Hum Genet (2007) 121:327–335
DOI 10.1007/s00439-007-0340-4

ORIGINAL INVESTIGATION

Common variation in KLKB1 and essential hypertension risk: tagging-SNP haplotype analysis in a case-control study

Xiangfeng Lu · Weiyan Zhao · Jianfeng Huang · Hongfan Li · Wei Yang · Laiyuan Wang · Wentao Huang · Shufeng Chen · Dongfeng Gu

Received: 20 September 2006 / Accepted: 2 February 2007 / Published online: 23 February 2007

Abstract The human plasma kallikrein gene (KLKB1) encodes plasma kallikrein, a serine protease that catalyzes the release of kinins and other vasoactive peptides and may be involved in the pathogenesis of hypertension. In this study, we performed a haplotype-based study to assess the effect of common genetic variation in the KLKB1 gene on the risk of essential hypertension. Eight common single nucleotide polymorphisms (SNPs) were selected from the HapMap database and used to determine the pattern of linkage disequilibrium (LD) and haplotype structure within the KLKB1 gene. Four tag SNPs were then identified with over 85% power to predict both common haplotypes and remaining common SNPs, and genotyped in 1,317 cases with essential hypertension and 1,269 healthy controls. Single SNP analyses indicated that SNPs rs2304595 and rs4253325 were significantly associated with hypertension, adjusted for covariates. Compared with the most common Hap2 CAGC, Hap1 AGAC and Hap3 CGAC, which carry the susceptible rs2304595 G allele and rs4253325 A allele, were found to significantly increase the risk of essential hypertension with adjusted odds ratios equal to 1.37 and 1.17, respectively ($P < 0.0001$ and 0.028). A strongly significant interaction with gene-drinking was also observed. Among drinkers, the adjusted OR for Hap1 relative to Hap2 was increased to 2.50 (95% CI, 2.40 to 2.61; $P < 0.0001$). This was the first study to perform association analysis of the KLKB1 gene with essential hypertension. Our findings suggested that common genetic variation in the KLKB1 gene might contribute to the risk of hypertension in the northern Han Chinese population.

Keywords Association study · Haplotype · HapMap · Human plasma kallikrein · Hypertension · Tag SNP

Conflict of interests: None.

Electronic supplementary material The online version of this article (doi:10.1007/s00439-007-0340-4) contains supplementary material, which is available to authorized users.

X. Lu · W. Zhao · J. Huang · H. Li · W. Yang · S. Chen · D. Gu (✉)
Department of Evidence Based Medicine and Division of Population Genetics, Cardiovascular Institute and Fu Wai Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, 167 Beilishi Rd, Beijing 100037, People's Republic of China
e-mail: gudf@yahoo.com

X. Lu · L. Wang · D. Gu
National Human Genome Center at Beijing, Beijing 100176, People's Republic of China

W. Huang
Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, People's Republic of China

Introduction

Essential hypertension is a complex disease where both genetic and environmental factors interact to produce the phenotype. The genetic contribution to blood pressure (BP) variation ranges from 30 to 50% (Ward 1990). It is likely that a number of genes with smaller effects account for the heritability of this complex disorder. Candidate genes that determine BP variation include those whose products have a direct role in BP regulation such as plasma kallikrein.

Plasma kallikrein is a serine protease that is synthesized in the liver as plasma prekallikrein, secreted into the blood, and converted into plasma kallikrein. Then plasma kallikrein acts on high molecular weight kininogen substrate to release bradykinin and converts prorenin to renin. By controlling the release of bradykinin (a potent vasodilator) and the activation of renin (the protease converting angiotensinogen to angiotensin I), plasma kallikrein is deeply involved in BP regulation (Dielis et al. 2005; Marcondes and Antunes 2005; Schmaier 2003; Tang et al. 2005). Based on the physiological effects, human plasma kallikrein gene (KLKB1) encoding plasma kallikrein can be considered as a good candidate gene for essential hypertension. The KLKB1 gene is mapped to the chromosome 4q34–q35, consists of 15 exons, and spans approximately 31 kb (Beaubien et al. 1991; Yu et al. 2000, 1998). So far, the association of the variation in the KLKB1 gene with hypertension has not been explored.

The haplotype-based association studies may be inherently more powerful than individual SNP analysis to identify causal genetic variants underlying complex disease, since the method incorporates LD information from multiple markers (Daly et al. 2001; Gabriel et al. 2002). This approach does not require the causal variants to be identified or directly tested, but rather has the potential to highlight physical regions that harbor putative disease-associated variants. The increasing knowledge of how the pattern of LD varies across human genome has enabled the design of selecting a minimum number of SNPs (tagging SNPs [tag SNPs]) to capture most of the haplotypic diversity, and several approaches have been suggested for identifying these optimal tag SNPs (Carlson et al. 2004; Johnson et al. 2001; Stram et al. 2003; Weale et al. 2003). The international HapMap project (2003) is a resource that provides empirical genome-wide data to support such analyses . Furthermore, recent developments of indirect approaches have made it feasible to use the tag SNPs to predict the association between those remaining (ungenotyped) SNPs and the trait (Chapman et al. 2003).

In the present study, we have employed the haplotype-based approach to examine the contribution of common variation in the KLKB1 gene to the risk of hypertension in a population-based case-control study. Tag SNPs were selected from HapMap data by optimizing the predictability of both common haplotypes and remaining SNPs. The potential association between those remaining SNPs and essential hypertension was also predicted.

Methods

Study population

This study was based on the International Collaborative Study of Cardiovascular Disease in Asia (InterASIA), from which all the DNA samples and clinical data for participants were obtained (Gu et al. 2002, 2005). The local bioethical committee approved the protocol, and informed consent was obtained from each participant. InterASIA selected a nationally representative sample of the general population aged 35–74 years in China. A total of 15,838 persons completed the survey and examination. Among these, we enrolled 1,317 unrelated hypertensive patients and 1,269 age- and gender-matched unrelated normotensive controls from four northern field centers of InterASIA, namely Beijing, Jilin, Shandong, and Shaanxi province, where high prevalence of cardiovascular morbidity and mortality have been observed. All measurements and interviews were taken under standard conditions as previously described (Gu et al. 2002). Three BP measurements were obtained from each participant by trained and certified observers according to a standard protocol recommended by the American Heart Association (Perloff et al. 1993). Hypertension was defined as an average systolic BP of $\geq$ 150 mmHg, an average diastolic BP $\geq$ 95 mmHg, or current treatment for hypertension with antihypertensive medication. Control subjects had systolic BP < 140 mmHg and diastolic BP < 90 mmHg. Subjects with a clinical history of secondary hypertension, coronary heart disease and diabetes were excluded from the study.

SNP identification and genotyping

We searched the Phase I HapMap data (http://www.hapmap.org, public release up to June 2005), and selected common SNPs (minor allele frequency [MAF] $\geq$ 5%) within a 36-kb region spanning the KLKB1 gene (including 5-kb upstream and downstream of the gene).

The selected tag SNPs (see section below) were genotyped in all 2,586 subjects by polymerase chain reaction/restriction fragment length polymorphism (PCR/RFLP) protocols. The primers, lengths of PCR products, related restriction endonuclease, as well as digested bands are shown in Table 1. Each 10 μl reaction consisted of approximately 25 ng of genomic DNA, 0.1 μM each primer, 0.2 mM dNTPs, 2.5 mM $MgCl_2$, 1.0 unit of Taq polymerase (TaKaRa Biotechnology Co. Ltd., Kyoto, Japan) and 1.0 μl reaction buffer (TaKaRa). Samples were subjected to denaturation at 94°C for 5 min,

Table 1 Primer sequences and restriction enzymes for KLKB1 tag SNPs

SNPs	Primer sequences	PCR product	Restriction enzymes	Digested bands
rs2278542	5′ TCTAAATGGCAAACAGTT3′ (F) 5′ AACATCCACATCTGAGAGC3′ (R)	226 bp	SacI[a]	A allele: 203 C allele: 226
rs2304595	5′ CCCAGAAGAATAAAATTTGAT3′ (F) 5′ TTGCAGGTTAAAAGGCTATAT3′ (R)	212 bp	Mph1103I[b]	G allele: 212 A allele: 190
rs4253325	5′ ATACACTGCTCTGATTCACTT 3′ (F) 5′ TTATAGCCAGCACAGACCAGC 3′ (R)	256 bp	PvuII[a]	A allele: 236 G allele: 256
rs925453	5′ GTCCCTTAGTTTGCAAACACCA 3′ (F) 5′ GCTCAGAATTTGACTTGAACT 3′ (R)	238 bp	Eco130I[b]	C allele: 238 T allele: 217

Underlined loci were mismatched

F forward primer, *R* reverse primer

[a] MBI Biotechnology Co. Ltd., Burlington, Canada

[b] TaKaRa Biotechnology Co. Ltd., Kyoto, Japan

followed by 30 cycles of 94°C for 15 s, annealing at 48°C (for rs2304595) or 60°C (for rs2278542, rs4253325 and rs925453) for 20 s, then extension at 72°C for 20 s, and a final step at 72°C for 8 min. PCR products (5 μl) were digested with five unit restriction enzymes following the manufacturer's instructions. Digested fragments were separated by electrophoresis on 3% agarose gel and identified by ethidium bromide staining. All genotyping was done blindly to the case–control status in Chinese National Human Genome Center, Beijing.

LD characterization, tag SNPs Selection and Prediction

The pattern of pairwise LD between the SNPs was measured by D′and r^2. Visualization of LD measures was performed using program HAPLOVIEW (http://www.broad.mit.edu/mpg/haploview). The KLKB1 haplotypes and their estimated frequencies were defined using the PLEM algorithm (Qin et al. 2002) implemented in the tagSNPs program (Stram et al. 2003).

Two programs were used to identify optimal subset of markers by different criteria. The program tagSNPs (Stram et al. 2003) was used to identify a set of tag SNPs that optimized the predictability of common haplotypes by use of the statistic R_h^2between tag SNPs haplotypes and common full haplotypes. This value can range from 0 to 1, with 1 indicating that tag SNPs haplotypes can perfectly predict all common SNPs haplotypes. We ran the program with the following criteria: common haplotypes were defined as 'the minimal set of haplotypes that covers 95% of existing haplotypes', and sets of tag SNPs resolving the common haplotypes were selected at a minimum R_h^2threshold of 0.85. The other program htSNP2, described by Chapman et al (2003), chooses an optimal set of tag SNPs that capture the remaining allelic variants. Predictive performance was assessed using a locus R^2 measure (R_L^2), which measures the proportion of variance of each remaining SNP "explained" by regression on the tag SNP alleles. The set of tag SNPs was selected to predict the remaining SNPs with a minimum $R_L^2$of 0.85. Exceptionally, all nonsynonymous SNPs were forced in as a set of tag SNPs. The program mlpop implemented in the package Genassoc was used to predict the association between essential hypertension and those remaining SNPs (Chapman et al. 2003).

Statistical analysis

The main purpose of our analyses was to test the association between tag SNPs and haplotype variation in the KLKB1 gene with essential hypertension. Analyses were done separately for each of the tag SNPs and followed by haplotype.

For each SNP, the genotypic and allelic frequencies between cases and controls were compared by the chi square (χ^2) test. Logistic regression analysis was used to assess whether the genetic variation was associated independently with hypertension after adjustment for covariates, including gender, age, smoking, alcohol drinking, body mass index (BMI), total cholesterol (TC), triglyceride (TG), creatinine (Cr), HDL cholesterol (HDL_C) and glucose (Glu). Hardy-Weinberg equilibrium was tested by Fisher's exact test. The statistical analysis was performed with STATA 8.0 for Windows. Adjustment was made for multiple testing using the Bonferroni correction by multiplying the nominal *P* value of each test by the number of SNPs or haplotypes. For the association of KLKB1 variation with hypertension stratified by alcohol drinking, the number of tests that was adjusted for doubled.

The haplo.score program (Schaid et al. 2002) was used to test for association of statistically inferred

haplotypes with hypertension. It is based on score statistics, which provide both global tests and haplotype-specific tests. The method models an individual's phenotype as a function of each inferred haplotype, weighted by their estimated probability, to account for haplotype ambiguity. Simulated *P* values were obtained from 1,000 replicates. Haplo.glm (Lake et al. 2003) 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. Both Haplo.score and Haplo.glm were implemented in the program Haplo.stats software developed using the R language (http://www.r-project.org). Only the haplotypes with frequencies > 1% were considered for the haplotype analyses.

Gene–environment interactions were tested in the multivariable-adjusted model, which included an extra interaction term, genotype/haplotype, and variables for exposure (alcohol drinking or smoking).

Power for individual SNPs was calculated using Quanto (http://www.hydra.usc.edu/gx), assuming an effective sample size of N/R_L^2 to adjust for the loss in power inherent in genotyping surrogate tag SNPs. Here N is the nominal sample size and R_L^2 is a sample size inflation factor.

Results

Clinical characteristics

The demographic and clinical characteristics of all individuals are shown in Table 2. Besides BP measurements, hypertensives exhibited significantly higher BMI, TG, Glu, TC, and low-density lipoprotein cholesterol (LDL_C) than controls. There were no significant differences between the cases and controls for gender, age, smoking and drinking status.

LD and tag SNPs selection

We used genotype data from the HapMap project. A total of 15 SNPs in the KLKB1 gene have been successfully genotyped in 45 unrelated Han Chinese in Beijing (CHB). Of these SNPs, eight SNPs were common (MAF ≥ 0.05) and spanned 31 kb of the gene, which yielded a mean density of one SNP per 4 kb. The rs number and relative position of these eight SNPs are shown in Fig. 1. No significant deviation from Hardy-Weinberg equilibrium was found for any polymorphism, after Bonferroni correction was applied. Thus, we used the eight SNPs to characterize the LD pattern and haplotype structure within KLKB1. Pairwise LD was measured by D′ among these eight SNPs, with value of 1 indicating the rarity of recombination between a pair of SNPs. It is worth noting that the LD decreased between rs2278542 and some other SNPs (Fig. 2). This could be due to its low frequency, for which test of LD based on 45 samples from HapMap might not have much power (Thompson et al. 1988). This assumption was confirmed by the fact that a strong LD between rs2278542 and rs925453 existed (D′ = 0.946, $P < 0.0001$) in our study subjects. We identified a total of ten haplotypes from the eight SNPs, five of which had frequencies > 1% and accounted for over 95% of all haplotypes in our population. This low haplotype diversity also suggested strong LD across KLKB1 without extensive recombination. Thus, we treated the whole gene as a single haplotype block to select tag SNPs.

Table 2 Comparison of clinical characteristics between cases and controls

	Cases (n = 1,317)	Controls (n = 1,269)	P value
Gender, m/f	655/662	658/611	NS
Age, years	54.15 ± 10.19	53.51 ± 9.49	NS
SBP, mmHg	159.70 ± 26.13	115.21 ± 10.72	<0.001
DBP, mmHg	95.85 ± 12.70	73.69 ± 7.81	<0.001
BMI, kg/m^2	25.99 ± 3.64	24.02 ± 3.44	<0.001
TC, mg/dl	199.21 ± 36.85	191.32 ± 39.40	<0.001
HDL-C, mg/dl	48.35 ± 11.37	49.97 ± 12.10	<0.001
LDL-C, mg/dl	122.85 ± 33.34	116.91 ± 34.38	<0.001
TG, mg/dl	142.92 ± 83.09	124.06 ± 78.93	<0.001
Glu, mg/dl	100.89 ± 22.17	96.98 ± 21.16	<0.001
Cr, umol/l	70.85 ± 14.08	70.18 ± 12.67	NS
Smokers, yes/no	527/790	546/723	NS
Drinkers, yes/no	418/891	385/876	NS

Mean ± SD values for continuous

SBP systolic blood pressure,*DBP* diastolic blood pressure, *BMI* body mass index, *TC* total cholesterol,*HDL-C* HDL cholesterol, *LDL-C* LDL cholesterol, *TG* triglyceride, *Glu* glucose, *Cr* creatinine, *Drinkers* the number of alcohol consumers who drank not less than 12 times during the year ahead of the interview, *Smokers* the number of cigarette consumers who had smoked not less than 100 cigarettes

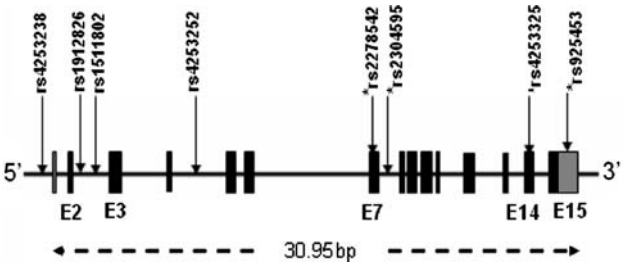

Fig. 1 The location of eight KLKB1 SNPs identified in HapMap database. The exons are indicated by *black boxes* and the 5′and 3′-UTR are denoted by *gray boxes*. *Tag SNPs

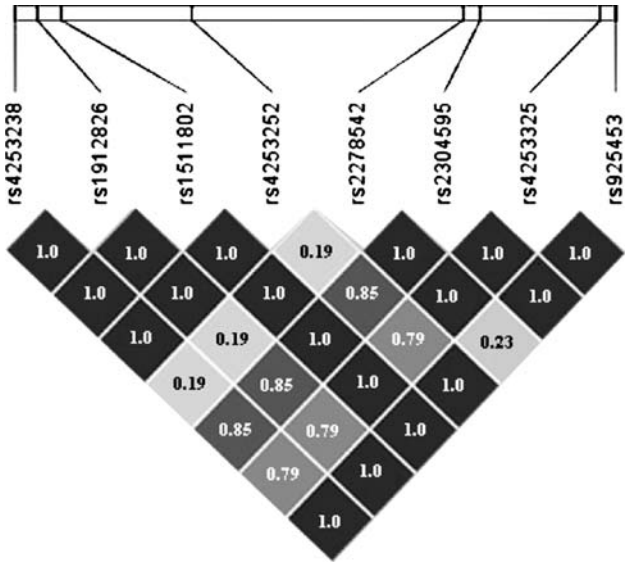

Fig. 2 Pair-wise LD among the eight KLKB1 SNPs in 45 Chinese in the HapMap project. LD strength between the eight SNPs is measured by the D′ statistic. The numbers in the boxes are D′ value

The same set of four tag SNPs, rs2278542, rs2304595, rs4253325, and rs925453, were selected by tagSNPs (Stram et al. 2003) and htSNP2 programs (Chapman et al. 2003). This set of tag SNPs could accurately predict the five common haplotypes with a minimum $R^2_h$of 0.85 and the four remaining common SNPs with a minimum $R^2_L$of 0.85 (Table 3).

Association analyses of single polymorphism

Genotyping success rate was 99.7%. Table 4 shows the genotype and allele frequencies of the four tag SNPs in the study group. Single SNP analyses indicated that SNPs rs2304595 and rs4253325 were significantly associated with hypertension. Results were also similar when adjusted for covariates in a logistic regression analysis. The frequency of the G allele and the prevalence of the GG genotype of rs2304595 polymorphism were significantly higher in hypertensive patients than in control subjects ($P = 0.022$ and 0.035, respectively). Compared with the rs4253325 GG homozygotes, rs4253325 A allele carriers had a significantly elevated risk of hypertension (adjusted OR = 1.20; 95%CI 1.02–1.41; $P = 0.026$), although the frequency of rs4253325 A allele was not statistically significantly higher in hypertensive patients than in control subjects ($P = 0.079$).

The approach we used to select an optimal set of tag SNPs (Chapman et al. 2003) can also be used to predict which of those SNPs not genotyped might also show association with the trait. Figure 3 shows the results based on prediction of single SNPs from the combination of the four tag SNPs. Significant association was predicted with SNP rs1511802 ($P = 0.031$).

Haplotype analysis

Five haplotypes of the tag SNPs had frequencies above 1% in controls, with a cumulative frequency above 99.4% (Table 5). The overall association between the haplotypes and disease status was not significant (global score statistic = 7.02, global $P = 0.219$). However, Hap2 (CAGC) was significantly underrepresented in cases and could be protective; the haplotype-specific score value was −2.35, with a P value of 0.019, which was in close agreement with the empirical P value of 0.015 obtained from simulation.

To find possible risk haplotypes and evaluate the effect of each haplotype, haplo.glm model was further performed, in which hap2 (CAGC), the protective one, was chosen as the baseline. Hap1 (AGAC) and Hap3 (CGAC) were found to significantly increase the risk

Table 3 8 KLKB1 SNPs identified in HapMap database

dbSNP	KLKB1 region	Alleles[a]	MAF[b]	Locus R^2_L	Position (NCBI B34)[c]	Intermarker distance (bp)
rs4253238	5′flanking	T/C	0.367	0.91	187844277	1,153
rs1912826	intron2	A/G	0.367	0.91	187845430	1,266
rs1511802	intron2	T/C	0.300	0.86	187846696	6,652
rs4253252	intron4	G/T	0.367	0.91	187853348	13,969
rs2278542	exon7	C/A	0.078	tag SNP	187867317	853
rs2304595	intron7	G/A	0.333	tag SNP	187868170	6,193
rs4253325	exon14	G/A	0.322	tag SNP	187874363	737
rs925453	exon15	C/T	0.156	tag SNP	187875100	–

SNPs are presented in chromosomal order and their location within the gene indicated

R^2_L was not calculated for SNPs with MAF < 0.05

[a] Major/minor allele

[b] Minor allele frequency

[c] Map positions on human chromosome 4 were from NCBI build 34

Table 4 Genotype distributions of KLKB1 tag SNPs in hypertension patients and controls

Tag SNPs	Alleles[a] (1/2)	Group	No. of genotypes (frequency)			MAF	Adjusted OR[b](95%CI, *P* value)	
			1/1	1/2	2/2		2 versus 1	1/2 + 2/2 versus 1/1
rs2278542	C/A	Control	1169(0.92)	98(0.08)	1(0.00)	0.04	1.23(0.93–1.64;0.152)	1.23(0.97–1.65,0.172)
		Case	1189(0.91)	115(0.09)	3(0.00)	0.05		
rs2304595	G/A	Control	528(0.42)	590(0.46)	151(0.12)	0.35	0.87(0.77–0.98;0.022)	0.84(0.71–0.99;0.035)
		Case	599(0.45)	584(0.44)	133(0.11)	0.33		
rs4253325	G/A	Control	640(0.50)	517(0.41)	111(0.09)	0.29	1.12(0.99–1.27,0.079)	1.20(1.02–1.41;0.026)
		Case	605(0.46)	589(0.45)	116(0.09)	0.31		
rs925453	C/T	Control	922(0.73)	331(0.26)	15(0.01)	0.14	1.01(0.86–1.19,0.874)	0.94(0.79–1.13;0.528)
		Case	961(0.73)	320(0.24)	30(0.02)	0.14		

MAF minor allele frequency, *OR* odds ratio, *95% CI* 95% confidence intervals

[a] The major allele was always referred to as allele 1 and the minor allele as allele 2

[b] OR estimated by logistic regression analysis, adjusted for Gender, age, smoking, alcohol drinking, BMI, TC, TG, Cr, HDL_C and Glu

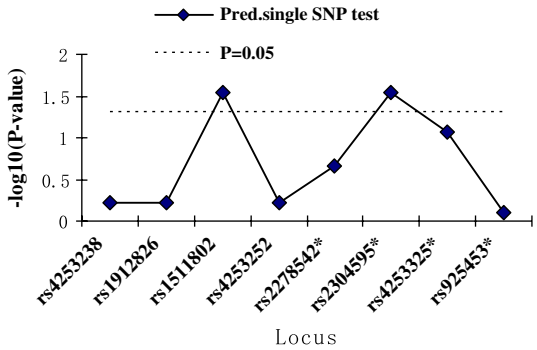

Fig. 3 Prediction of all association from tag SNPs. Associations of tag SNPs and predicted values for ungenotyped SNPs are indicated by −log10 (*P* value). Points above the *dotted line* represent SNP associations found or predicted to be significant at the $P = 0.05$ level. *Tag SNPs

of hypertension. Compared with the Hap2, the adjusted odds ratio was 1.37 (95% CI 1.29–1.45, $P < 0.0001$) for Hap1 and 1.17 (95% CI 1.02–1.34, $P = 0.028$) for Hap3, respectively. Hap1 remained significant after Bonferroni correction for multiple comparisons (five tests).

For the association analyses of BP, only control subjects were used because treatment of hypertension might have affected BP in case subjects. Haplotype analysis found no significant association of the haplotypes with BP (Table 6). However, there was a trend towards Hap1 being positively associated with BP variation. Hap1 achieved the highest positive score statistic in association tests with systolic BP (hap-score = 1.37, simulated $P = 0.167$) and diastolic BP (hap-score = 1.73, simulated $P = 0.067$).

SNP rs2278542, Hap1 and alcohol drinking

No significant difference was found between cases and controls for SNP rs2278542. The following interaction model indicated that there was a significant interaction between alcohol drinking and rs2278542 (interaction term $P = 0.013$).

We then separated the study population into drinkers and nondrinkers (Table 7). Only among drinkers was the frequency of rs2278542 A allele significantly higher in cases than in controls (6.14%, 2.99%; $P = 0.003$). This difference was still significant after

Table 5 Association between KLK1 gene haplotypes and essential hypertension

Variables	Haplotype[a]	Frequencies			OR (95% CI)[b]	*P* value[b]
		All	Cases	Controls		
Hap1	AGAC	0.040	0.044	0.037	1.37(1.29–1.45)	0.000
Hap2[c]	CAGC	0.335	0.320	0.350	Reference	–
Hap3	CGAC	0.261	0.268	0.252	1.17(1.02–1.34)	0.028
Hap4	CGGC	0.218	0.220	0.216	1.14(0.98–1.32)	0.089
Hap5	CGGT	0.140	0.142	0.139	1.12(0.94–1.33)	0.238

Haplotypes with frequencis < 0.05 were not included in the table. Hap1–Hap5 cover 99.4% of existing haplotypes

[a] Loci are arranged in the order rs2278542, rs2304595, rs4253325, rs925453

[b] Gender, age, smoking, alcohol drinking, BMI, TC, TG, Cr, HDL_C and Glu were adjusted

[c] Hap2 (CAGC) was chosen to be the baseline haplotype

Table 6 Association of KLKB1 gene haplotypes with blood pressure in 1,269 controls

Variables	Haplotype[a]	Frequency	SBP		DBP	
			Hap-score[b]	Sim *P* value	Hap-score[b]	Sim *P* value
Hap1	AGAC	0.037	1.37	0.167	1.73	0.067
Hap2	CAGC	0.350	0.60	0.537	−0.23	0.824
Hap3	CGAC	0.253	0.21	0.846	0.29	0.783
Hap4	CGGC	0.216	−1.38	0.164	−0.77	0.460
Hap5	CGGT	0.139	−0.51	0.581	−0.47	0.647

[a] Loci are arranged in the order rs2278542, rs2304595, rs4253325, rs925453

[b] The haplo.score function was used to infer the haplotype frequencies and to perform the score test statistics. Gender, age, smoking, alcohol drinking, BMI, TC, TG, Cr, HDL_C and Glu were adjusted

Table 7 Adjusted OR for hypertension associated with rs2278542 and Hap1, according to drinking status

		Drinkers (n = 803)		Nondrinkers (n = 1,767)	
		OR (95% CI)[a]	*P* value	OR (95% CI)[a]	*P* value
rs2278542	CC[b]	–	–	–	–
	CA + AA	2.22 (1.28–3.84)	0.005	0.94 (0.66–1.34)	0.745
Haplotype	Hap2[c]	–	–	–	–
	Hap1	2.50 (2.40–2.61)	0.000	1.04 (0.99–1.07)	0.404

[a] Age, gender, smoking, BMI, TC, TG, Cr, HDL_C and Glu were adjusted

[b] GG genotype was used as the reference

[c] Hap2 was used as the reference

Bonferroni correction ($P = 0.024$). With the rs2278542 CC genotype used as the reference, the adjusted OR for A allele carriers was 2.22 (95% CI, 1.28 to 3.84; $P = 0.005$).

As expected, a significant interaction between Hap1 (carrying rs2278542 A allele) and drinking status was also observed (interaction term $P < 0.0001$). In drinkers, a significant association was found between Hap1 and hypertension (Hap-score = 2.96, simulated $P = 0.003$). Increased risk of hypertension was observed comparing Hap1 to Hap2 (OR, 2.50; 95% CI, 2.40 to 2.61; $P < 0.0001$) (Table 7).

Discussion

In this study, we genotyped four tag SNPs of the KLKB1 gene and found that the KLKB1 gene variation was associated with hypertension. The Hap 1 (AGAC) and Hap3 (CGAC), which carry the susceptible rs2304595 G allele and rs4253325 A allele, were found to significantly increase the risk of essential hypertension. Among the drinkers, the association between Hap1 and hypertension became stronger.

We adopted two approaches, tagSNPs (Stram et al. 2003) and htSNP2 (Chapman et al. 2003), to select tag SNPs. The main difference is that the former is based on prediction of extended haplotypes (in this case, based on eight SNPs) from the marker haplotypes (in this case, based on four tag SNPs), whereas the latter is based on prediction of single SNP loci. In our study, both common haplotypes and remaining common SNPs in KLKB1 can be accurately predicted by four tag SNPs. However, comprehensive tagging will require a high genotyping density to identify all existing polymorphisms, and the 1 SNP per 4 kb density available for KLKB1 in the Phase I HapMap might be insufficient. To test the dependence of tag SNPs' performance on density, the predictive performance of the four tag SNPs was further applied to denser versions of the Phase II HapMap (public release up to Jan 2007) by calculating R^2_L in which a total of 31 SNPs (MAF $\geq$ 5%) within the KLKB1 gene were identified. The 27 remaining HapMap polymorphisms were tagged with a mean R^2_L of 0.82 by the four selected tag SNPs, with 20 out of 27 having a R^2_L greater than 0.85. For a SNP tagged with $R^2_L = 0.85$, we had 82% power at the 5% significance level to detect a dominant allele with a frequency of 0.15 that confers a relative risk of hypertension of 1.3. Thus, loss of power was marginal for these 20 SNPs. Interestingly, the other seven SNPs, which were poorly predicted (mean $R^2_L = 0.57$), had low MAFs (mean MAF = 0.11). Tag SNPs performance for SNPs with low MAFs was heavily dependent on the size of the LD sample, and there was a larger improvement in performance with increasing sample size

(Ahmadi et al. 2005). So we inferred that, for the seven SNPs, the predictive performance was limited by the size (45 individuals) of HapMap LD sample and that predictive power might be conservative.

Plasma kallikrein is composed of a heavy chain and a light chain held together by a disulphide bond. The light chain contains the active site or catalytic domain of the enzyme (Chung et al. 1986; Nunes et al. 2003). The rs4253325 polymorphism is located on the exon 14 which leads to the replacement of argnine with a glutamine at position 560 in the light chain of the protein. In the present study, rs4253325 A allele carriers had a higher risk of hypertension. Although the exact mechanism remains to be understood, it is not irrational to speculate that the decrease of KLKB1 enzymatic activity that results from structural or quantitative changes in the KLKB1 protein will reduce the release of bradykinin from kininogen. Obviously this hypothesis needs careful examination in future functional studies.

The tag SNPs were also used to predict associations of the remaining SNPs with hypertension. We predicted additional significant association with rs1511802, which was expected because of the almost perfect LD between rs1511802 in intron 2 and rs2304595 in intron 7 ($r^2 = 0.84$). The associations of the intronic SNPs are likely to result from LD with the true functional variants. SNP rs4253238 is located in the 5′ flanking region and might influence transcriptional activity of the gene. However, significant association could not be predicted. Currently, there are no validated KLKB1 functional promoter variants in public databases. The identification of functional promoter SNPs will greatly facilitate further investigation of KLKB1 in hypertension.

Another interesting finding of this study was that KLKB1 variation interacted with alcohol drinking. SNP rs2278542 and Hap1 were significantly associated with the risk of hypertension in the drinkers but not in the nondrinkers, which reflected a gene-environment interaction. Such an interaction is biologically plausible because alcohol drinking is a major risk factor for hypertension (Wildman et al. 2005). Earlier evidence of animal study suggested that there existed a possible link between plasma kallikrein and Ethanol. Ethanol could elicit an increase in the kallikrein activity and bradykinin-decomposed activity in plasma and markedly potentiate the vascular permeability accelerated by bradykinin (Tanaka and Yamashita 2002). These results warrant further studies regarding the functional effect of this SNP and its relationship with hypertension.

Haplotype analysis may provide more power to identify causal genetic variants. Due to the high degree of LD between SNPs within block, ancestral disease variants may be uncovered through evaluation of the underlying haplotypes (Gabriel et al. 2002; Patil et al. 2001). In the present study, Hap1 (AGAC) and Hap3 (CGAC) relative to the protective Hap2 were found to significantly increase the risk of hypertension. Comparing the elements of protective Hap2 and the two susceptible haplotype (Hap1 and Hap3), two polymorphisms (rs2304595 G allele and rs4253325 A allele) were regarded as the determinants modifying the risk of hypertension, which was consistent with the results of single SNP analyses. Of note, the rs2278542 was discriminator between Hap1 and Hap3, and with the replacement of Ala with Glu at rs2278542, the relative risk of individuals having hypertension was increased from 1.17 (Hap3) to 1.37 (Hap1) relative to Hap2. Furthermore, the risk for Hap1 reached an OR of 2.50 in the drinkers. Thus, Hap1 was more likely to confer a higher risk of hypertension. Hap1 might also contribute to BP variation, though the results of quantitative analyses were not statistically significant at the 5% level. Considering the fact that our quantitative trait association analyses were based on control population, further studies in random samples of general population will be needed to elucidate the possible effect of KLKB1 on BP variation.

When interpreting any genetic association study, several epidemiological limitations potentially leading to false–positive findings should be considered, including inadequate sample size, selection of control groups, multiple testing and population stratification. Our study was based on a large community-based sample; cases and controls were matched for age, gender and place of residence, and were self-identified as having four Han Chinese grandparents; our control allele frequencies were similar to those obtained from HapMap; ORs were adjusted for potential confounders; Bonferroni correction was overly conservative for the four SNPs in LD. These considerations argue against the existence of the aforementioned causes of false–positive results. Nevertheless, our observations are based on marginal *P* values (for Hap3) or small numbers (for Hap1) and might be interpreted with caution since this initial finding has not yet been confirmed in a separate population.

In conclusion, our results have consistently demonstrated significant association between KLKB1 single SNPs and haplotypes with hypertension in the northern Han Chinese population. The association appeared to be influenced by alcohol drinking. We hope this report will stimulate studies of the KLKB1 gene and hypertension that will potentially replicate our findings, either via association designs or through functional examinations.

Acknowledgments This work was supported by National Basic Research Program of China (Grant No. 2006CB503805), the Ministry of Science and Technology of The People's Republic of China (Grant No. 2006AA02Z170) and Beijing Natural Science Foundation (Grant No. 7061006).

References

Ahmadi KR, Weale ME, Xue ZY, Soranzo N, Yarnall DP, Briley JD, Maruyama Y, Kobayashi M, Wood NW, Spurr NK, Burns DK, Roses AD, Saunders AM, Goldstein DB (2005) A single-nucleotide polymorphism tagging set for human drug metabolism and transport. Nat Genet 37:84–89

Beaubien G, Rosinski-Chupin I, Mattei MG, Mbikay M, Chretien M, Seidah NG (1991) Gene structure and chromosomal localization of plasma kallikrein. Biochemistry 30:1628–1635

Carlson CS, Eberle MA, Rieder MJ, Yi Q, Kruglyak L, Nickerson DA (2004) Selecting a maximally informative set of single-nucleotide polymorphisms for association analyses using linkage disequilibrium. Am J Hum Genet 74:106–120

Chapman JM, Cooper JD, Todd JA, Clayton DG (2003) Detecting disease associations due to linkage disequilibrium using haplotype tags: a class of tests and the determinants of statistical power. Hum Hered 56:18–31

Chung DW, Fujikawa K, McMullen BA, Davie EW (1986) Human plasma prekallikrein, a zymogen to a serine protease that contains four tandem repeats. Biochemistry 25:2410–2417

Daly MJ, Rioux JD, Schaffner SF, Hudson TJ, Lander ES (2001) High-resolution haplotype structure in the human genome. Nat Genet 29:229–232

Dielis AW, Smid M, Spronk HM, Hamulyak K, Kroon AA, ten Cate H, de Leeuw PW (2005) The prothrombotic paradox of hypertension: role of the renin-angiotensin and kallikrein-kinin systems. Hypertension 46:1236–1242

Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, Higgins J, DeFelice M, Lochner A, Faggart M, Liu-Cordero SN, Rotimi C, Adeyemo A, Cooper R, Ward R, Lander ES, Daly MJ, Altshuler D (2002) The structure of haplotype blocks in the human genome. Science 296:2225–2229

Gu D, Reynolds K, Wu X, Chen J, Duan X, Muntner P, Huang G, Reynolds RF, Su S, Whelton PK, He J (2002) Prevalence, awareness, treatment, and control of hypertension in china. Hypertension 40:920–927

Gu D, Reynolds K, Wu X, Chen J, Duan X, Reynolds RF, Whelton PK, He J (2005) Prevalence of the metabolic syndrome and overweight among adults in China. Lancet 365:1398–1405

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 (2001) Haplotype tagging for the identification of common disease genes. Nat Genet 29:233–237

Lake SL, Lyon H, Tantisira K, Silverman EK, Weiss ST, Laird NM, Schaid DJ (2003) Estimation and tests of haplotype-environment interaction when linkage phase is ambiguous. Hum Hered 55:56–65

Marcondes S, Antunes E (2005) The plasma and tissue kininogen-kallikrein-kinin system: role in the cardiovascular system. Curr Med Chem Cardiovasc Hematol Agents 3:33–44

Nunes VA, Gozzo AJ, Sampaio MU, Juliano MA, Sampaio CA, Araujo MS (2003) Mapping of human plasma kallikrein active site by design of peptides based on modifications of a Kazal-type inhibitor reactive site. J Protein Chem 22:533–541

Patil N, Berno AJ, Hinds DA, Barrett WA, Doshi JM, Hacker CR, Kautzer CR, Lee DH, Marjoribanks C, McDonough DP, Nguyen BT, Norris MC, Sheehan JB, Shen N, Stern D, Stokowski RP, Thomas DJ, Trulson MO, Vyas KR, Frazer KA, Fodor SP, Cox DR (2001) Blocks of limited haplotype diversity revealed by high-resolution scanning of human chromosome 21. Science 294:1719–1723

Perloff D, Grim C, Flack J, Frohlich ED, Hill M, McDonald M, Morgenstern BZ (1993) Human blood pressure determination by sphygmomanometry. Circulation 88:2460–2470

Qin ZS, Niu T, Liu JS (2002) Partition-ligation-expectation-maximization algorithm for haplotype inference with single-nucleotide polymorphisms. Am J Hum Genet 71:1242–1247

Schaid DJ, Rowland CM, Tines DE, Jacobson RM, Poland GA (2002) Score tests for association between traits and haplotypes when linkage phase is ambiguous. Am J Hum Genet 70:425–434

Schmaier AH (2003) The kallikrein-kinin and the renin-angiotensin systems have a multilayered interaction. Am J Physiol Regul Integr Comp Physiol 285:R1–13

Stram DO, Haiman CA, Hirschhorn JN, Altshuler D, Kolonel LN, Henderson BE, Pike MC (2003) 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 55:27–36

Tanaka Y, Yamashita Y (2002) Effects of ethanol administration at a high-dose level on the stimulatory action by bradykinin in vascular permeability. J Nutr Sci Vitaminol (Tokyo) 48:270–277

Tang J, Yu CL, Williams SR, Springman E, Jeffery D, Sprengeler PA, Estevez A, Sampang J, Shrader W, Spencer J, Young W, McGrath M, Katz BA (2005) Expression, crystallization, and three-dimensional structure of the catalytic domain of human plasma kallikrein. J Biol Chem 280:41077–41089

The International HapMap Project (2003) Nature 426:789–796

Thompson EA, Deeb S, Walker D, Motulsky AG (1988) The detection of linkage disequilibrium between closely linked markers: RFLPs at the AI-CIII apolipoprotein genes. Am J Hum Genet 42:113–124

Ward R (1990) Familial aggregation and genetic epidemiology of blood pressure. In: Laragh JH, Brenner BM (eds) Hypertension: pathophysiology, diagnosis, and management. Raven Press, New York, pp 81–100

Weale ME, Depondt C, Macdonald SJ, Smith A, Lai PS, Shorvon SD, Wood NW, Goldstein DB (2003) Selection and evaluation of tagging SNPs in the neuronal-sodium-channel gene SCN1A: implications for linkage-disequilibrium gene mapping. Am J Hum Genet 73:551–565

Wildman RP, Gu D, Muntner P, Huang G, Chen J, Duan X, He J (2005) Alcohol intake and hypertension subtypes in Chinese men. J Hypertens 23:737–743

Yu H, Anderson PJ, Freedman BI, Rich SS, Bowden DW (2000) Genomic structure of the human plasma prekallikrein gene, identification of allelic variants, and analysis in end-stage renal disease. Genomics 69:225–234

Yu H, Bowden DW, Spray BJ, Rich SS, Freedman BI (1998) Identification of human plasma kallikrein gene polymorphisms and evaluation of their role in end-stage renal disease. Hypertension 31:906–911