGCAGT	2.0	89.6	–
K	TAGTGAAGAGTGTAAT	1.2	90.8	–
L	ATTCGAAAAGTGTAAC	1.1	91.9	–
M	ATTTAAGAAGCGCCGT	1.0	92.9	–
N	ATTTAGAAAGTGCAGT	1.0	93.9	–
O	TAGTGAGAAGTGTAAT	1.0	94.9	–

[a]Haplotype designation; [b]Loci are arranged in the order SNP1-SNP16 (as in Table 2); [c]Frequencies estimated by the partition-ligation E-M algorithm; [d]The squared correlation between estimated and true haplotype dosage when SNP1, SNP6, SNP7, SNP13–SNP15 selected as htSNPs.

Fig. 1



Reduced median haplotype network based on data in Table 4. SNPs separating haplotypes are indicated on relevant branches.

we applied this strategy to evaluate the association of *AGTR1* gene and MI in Chinese males.

Among the 16 polymorphisms detected within the subsample, SNP5, SNP8-SNP10, SNP12 and SNP16 were newly identified. They were quite rare in our population (maximum frequency of the minor allele was 3.6%), and none of them was chosen as htSNPs. SNP1 (–1138A>T), SNP2 (–810T>A), SNP3 (–713T>G), SNP4 (–521C>T), SNP11 (ex4/+55T>C), SNP13 (ex5/+573T>C), SNP14 (+1166A>C) and SNP15 (+1878A>G) have been previously reported [4,11,13] (what in parentheses were names of the polymorphisms used in their publications). The allele frequencies for some of these polymorphisms differ among ethnic groups. For example, the frequency of SNP4 T allele was reported as 0.36 in Caucasians [11], whereas it is 0.14 in Japanese population [14] and 0.26 in our populations. The frequency of the SNP14 C allele was markedly low in Japanese population (0.079) [6] and Chinese population (0.04) as compared with Caucasians (0.29) [11].

Pairwise LD were measured by D′ among these 16 SNPs, with value of 1 indicating the rarity of recombination between a pair of SNPs. As noted, the number of pairs with $D' = 1$ was 87, however, only 13 of which were significant after being adjusted for multiple tests.

Table 5 Genotype distributions of *AGTR1* gene htSNPs in cases ($n = 419$) and controls ($n = 400$)

htSNP	Genotype counts			*P* value*
SNP1	AA	AT	TT	0.0153
Cases	300	101	16	
Controls	308	80	7	
SNP6	AA	AG	GG	0.0241
Cases	176	186	56	
Controls	145	174	71	
SNP7	AA	AG	GG	0.6414
Cases	273	126	18	
Controls	264	115	16	
SNP13	TT	TC	CC	0.0438
Cases	189	195	32	
Controls	204	158	35	
SNP14	AA	AC	CC	0.8702
Cases	372	44	1	
Controls	357	39	4	
SNP15	AA	AG	GG	0.0946
Cases	310	105	3	
Controls	319	71	6	

*Age, BMI, smoking, alcohol use, history of hypertension, TG, HDL-C and GLU were adjusted.

This could be due to the low minor frequencies of some SNPs, for which test of LD might not have much power [28]. Although the LD decreased between polymorphisms of the promoter region and the SNPs in exon 5 of the *AGTR1* gene, similar to the previous report [23], we also found strong LD between widely-separated SNPs, such as SNP5 and SNP15 (~35 kb separated), indicating that the level of recombination had not been enough to break down the LD throughout the gene. The low haplotype diversity, that is, only six major haplotypes (frequency > 3%) being yielded based on 16 polymorphisms, also suggested a large block of LD across the *AGTR1* gene. The presence of LD throughout the gene meant that the smallest sufficient htSNPs sets would be found by consideration of haplotype for the gene as a whole.

Among the approaches suggested for identifying optimal htSNPs, in present study we adopted the method developed by Stram *et al.* [20], which was based on the measure of the uncertainty in the prediction of common haplotypes from unphased SNP genotypes. Compared with the methods based on capturing as much as possible of the original haplotype diversity (diversity-based measures), this method (association-based measure) is concerned most directly with the issue of prediction and allows a statistically natural approach for assessing how well htSNPs are expected to perform in a genetic association study. In our study, six SNPs with a minimum $R^2_h \geqslant 0.7$ were selected as htSNPs. Although one study by Stram *et al.* [29] found exceedingly modest biases in the risk estimates so long as R^2_h was higher than 0.9 or so, the biases of risk estimate were still relatively mild when $R^2_h = 0.8$. In addition, Stram *et al.* [20] observed the sample size requirements for estimating logistic regression model were almost similar for $R^2_h = 0.9$ and $R^2_h = 0.7$ when the effect of the haplotype was not very small (OR > 1.4). Since at least four more SNPs would be identified as htSNPs and have to be further genotyped in the whole sample if we chose the threshold of $R^2_h = 0.8$, in consideration of cost-effective, the R^2_h threshold was determined as 0.7 in this study.

Among the selected six htSNPs, SNP6, located in intron 2, was a discrepant SNP separated two high-frequency haplotypes. If this SNP were excluded, only one major haplotype would be found (with frequency 0.61). The min R^2_h was only 0.46 when using the

Table 6 Results of haplo.glm applied to the present study using the haplotypes from 6 htSNPs in *AGTR1* gene and environment covariates

Model	Variables	Haplotype[a]	Frequency		Coefficient	Odds ratio (95% CI)	*P* value
			Case	Control			
Environmental covariates	Hypertension[b]				0.237805	1.27 (0.909–1.769)	0.0808
	Smoker[b]				0.061794	1.06 (0.727–1.556)	0.375
	Drinker[b]				0.055455	1.06 (0.757–1.476)	0.372
	Age				0.011098	1.01 (0.994–1.029)	0.103
	BMI				0.141451	1.15 (1.091–1.216)	**1.75×10^{-7}**
	TG				0.000269	1.00 (0.998–1.002)	0.384
	HDL-C				−0.069064	0.93 (0.916–0.951)	**1.12×10^{-12}**
	GLU				0.006638	1.01 (1.001–1.013)	**0.0142**
Haplotypes[c]	Hap1	**AGATAA**	0.248	0.290	–	Reference	–
	Hap2	AAGCCA	0.027	0.036	−0.068784	0.93 (0.46–1.88)	0.424
	Hap3	AAGCAA	0.047	0.061	0.112343	1.12 (0.64–1.95)	0.346
	Hap4	AGACAG	0.063	0.074	0.103836	1.11 (0.65–1.90)	0.353
	Hap5	TAATAA	0.049	0.050	0.34088	1.41 (0.78–2.53)	0.127
	Hap6	**AAATAA**	0.355	0.338	0.286496	1.33 (0.99–1.79)	**0.029**
	Hap7	**TAGCAA**	0.057	0.043	0.560977	1.75 (1.00–3.08)	**0.026**
	Hap8	**AAACAG**	0.037	0.021	0.972031	2.64 (1.10–6.33)	**0.015**

*Significant variables were shown in bold.
[a]Loci are arranged in the order SNP1–6–7–13–14–15; [b]1, with hypertension, 0, without hypertension; 1, smoker, 0, nonsmoker; 1, drinker, 0, nondrinker;
[c]Corresponding to haplotypes in Fig. 1: Hap1 = A/H; Hap2 = G; Hap3 = E; Hap4 = J/N; Hap5 = D/K; Hap6 = B/I/L; Hap7 = C; Hap8 = F.

remaining five htSNPs to predict the common haplotypes, which indicated that lots of efficiencies were lost. In most previous association studies, only the SNPs located in promoter region and/or exons were used to define the haplotypes within a candidate gene. Here we found some SNPs in introns might be more important as htSNPs and haplotype structures could be more clearly revealed when including such SNPs. It is possible that some other discrepant SNPs are yet undiscovered, because large regions of introns were left unscreened and two high-frequency haplotypes existed in our samples. If this is true for *AGTR1* gene, a more exhaustive search should be conducted in different ethnic populations to describe the haplotype structures before further association analysis.

The contributions of variants of *AGTR1* gene to coronary heart disease are conflicting in previous reports. Best evaluated with respect to the association with cardiovascular phenotypes is the +1166A>C polymorphism (SNP14). The possible synergistic interaction of this polymorphism with an I/D polymorphism of angiotensin-converting enzyme (*ACE*) gene on MI or coronary atherosclerosis was also explored, although results were not consistent [9,30–33]. It remains unclear whether the +1166A>C polymorphism in the 3′ UTR of *AGTR1* gene has functional implications. Recently, more concerns were transferred to the functional promoter region of the *AGTR1* gene, which may have the potential to influence *AGTR1* gene expression. In a rat model it has been shown that the *AGTR1* promoter is active in cardiac muscle and its expression is induced by pressure overload, and this response is mediated by a functional interaction between AP-1 and GATA-4 transcription factors [34]. According to a database search, Erdmann *et al.* [13] indicated that –810T>A (SNP2; In perfect LD with SNP1 and SNP3) might destroy a transcription factor binding site for GATA binding factors. Poirier *et al.* [11] found that this polymorphism was associated with the risk of MI in Caucasian males, while a study in Chinese population [15] failed to demonstrate the association.

In the present study, the analyses based on single SNP and haplotype were both conducted. The single SNP analysis showed that SNP1 T allele, SNP6 A allele and SNP13 C allele significantly increased the risk of MI after adjusting covariates. The results of haplotype-specific score test showed that the frequency of AGATAA (Hap1) was significantly lower in cases than in controls (0.248 vs. 0.290, $P = 0.006$). To find possible risk haplotypes and evaluate the effect of each haplotype, we further performed the analysis of haplo.glm, in which the Hap1 was chosen as the baseline haplotype. Three patterns of haplotypes, AAATAA (Hap6), TAGCAA (Hap7) and AAACAG (Hap8), were found to significantly increase the risk of MI, in comparison with the Hap1 (Table 6). An interesting finding is that the Hap6–8 featured a common constellation of the second site (SNP6 A allele). The only difference between Hap1 and Hap6 was at the SNP6. Compared with Hap1, the adjusted OR was 1.33 for Hap6. Furthermore, with the SNP1 T allele, SNP6 A allele and SNP13 C allele presenting in Hap7, and with the SNP6 A allele, SNP13 C allele and SNP15 G allele presenting in Hap8, the relative risks of individuals having MI were increased to 1.75 and 2.64 for Hap7 and Hap8 respectively, compared with Hap1. The results based on haplotype supported the finding of single SNP analysis, that SNP1, SNP6 and SNP13 might be associated with MI. In addition, the haplotype-based analyses also found some potential synergistic interactions among SNP1, SNP6, SNP13 and SNP15. Our findings about SNP1 (in perfect LD with –810T>A) are consistent with the report in Caucasian males [11], whereas SNP6 (located in intron 2) and SNP13 (a synonymous SNP in exon 5) were first explored in a population-based association study. These polymorphisms may be causative or in LD with a functional variant that could affect stability of the mRNA or the regulation of expression of the gene. There is also suggestive evidence that some haplotype patterns (multiple SNPs distributed along the entire gene) might exert a stronger impact on the variability of AGTR1-mediated physiological response. Further functional studies are necessary to explore if these polymorphisms or specific haplotype patterns have impact on *AGTR1* gene expression.

Two limitations of the present study were that, first, only polymorphisms located in promoter region, exons and exon/intron boundaries were identified, and relative large regions of introns were left unscreened. To describe the haplotype structure of *AGTR1* gene completely, it might be necessary to screen more polymorphisms in introns in the future study. Second, using the htSNPs, the association between the *AGTR1* gene and MI was detected. However, we couldn't predict those SNPs not yet genotyped in our study but might be associated with the trait (maybe causal variants). Chapman *et al.* [35] have discussed this issue by performing their methods termed as locus scoring. However, use of this approach may require selection of more markers to achieve the target value. Additional work is needed in this area.

In conclusion, the present association study suggests that common genetic variations in the *AGTR1* gene may affect the risk of MI in Chinese males. The results based on single SNP and haplotypes indicate that there might be several functional variants in *AGTR1* gene and the combined effect of these variants seems to have a larger effect on the risk of MI in Chinese males. The htSNPs analysis suggests that some polymorph-

isms in introns are also required to reveal haplotype structure completely.

Acknowledgements

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 Regitz-Zagrosek V, Friedel N, Heymann A, Bauer P, Neuss M, Rolfs A, *et al.* Regulation, chamber localization, and subtype distribution of angiotensin II receptors in human hearts. *Circulation* 1995; **91**(5):1461–1471.
2 Zhuo J, Moeller I, Jenkins T, Chai SY, Allen AM, Ohishi M, *et al.* Mapping tissue angiotensin-converting enzyme and angiotensin AT1, AT2 and AT4 receptors. *J Hypertens* 1998; **16**(12 Pt 2):2027–2037.
3 Allen AM, Zhuo J, Mendelsohn FA. Localization and function of angiotensin AT1 receptors. *Am J Hypertens* 2000; **13**(1 Pt 2):31S-38S.
4 Bonnardeaux A, Davies E, Jeunemaitre X, Fery I, Charru A, Clauser E, *et al.* Angiotensin II type 1 receptor gene polymorphisms in human essential hypertension. *Hypertension* 1994; **24**(1):63–69.
5 Ono K, Mannami T, Baba S, Yasui N, Ogihara T, Iwai N. Lack of association between angiotensin II type 1 receptor gene polymorphism and hypertension in Japanese. *Hypertens Res* 2003; **26**(2):131–134.
6 Kikuya M, Sugimoto K, Katsuya T, Suzuki M, Sato T, Funahashi J, *et al.* A/C1166 gene polymorphism of the angiotensin II type 1 receptor (AT1) and ambulatory blood pressure: the Ohasama Study. *Hypertens Res* 2003; **26**(2):141–145.
7 Benetos A, Gautier S, Ricard S, Topouchian J, Asmar R, Poirier O, *et al.* Influence of angiotensin-converting enzyme and angiotensin II type 1 receptor gene polymorphisms on aortic stiffness in normotensive and hypertensive patients. *Circulation* 1996; **94**(4):698–703.
8 Lajemi M, Labat C, Gautier S, Lacolley P, Safar M, Asmar R, *et al.* Angiotensin II type 1 receptor-153A/G and 1166A/C gene polymorphisms and increase in aortic stiffness with age in hypertensive subjects. *J Hypertens* 2001; **19**(3):407–413.
9 Tiret L, Bonnardeaux A, Poirier O, Ricard S, Marques-Vidal P, Evans A, *et al.* Synergistic effects of angiotensin-converting enzyme and angiotensin-II type 1 receptor gene polymorphisms on risk of myocardial infarction. *Lancet* 1994; **344**(8927):910–913.
10 Nakauchi Y, Suehiro T, Yamamoto M, Yasuoka N, Arii K, Kumon Y, *et al.* Significance of angiotensin I-converting enzyme and angiotensin II type 1 receptor gene polymorphisms as risk factors for coronary heart disease. *Atherosclerosis* 1996; **125**(2):161–169.
11 Poirier O, Georges JL, Ricard S, Arveiler D, Ruidavets JB, Luc G, *et al.* New polymorphisms of the angiotensin II type 1 receptor gene and their associations with myocardial infarction and blood pressure: the ECTIM study. Etude Cas-Temoin de l'Infarctus du Myocarde. *J Hypertens* 1998; **16**(10):1443–1447.
12 Fernandez-Arcas N, Dieguez-Lucena JL, Munoz-Moran E, Ruiz-Galdon M, Espinosa-Caliani S, Aranda-Lara P, *et al.* The genotype interactions of methylenetetrahydrofolate reductase and renin-angiotensin system genes are associated with myocardial infarction. *Atherosclerosis* 1999; **145**(2):293–300.
13 Erdmann J, Riedel K, Rohde K, Folgmann I, Wienker T, Fleck E, *et al.* Characterization of polymorphisms in the promoter of the human angiotensin II subtype 1 (AT1) receptor gene. *Ann Hum Genet* 1999; **63 (Pt 4)**:369–374.
14 Takahashi N, Murakami H, Kodama K, Kasagi F, Yamada M, Nishishita T, *et al.* Association of a polymorphism at the 5′-region of the angiotensin II type 1 receptor with hypertension. *Ann Hum Genet* 2000; **64 (Pt 3)**: 197–205.
15 Jin W, Liu Y, Sheng HH, Jin L, Shen YY, Hua Q, *et al.* Single nucleotide polymorphisms in promoter of angiotensin II type 1 receptor gene associated with essential hypertension and coronary heart disease in Chinese population. *Acta Pharmacol Sin* 2003; **24**(11):1083–1088.
16 Akey J, Jin L, Xiong M. Haplotypes vs single marker linkage disequilibrium tests: what do we gain? *Eur J Hum Genet* 2001; **9**(4):291–300.
17 Zaykin DV, Westfall PH, Young SS, Karnoub MA, Wagner MJ, Ehm MG. Testing association of statistically inferred haplotypes with discrete and continuous traits in samples of unrelated individuals. *Hum Hered* 2002; **53**(2):79–91.
18 Johnson GC, Esposito L, Barratt BJ, Smith AN, Heward J, Di Genova G, *et al.* Haplotype tagging for the identification of common disease genes. *Nat Genet* 2001; **29**(2):233–237.
19 Zhang K, Deng M, Chen T, Waterman MS, Sun F. A dynamic programming algorithm for haplotype block partitioning. *Proc Natl Acad Sci U S A* 2002;**99**(11):7335–7339.
20 Stram DO, Haiman CA, Hirschhorn JN, Altshuler D, Kolonel LN, Henderson BE, *et al.* 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** (1): 27–36.
21 Weale ME, Depondt C, Macdonald SJ, Smith A, Lai PS, Shorvon SD, *et al.* Selection and evaluation of tagging SNPs in the neuronal-sodium-channel gene SCN1A: implications for linkage-disequilibrium gene mapping. *Am J Hum Genet* 2003; **73**(3):551–565.
22 Carlson CS, Eberle MA, Rieder MJ, Yi Q, Kruglyak L, Nickerson DA. Selecting a maximally informative set of single-nucleotide polymorphisms for association analyses using linkage disequilibrium. *Am J Hum Genet* 2004; **74**(1):106–120.
23 Zhu X, Cooper RS. Linkage disequilibrium analysis of the renin-angiotensin system genes. *Curr Hypertens Rep* 2003; **5**(1):40–46.
24 Zhu X, Yan D, Cooper RS, Luke A, Ikeda MA, Chang YP, *et al.* Linkage disequilibrium and haplotype diversity in the genes of the renin-angiotensin system: findings from the family blood pressure program. *Genome Res* 2003; **13**(2):173–181.
25 Bandelt H-J, Forster P, Sykes BC, Richards MB. Mitochondrial portraits of human populations using median networks. *Genetics* 1995; **141**(2):743–753.
26 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**(2):425–434.
27 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**(1):56–65.
28 Thompson EA, Deeb S, Walker D, Motulsky AG. The detection of linkage disequilibrium between closely linked markers: RFLPs at the AI-CIII apolipoprotein genes. *Am J Hum Genet* 1988; **42**(1):113–124.
29 Stram DO, Leigh Pearce C, Bretsky P, Freedman M, Hirschhorn JN, Altshuler D. Modeling and E-M estimation of haplotype-specific relative risks from genotype data for a case-control study of unrelated individuals. *Hum Hered* 2003; **55**(4):179–190.
30 Alvarez R, Reguero JR, Batalla A, Iglesias-Cubero G, Cortina A, Alvarez V, *et al.* Angiotensin-converting enzyme and angiotensin II receptor 1 polymorphisms: association with early coronary disease. *Cardiovasc Res* 1998; **40**(2):375–379.
31 Ye S, Dhillon S, Seear R, Dunleavey L, Day LB, Bannister W, *et al.* Epistatic interaction between variations in the angiotensin I converting enzyme and angiotensin II type 1 receptor genes in relation to extent of coronary atherosclerosis. *Heart* 2003; **89**(10):1195–1199.
32 Gardemann A, Nguyen QD, Humme J, Stricker J, Katz N, Tillmanns H, *et al.* Angiotensin II type 1 receptor A1166C gene polymorphism. Absence of an association with the risk of coronary artery disease and myocardial infarction and of a synergistic effect with angiotensin-converting enzyme gene polymorphism on the risk of these diseases. *Eur Heart J* 1998; **19**(11):1657–1665.
33 Hernandez Ortega E, Medina Fernandez-Aceituno A, Rodriguez Esparragon FJ, Hernandez Perera O, Melian Nuez F, Delgado Espinosa A, *et al.* [The involvement of the renin-angiotensin system gene polymorphisms in coronary heart disease]. *Rev Esp Cardiol* 2002; **55**(2):92–99.
34 Herzig TC, Jobe SM, Aoki H, Molkentin JD, Cowley AW Jr, Izumo S, *et al.* Angiotensin II type1a receptor gene expression in the heart: AP-1 and GATA-4 participate in the response to pressure overload. *Proc Natl Acad Sci U S A* 1997; **94**(14):7543–7548.
35 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**(1–3):18–31.

Web site references

http://www.ncbi.nlm.nih.gov/Omim;
Online Mendelian Inheritance in Man (OMIM); for *AGTR1* [MIM 106165]
http://www-rcf.usc.edu/~stram/tagSNPs.html;
Haplotype tagging SNPs (htSNPs) selection (for tagSNPs software)
http://www.fluxus-engineering.com/sharenet.htm;
NETWORK 4.0 software
http://www.mayo.edu/hsr/people/schaid.html;
Haplo.stats software (haplo.score and haplo.glm)