Published in final edited form as:

Pharmacogenet Genomics. 2007 April ; 17(4): 267–275. doi:10.1097/FPC.0b013e3280105235.

Single nucleotide polymorphism discovery and haplotype analysis of Ca²⁺-dependent K⁺ channel beta-1 subunit

Yan Gong^{a,b}, Amber L. Beitelshes^{a,b}, Jennifer Wessel^d, Taimour Y. Langaee^{a,b}, Nicholas J. Schork^d, and Julie A. Johnson^{a,b,c}

^aCenter for Pharmacogenomics, College of Pharmacy, University of Florida, Gainesville, Florida

^bDepartment of Pharmacy Practice, College of Pharmacy, University of Florida, Gainesville, Florida

^cDepartment of Medicine, Division of Cardiology, University of Florida, Gainesville, Florida

^dDepartment of Psychiatry, University of California at San Diego, La Jolla, California, USA

Abstract

The large-conductance, Ca²⁺-dependent K⁺ channel plays a key role in the control of vascular tone. Variation in the gene encoding the β -1 subunit of the Ca²⁺-dependent K⁺ channel (*KCNMB1*) has been reported to be associated with hypertension, however, variants in *KCNMB1* have not been systematically characterized to date. In this study, we have performed the most comprehensive evaluation to date of single nucleotide polymorphisms in *KCNMB1* using genomic DNA from 60 individuals of European, African and native American ancestry. We identified and characterized single nucleotide polymorphisms in the exons, intron/exon junctions, upstream region and 3' untranslated regions of *KCNMB1* using denaturing high-performance liquid chromatography combined with direct DNA sequencing. A total of 25 single nucleotide polymorphisms in *KCNMB1* were identified. Seven of the polymorphisms (28%) are novel single nucleotide polymorphisms not reported previously. Allele frequencies range from less than 1.7 to 50% and 19 single nucleotide polymorphisms had a minor allele frequency greater than 5%. A lack of strong linkage disequilibrium among the 25 single nucleotide polymorphisms was observed in all three race/ethnicity groups; therefore the identification of haplotype 'tag' single nucleotide polymorphisms for genetic association studies is not likely to be appropriate for *KCNMB1*. Multiple species comparative analysis and in-silico functional analysis were performed to identify potential functionally important single nucleotide polymorphisms within the gene. These data highlight that a tag single nucleotide polymorphism approach will not be appropriate for the study of genes such as *KCNMB1*, although potentially important functionally significant single nucleotide polymorphisms are suggested for future studies investigating the influence of this gene's variability on disease and drug response.

Keywords

Ca²⁺-dependent K⁺ channel; haplotype; *KCNMB1*; single nucleotide polymorphisms

Introduction

Hypertension is the most common chronic disease in the Western world, although there are many drug classes from which to choose therapy, only 34% of North Americans with hypertension currently have their blood pressure controlled [1]. The inability to determine the most appropriate drug therapy for a hypertensive patient *a priori* likely contributes to this poor blood pressure control. Pharmacogenetics research has the potential to significantly improve blood pressure control and outcomes.

Currently, there are limited data on genetic associations with response to one of the most prescribed classes of antihypertensive drugs, calcium channel blockers. The Ca^{2+} -dependent K^+ (BK) channel (also referred to as MaxiK channels) is a large conductance, voltage and calcium-sensitive potassium channel that is fundamental to the control of smooth muscle tone and neuronal excitability. Increased arterial tone is a hallmark of elevated blood pressure, and pressure-induced depolarization of vascular smooth muscle cells contributes to increased vascular tone by increasing Ca^{2+} influx through BK channels [2,3]. Our long-term goal is to study the association between the variants in the BK channel and blood pressure response to calcium channel blocker therapy.

In-vitro evidence suggests the BK channel limits Ca^{2+} entry and thereby arterial contraction by repolarizing smooth muscle cells and closing voltage-dependent Ca^{2+} channels opened previously by pressure or vasoconstrictors [4]. BK channels also mediate steady hyperpolarization and vasorelaxation as a result of transient outward currents carried by BK channels spontaneously activated by local release of Ca^{2+} from intracellular stores via ryanodine receptor channels [5].

BK channels can be formed by two subunits: the pore-forming α subunit and the modulatory β -1 subunit. The β -1 subunit, which is restricted to smooth muscle, maintains the normal high voltage and Ca^{2+} sensitivity of the pore-forming α subunit [6]. The role of the BK channel auxiliary β -1 subunit in blood pressure regulation has been tested previously by deleting its gene in mice, resulting in impaired coupling between Ca^{2+} release and activation of hyperpolarizing BK currents, and leading to systemic hypertension [7]. The β -1 subunit of the BK channel is encoded by *KCNMB1* gene (gene ID: 3779; HGNC: 6285, MIM: 603951), which is located on chromosome 5q34. *KCNMB1* has a sequence length of 11 472 bp, and contains four exons and three introns with a total coding region of 1517 bp.

In a previously reported study, in which a nonsynonymous *KCNMB1* single nucleotide polymorphism (SNP) E65K was genotyped in hypertensive and normotensive individuals, this polymorphism was found to be associated with low prevalence of moderate and severe diastolic hypertension [8]. In a series of in-vitro functional studies, this polymorphism showed a gain of function consistent with a more efficient feedback mechanism for the control of vascular tone, compatible with a protective effect of this polymorphism against severity of diastolic hypertension [8]. A recent epidemiological study by the same group has demonstrated the protective effect of this polymorphism against diastolic hypertension. Compared with EE homozygotes, K-carriers had a reduced risk of moderate-to-severe diastolic hypertension (odds ratio 0.35; $P = 0.006$). A 5-year follow-up study also demonstrated a reduced risk [hazard ratio of 0.11, 95% confidence interval (0.01, 0.79)] of K-carriers for adverse cardiovascular events (myocardial infarction and stroke) compared with EE homozygotes after adjusting for age and sex [9]. BK channels are also involved in calcium signaling and act in a negative feedback manner on L-type calcium channels, the binding site of the calcium channel blockers used in the treatment of hypertension and angina.

To determine the potential clinical importance of this gene, more comprehensive studies of its variations are essential. Thus, we resequenced the coding region, intron/exon boundaries and

presumed regulatory regions of the gene in 60 ethnically diverse human participants. We then compared SNP selection by haplotype tagging approaches, examination of conserved regions and the use of computational methods for identification of SNPs likely to have functional significance.

Methods

Samples

We included 60 individuals in this analysis, which provided > 99% power to detect polymorphisms of > 5% frequency, as well as reasonable power to detect polymorphisms with frequencies as low as 1%.

Genomic DNA was obtained from lymphoblastoid cell lines selected from the Human Genetic Cell Repository at the Coriell Institute, Camden, New Jersey [10]. These included 20 Caucasians (W1-W20, Coriell Cell Repository DNA: NA05377, NA05380, NA05392, NA05398, NA05408, NA05422, NA05423, NA06051, NA06859, NA06861, NA00130, NA00536, NA00546, NA00558, NA00607, NA00621, NA00893, NA00946, NA01310, NA01805, respectively); 20 African-Americans (AA1-AA20, from Human variation panel of African-American Panel of 50, Coriell Cell Repository DNA: NA17101-NA17120) and 20 native Americans [NA1-NA20, from Human variation panels: South American (Andes) panel and South American (Brazil, Guyana and Venezuela) panel: NA17301-NA17320]. We also resequenced *KCNMB1* from a chimpanzee (also obtained from Coriell) to assist in assessment of phylogenetic patterns associated with the discovered SNPs.

Single nucleotide polymorphism discovery and genotyping

Regions including 5' and 3' untranslated regions (2 kb flanking), exons and intron/exon junctions (approximately 100 bp into intron) of *KCNMB1* were amplified in fragments of approximately 200–500 bp, which was the optimal size for our screening method. A denaturing high-performance liquid chromatograph (DHPLC) system (WAVE, Transgenomics Limited, Crewe, UK) was used for the initial mutation detection [11]. The Navigator software (version 1.5.4, Transgenomic WAVE system, Omaha, Nebraska, USA) was used for data handling, optimizing gradients and choosing temperatures for the runs. For each amplicon, a reference sample was selected on the basis of its failure to form a heteroduplex (indicating a lack of variation in that amplicon). Polymerase chain reaction (PCR) product from the chosen reference sample was then pooled with each of the other 59 individual sample PCR products. This pooled DNA was partially denatured at 95°C, reannealed over 30 min and run at the predetermined temperatures under partially denaturing conditions. The reference sample, samples forming heteroduplexes with the reference sample, those with questionable DHPLC chromatograms and several samples failing to form heteroduplexes (presumed major allele homozygotes) were sent for direct sequencing, which was accomplished using Amersham Biosciences ET-terminator chemistry at the University of Florida Interdisciplinary Center for Biotechnology Research. Those samples that did not form a heteroduplex with the reference sample were assigned the same genotype as the sequencing-determined genotype of the reference samples. DNA sequence data were compiled, and polymorphic sites were identified using PolyPhred (University of Washington, Seattle, Washington, USA). Genotypes were called if the readings from forward and reverse PCR reactions were consistent.

Owing to incomplete sequence data in the sample of 60 participants, two SNPs (SNP8 G/A at -691 and SNP9 T/C at -338) in the upstream region of *KCNMB1* were genotyped by Pyrosequencing [12] using a PSQHS96A system and SNP reagent kit according to the manufacturer's protocol (Biotage AB, Uppsala, Sweden). The following primers were used. SNP8: (5'-biotin-GACCA CAGGCATGCACCACTA-3' (PCR-forward), 5'-GATT

GCTCAAGCCCAGGAGTTT-3' (PCR-reverse) and sequencing primer: 5'-CATCCTGGGCAACAT-3' (reverse). SNP9: 5'-GCTTCAGCTCAGTTGGTCCTTAAA-3' (PCR-forward); 5'-biotin-GGGTAGTGATGGGAGTTAGAGACG-3' (PCR-reverse); sequencing primer (forward): 5'-AGT CAGTTCTCTCCCCT. PCR mixture (12.5 µl) consisted of 6.25 µl HotStarTaq Master Mix Kit (Qiagen Inc., Valencia, California, USA), 1 µl PCR primers (10pmol/µl), 0.75 ml of dimethyl sulfoxide, 1.25 µl of water and 40 ng of DNA. PCR was performed under the following conditions: 95°C for 15 min; 40 cycles consisting of denaturation at 95°C for 30s, annealing at 62°C for 40 s and extension at 72°C for 30s and final extension for 7 min. Pyrosequencing was performed under standard conditions for sequence determination and allele designation and data were captured with PSQ HS96A SNP software (Biotype AB, Uppsala, Sweden).

Computational methods and statistical analysis

Comparative analysis of *KCNMB1* sequences from multiple species (rat and mouse versus human) was performed with the VISTA genome browser tools (<http://pipeline.lbl.gov/cgi-bin/gateway2>) [13–15], which has been shown to be effective in identifying functionally significant variations in genes [16,17]. The human [Human May 2004, University of California Santa Cruz (UCSC) hg17], mouse (May 2004, UCSC mm5) and rat (June 2003, UCSC rn3) genome sequence were obtained from the UCSC Genome Browser. VISTA plots present comparative data in the form of a curve to display the level of sequence conservation (from 50 to 100%) in a predefined window of a global alignment.

Systematic studies of common genetic variants are facilitated by linkage disequilibrium (LD) [18]. The observed structure of LD in any particular genomic interval depends on a complex interplay of demographic history, stochastic events and functional constraints [19]. We used both D' and r^2 methods for measuring LD between pairs of SNPs. Haploview software (Broad Institute of MIT, Cambridge, MA, USA) [20] was used to analyze and visualize the pairwise LD between each common SNP in *KCNMB1* gene.

Haplotype assignment was performed using HAP software (California Institute for Telecommunications and Information Technology and National Biomedical Computation Resource, University of California, San Diego, California, USA) [21]. Using SNP-EM software [22] (<http://polymorphism.ucsd.edu/cgi-bin/PRL/snpem/snpem.cgi>), likelihood calculations were performed to estimate the haplotype frequencies for each race/ethnicity group as well as the combined sample. An omnibus likelihood ratio statistic was then computed by taking the ratio of the likelihoods to test the hypothesis of equal haplotype frequencies between the racial groups. The null distribution of the likelihood ratio statistic was then approximated via randomization tests, in which the race/ethnicity indicators were randomly permuted among the individuals in the sample and the likelihood ratio statistics recomputed.

One major evolutionary mechanism that causes changes in SNP frequencies is natural selection, which occurs when one allele confers higher average fitness on its bearers. Alleles, however, may also change in their relative frequencies because of random drift or chance. To test the hypothesis that there was no evidence for selection of the genetic variations, intrapopulation selective neutrality tests, including Tajima's D [23] and Fu's F_s [24], were performed. These tests were complemented by analysis of molecular variance framework tests [25]. In addition, pairwise F_{ST} statistics were computed between the three populations as a relative measure of population differentiation [26]. The significance of the genetic distances is tested by permutating the haplotypes between the populations [26]. All population genetics analyses were performed in ARLEQUIN population genetics software (version 4.0; Genetics and Biometry Laboratory, University of Geneva, Geneva, Switzerland) [27]. Minimum spanning network among haplotypes was also computed and plugged in software TreeView

(University of Glasgow, Scotland, UK) [28] to construct the phylogenetic tree of the haplotypes. The Chimp was used as the outgroup for the relevant calculations.

Polymorphism Database Mining, Annotation Programs (PolyMAPr, Washington University, School of Medicine, St. Louis, Missouri, USA) [29] was used to predict the functional effects of nonsynonymous coding-region SNPs (PolyPhen, European Molecular Biology Laboratory, Heidelberg, Germany) and any variants that might alter exon splicing enhancer sites (ESEfinder, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, USA), putative transcription factor binding sites (JASPAR, Karolinska Institute, Stockholm, Sweden) or intron–exon splice sites (ASD: alternative splicing database, splice junction sequences). Another web tool, PuPaSNP Finder (<http://pupasnp.bioinfo.ocha.fib.es/>) was also used to search for SNPs with potential phenotypic effect [30].

Results

The genomic sequence of *KCNMB1* is depicted visually in Fig. 1b. Using DHPLC in combination with big-dye terminator chemistry-based direct sequencing, a total of 25 SNPs were identified in the *KCNMB1* gene. The detailed information on the variants, including chromosomal location and estimated minor allelic frequencies in three race/ethnicity groups are listed in Table 1. Seventeen of the 25 SNPs are purine–purine or pyrimidine–pyrimidine transitions (10 A↔G and 7 C↔T) and eight are purine↔pyrimidine transversions. Eighteen of the SNPs (72%) have been reported previously in the NCBI SNP database, whereas seven (28%) were novel. Eleven of the previously reported SNPs had not yet been validated by multiple independent submissions and five had no frequency information. Ten of 25 SNPs occurred within the 5' and promoter region. Of the other 15 SNPs, six occurred in exons, five in introns and six in the 3' untranslated region (Table 1). Among the novel SNPs, four were common (MAF > 5%) in at least one population, although three of these four were not detected in Caucasians. Among the three coding-region polymorphisms, two encode amino acid substitutions (SNP19: Glu65Lys and SNP21: Val110Leu) (Fig. 1d). The third is synonymous (SNP15: Cys18Cys) (Table 1). Detailed genotype data for each of the 60 individuals is available online at The Pharmacogenetics and Pharmacogenomics Knowledge Base (PharmGKB) website (<http://www.pharmgkb.org/>).

Departure from Hardy–Weinberg equilibrium was tested for each SNP in each of the three different race/ethnicity groups using χ^2 test with one degree of freedom. Among the 64 tests that were run, only SNP6 ($\chi^2 = 5.915$, $P = 0.015$) and SNP12 ($\chi^2 = 8.183$, $P = 0.004$) in Caucasians and SNP7 ($\chi^2 = 3.874$, $P = 0.049$) in African Americans deviated from Hardy–Weinberg equilibrium. This represents the approximate percentage of P values < 0.05 that would be expected to occur by chance with this number of tests.

The overall minor allele frequencies (MAF) of the *KCNMB1* SNPs ranged from 0.017 to 0.50. The location and MAF of common SNPs (MAF > 0.05) are shown in Fig. 1c. Eighteen SNPs were polymorphic in each of the three population samples genotyped in our study. The majority of these SNPs have a MAF greater than 10% in each population. Although the sample sizes in our study preclude any definite conclusion regarding the absence of a particular allele in any given population, we observed three SNPs (SNP4, 17 and 24) only in African Americans, but not in Caucasians or native Americans and SNP10 only in native Americans. A schematic representation of population-specific SNPs is shown in Fig. 2. Although population-specific SNPs tend to have lower MAFs than other SNPs, 50% of the population-specific SNPs reported here have MAF $\geq$ 0.05 in the specific population.

The VISTA plot of the multiple species comparative analysis of *KCNMB1* sequence from rat and mouse versus human is shown in Fig. 1a. Comparison of the human, rat and mouse genomic

sequences of the *KCNMB1* locus showed certain strongly conserved (> 75%) regions of the protein-coding (in blue) and 5' untranslated region (in light green) (Fig. 1a). The SNPs that occurred in the highly conserved regions were the following three coding-region polymorphisms: SNP15 (Cys18Cys, novel), SNP19 (Glu65Lys, rs11739136) and SNP21 (Val110Leu, rs2301149), three 5' untranslated region SNPs: SNP10 (novel), SNP11 (rs827778) and SNP12 (novel); and three SNPs in the 3' untranslated regions: SNP22 (rs2656842), SNP23 (rs2656841) and SNP25 (rs314156).

Pairwise LD (r^2) of the 25 SNPs in *KCNMB1* for the three race/ethnicity groups obtained from Haploview software are shown in Fig. 3. A relative lack of strong LD was observed for all three groups, especially in African-Americans. Similar results were obtained when D' was used as a measure of LD in all three populations (data not shown).

We inferred haplotypes in 32 individuals for whom there were no missing genotypes, using the 19 most common SNPs ($MAF > 0.05$), through use of HAP software [21]. A total of 50 haplotypes were derived from the 19 SNPs. The phylogeny of these haplotypes from three populations and the chimp (our chosen outgroup) is shown in Fig. 4. The large number of haplotypes is consistent with the relatively low LD between most SNPs in this gene.

The omnibus likelihood ratio test to assess the equality of haplotype frequencies between different race/ethnicity groups revealed differences between African-American and native Americans ($P = 0.004$), and African-Americans versus Caucasians ($P = 0.032$), but not between Caucasians and native Americans ($P = 0.389$). The significant difference between Caucasians and African-Americans is mostly driven by the haplotype 13. The difference between African-American and native American is driven by several unique haplotypes in either population.

Results for the testing for departure of the *KCNMB1* variants from selective neutrality are shown in Table 2. The Tajima's D_s are positive for all three populations, but none reached the 0.05 significance level for their derived P values. The Fu's F_s are negative and significant for all three populations, especially in African-Americans, with a derived P value of <0.0001. This test statistic was of borderline significance in Caucasians ($P = 0.011$; relative to the adjusted P of 0.02 corresponding to α level of 0.05).

In-silico tests of functional significance of common SNPs of *KCNMB1* using PolyMAPr predicted the two nonsynonymous SNPs: SNP19 (Glu65Lys, rs11739136) and SNP21 (Val110Leu, rs2301149) are likely to be benign. SNP1 (rs314127), SNP3 (rs314129) and SNP14 (rs2075612) are located in putative transcription factor-binding sites FREAC-7, E47-Hand1 and a C2H2 zinc finger MZF1-4, respectively (Fig. 1e). PupaSNP finder program [30] also suggested that SNP3 overlapped with the domain of the basic helix-loop-helix transcriptional factor gene E47-Hand1. SNP12 and nonsynonymous SNP 19 (Glu65Lys, rs 11739136) overlap with exon splicing enhancer (ESE) for serine/arginine-rich (SR) protein SF2. SNP13 and SNP18 are found at an ESE site for SR protein SRp40 (Fig. 1f).

The above analyses were conducted to identify the most appropriate SNPs for future clinical association studies. A lack of strong LD among the 25 SNPs suggests that a haplotype tagging SNP approach would add little efficiency in the study of the *KCNMB1* gene. Therefore, on the basis of considerations of allele frequency ($MAF > 0.05$), LD, conservation and potential functional importance, we suggest a list of eight SNPs for future genetic association studies (shown in Table 3).

Discussion

Our study provides the most comprehensive analysis to date of SNPs and haplotype structure in the *KCNMB1* gene in three ethnically diverse populations. We chose panels of Caucasians, African-Americans and native Americans on the basis of the fact that our patient population of interest (INVEST-GENES) [31] contains a large percentage of Hispanics, which previous studies have suggested have a high degree of African, European and native American ancestry [32].

A total of 25 variants were detected in *KCNMB1*, 28% of which have not been reported previously. Additionally, our own genotyping data suggest that 10 previously unvalidated SNPs are, in fact, valid SNPs, and we provide allele frequencies in multiple populations for them. The novel SNPs were found mostly in African-Americans and native Americans.

It is well accepted that the pattern of LD across the human genome is not uniform, with regions of strong LD, of varying size, often referred to as 'LD blocks', and regions of relatively low LD. In regions of high LD, a reduced set of haplotype tag SNPs may be selected to detect efficient associations between variations in that gene or region and a trait of interest. Our data suggest that *KCNMB1* resides in a region of relatively low LD in all three populations we studied. The fact that there are 50 different haplotypes from 19 common SNPs confirmed our observed lack of strong LD in the gene. Therefore, a haplotype tag SNP approach for selection of SNPs for study would not have the desired efficiencies for this gene. Thus, selection of SNPs for study of this and other genes that lie in regions of low LD must rely on different approaches, such as conservation and predicted functional effects. These data highlight that the increasing focus on a HapMap-based tag SNP approach will not be relevant or appropriate for all genes of interest. Thus, for genes that lie in regions of low LD, other approaches for SNP selection must be taken, and based on these other approaches, we propose eight SNPs that investigators should consider for the study of *KCNMB1*. One limitation of this study is that we did not sequence the entire intronic sequence. The low LD in this gene, however, and the fact that relatively few functional SNPs lie deep in intronic regions, reduces the likelihood that we have failed to identify functionally important SNPs in this gene.

One question of interest is the pattern of polymorphism we observed is consistent with a neutral model of evolution. We used two neutrality tests in our study: Tajima's *D* and Fu's *F_s*, both assuming an infinite-site model. Fu's *F_s* test has been shown to be especially sensitive to population demographic expansion, which generally leads to large negative *F_s* values [24]. In our neutrality test result, Fu's *F_s* tests showed significant negative values for all three populations, providing evidence for the effects of selection on this gene, especially in African-Americans. We note that this is consistent with the interpopulation allele frequency and haplotype frequency differences that we observed, as increased interpopulation differences are indications of selection pressures [33,34].

Acknowledgement

This study was funded by NIH grant HL74730. There are no conflicts of interests.

References

1. Johnson JA, Turner ST. Hypertension pharmacogenomics: current status and future directions. *Curr Opin Mol Ther* 2005;7:218–225. [PubMed: 15977418]
2. Harder DR, Smeda J, Lombard J. Enhanced myogenic depolarization in hypertensive cerebral arterial muscle. *Circ Res* 1985;57:319–322. [PubMed: 4017200]

3. Wellman GC, Cartin L, Eckman DM, Stevenson AS, Saundry CM, Lederer WJ, Nelson MT. Membrane depolarization, elevated Ca^{2+} entry, and gene expression in cerebral arteries of hypertensive rats. *Am J Physiol Heart Circ Physiol* 2001;281:H2559–H2567. [PubMed: 11709423]
4. Brayden JE, Nelson MT. Regulation of arterial tone by activation of calcium-dependent potassium channels. *Science* 1992;256:532–535. [PubMed: 1373909]
5. Mark DB, Hlatky MA, Califf RM, Naylor CD, Lee KL, Armstrong PW, et al. Cost effectiveness of thrombolytic therapy with tissue plasminogen activator as compared with streptokinase for acute myocardial infarction. *N Engl J Med* 1995;332:1418–1424. [PubMed: 7723799]
6. Wallner M, Meera P, Ottolia M, Kaczorowski GJ, Latorre R, Garcia ML, et al. Characterization of and modulation by a beta-subunit of a human maxi K Ca channel cloned from myometrium. *Receptors Channels* 1995;3:185–199. [PubMed: 8821792]
7. Shklyayev S, Aslanidi G, Tennant M, Prima V, Kohlbrenner E, Kroutov V, et al. Sustained peripheral expression of transgene adiponectin offsets the development of diet-induced obesity in rats. *Proc Natl Acad Sci U S A* 2003;100:14217–14222. [PubMed: 14617771]
8. Fernandez-Fernandez JM, Tomas M, Vazquez E, Orio P, Latorre R, Senti M, et al. Gain-of-function mutation in the KCNMB1 potassium channel subunit is associated with low prevalence of diastolic hypertension. *J Clin Invest* 2004;113:1032–1039. [PubMed: 15057310]
9. Senti M, Fernandez-Fernandez JM, Tomas M, Vazquez E, Elosua R, Marrugat J, Valverde MA. Protective effect of the KCNMB1 E65K genetic polymorphism against diastolic hypertension in aging women and its relevance to cardiovascular risk. *Circ Res*. 2005
10. Collins FS, Brooks LD, Chakravarti A. A DNA polymorphism discovery resource for research on human genetic variation. *Genome Res* 1998;8:1229–1231. [PubMed: 9872978]
11. Crepin M, Pigny P, Escande F, Bauters CC, Calender A, Lefevre S, et al. Evaluation of denaturing high performance liquid chromatography for the mutational analysis of the MEN1 gene. *J Mol Endocrinol* 2006;36:369–376. [PubMed: 16595707]
12. Langaee T, Ronaghi M. Genetic variation analyses by Pyrosequencing. *Mutat Res* 2005;573:96–102. [PubMed: 15829240]
13. Frazer KA, Pachter L, Poliakov A, Rubin EM, Dubchak I. VISTA: computational tools for comparative genomics. *Nucleic Acids Res* 2004;32:W273–W279. [PubMed: 15215394]
14. Mayor C, Brudno M, Schwartz JR, Poliakov A, Rubin EM, Frazer KA, et al. VISTA: visualizing global DNA sequence alignments of arbitrary length. *Bioinformatics* 2000;16:1046–1047. [PubMed: 11159318]
15. Brudno M, Do CB, Cooper GM, Kim MF, Davydov E, Green ED, et al. LAGAN and Multi-LAGAN: efficient tools for large-scale multiple alignment of genomic DNA. *Genome Res* 2003;13:721–731. [PubMed: 12654723]
16. Boffelli D, McAuliffe J, Ovcharenko D, Lewis KD, Ovcharenko I, Pachter L, et al. Phylogenetic shadowing of primate sequences to find functional regions of the human genome. *Science* 2003;299:1391–1394. [PubMed: 12610304]
17. Wen G, Mahata SK, Cadman P, Mahata M, Ghosh S, Mahapatra NR, et al. Both rare and common polymorphisms contribute functional variation at CHGA, a regulator of catecholamine physiology. *Am J Hum Genet* 2004;74:197–207. [PubMed: 14740315]
18. Altshuler D. A haplotype map of the human genome. *Nature* 2005;437:1299–1320. [PubMed: 16255080]
19. Hinds DA, Stuve LL, Nilsen GB, Halperin E, Eskin E, Ballinger DG, et al. Whole-genome patterns of common DNA variation in three human populations. *Science* 2005;307:1072–1079. [PubMed: 15718463]
20. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005;21:263–265. [PubMed: 15297300]
21. Halperin E, Eskin E. Haplotype reconstruction from genotype data using imperfect phylogeny. *Bioinformatics* 2004;20:1842–1849. [PubMed: 14988101]
22. Fallin D, Cohen A, Essioux L, Chumakov I, Blumenfeld M, Cohen D, Schork NJ. Genetic analysis of case/control data using estimated haplotype frequencies: application to APOE locus variation and Alzheimer's disease. *Genome Res* 2001;11:143–151. [PubMed: 11156623]

23. Tajima F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 1989;123:585–595. [PubMed: 2513255]
24. Fu YX. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* 1997;147:915–925. [PubMed: 9335623]
25. Excoffier L, Smouse PE, Quattro JM. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 1992;131:479–491. [PubMed: 1644282]
26. Slatkin M. A measure of population subdivision based on microsatellite allele frequencies. *Genetics* 1995;139:457–462. [PubMed: 7705646]
27. Schneider, S.; Roessli, D.; Excoffier, L. Arlequin: a software for population genetics data analysis. Ver 2.000. Dept. of Anthropology, University of Geneva.: Genetics and Biometry Lab; 2000.
28. Page RD. TreeView: an application to display phylogenetic trees on personal computers. *Comput Appl Biosci* 1996;12:357–358. [PubMed: 8902363]
29. Freimuth RR, Stormo GD, McLeod HL. PolyMAPr: programs for polymorphism database mining, annotation, and functional analysis. *Hum Mutat* 2005;25:110–117. [PubMed: 15643605]
30. Conde L, Vaquerizas JM, Santoyo J, Al-Shahrour F, Ruiz-Llorente S, Robledo M, Dopazo J. PupaSNP Finder: a web tool for finding SNPs with putative effect at transcriptional level. *Nucleic Acids Res* 2004;32:W242–W248. [PubMed: 15215388]
31. Pepine CJ, Handberg-Thurmond E, Marks RG, Conlon M, Cooper-DeHoff R, Volkers P, Zellig P. Rationale and design of the international verapamil SR/trandolapril study (INVEST): an internet-based randomized trial in coronary artery disease patients with hypertension. *J Am Coll Cardiol* 1998;32:1228–1237. [PubMed: 9809930]
32. Choudhry S, Coyle NE, Tang H, Salari K, Lind D, Clark SL, et al. Population stratification confounds genetic association studies among Latinos. *Hum Genet* 2006;118:652–664. [PubMed: 16283388]
33. Voight BF, Kudaravalli S, Wen X, Pritchard JK. A map of recent positive selection in the human genome. *PLoS Biol* 2006;4:e72. [PubMed: 16494531]
34. Weir BS, Cardon LR, Anderson AD, Nielsen DM, Hill WG. Measures of human population structure show heterogeneity among genomic regions. *Genome Res* 2005;15:1468–1476. [PubMed: 16251456]

Fig. 1. *KCNMB1* gene structure and identified common SNPs. Orientation of 5' to 3': from right to left. (a) VISTA plot of *KCNMB1* gene sequence conservation (including 2 kb upstream and downstream of the genomic sequence) in mouse and rat versus human (base sequence). Colors: coding blocks in light blue, untranslated region in light green and conserved noncoding sequence in pink. (b) Gene structure of human *KCNMB1* gene, including four exons in black boxes, three introns (in lines). The upstream region and 3' untranslated region are shown in grey boxes. (c) Locations and minor allele frequencies of common single nucleotide polymorphisms (SNPs) (minor allele frequency greater than 0.05) relative to exons and conserved noncoding sequences is indicated by position. (d) Nonsynonymous SNPs. (e) Functionally important SNPs identified by PolyMAPr program. (f) Functionally important SNPs identified by PuPaSNP program. (Note: SNP12 has a minor allele frequency of 0.04, and is therefore not shown in this figure.)

Fig. 2.
Population-specific single nucleotide polymorphisms.

Fig. 3. Pairwise linkage disequilibrium (r^2) of 25 single nucleotide polymorphisms (SNPs) in *KCNMB1* in African-Americans (a), native Americans (b) and Caucasians (c). Region: Ch5 169751147 – Ch5 169737989. Color: black: $r^2 = 1$; shades of grey: $0 < r^2 < 1$; white: $r^2 = 0$.

Fig. 4.

Phylogeny and frequencies of the common *KCNMB1* haplotypes in chimp and three human populations. Note: chimp was used as outgroup (in red). AA, African-Americans (in orange); NA, native Americans (in blue); Cauc, Caucasians (in black).

Table 1

Identified *KCNMB1* SNPs

SNP	Region	Golden path position	Alleles	S/NS AA change	Estimated			Frequency			Reported in dbSNP?	rs no.	Freq info ^a ?	Validated ^b ?
					Overall	Minor	Allele	AA	NA	CAUC				
1	Promoter	ch5-169751147	T>C		0.314	0.275	0.368	0.300	0.300	0.300	Yes	rs314127	0.340	No
2	Promoter	ch5-169750902	A>G		0.458	0.450	0.425	0.500	0.500	0.490	Yes	rs314128	0.490	No
3	Promoter	ch5-169750695	A>G		0.388	0.395	0.300	0.474	0.474	0.353	Yes	rs314129	0.353	Yes
4	Promoter	ch5-169750376	C>T		0.025	0.000	0.075	0.000	0.000		No			
5	Promoter	ch5-169750337	G>T		0.136	0.053	0.275	0.075	0.075	0.229	Yes	rs314130	0.229	No
6	Promoter	ch5-169749971	G> A		0.167	0.286	0.133	0.094	0.094	NR	Yes	rs10066233	NR	No
7	Promoter	ch5-169749943	G>A		0.500	0.357	0.438	0.441	0.441	NR	Yes	rs7722205	NR	No
8	Promoter	ch5-169749908	G>A		0.190	0.306	0.125	0.150	0.150	NR	Yes	rs10066171	NR	No
9	Promoter	ch5-169749555	T>C		0.475	0.350	0.400	0.475	0.475	NR	Yes	rs10050842	NR	Yes
10	Promoter	ch5-169749362	C>G		0.009	0.000	0.000	0.028	0.028		No			
11	5'UTR	ch5-169749143	C>T		0.092	0.100	0.100	0.075	0.075	0.152	Yes	rs827778	0.152	No
12	5'UTR	ch5-169748952	G>A		0.042	0.075	0.000	0.050	0.050		No			
13	5'UTR	ch5-169748878	C>T		0.117	0.100	0.125	0.125	0.125	0.114	Yes	rs703508	0.114	No
14	Intron	ch5-169748777	T>A		0.375	0.306	0.472	0.350	0.350	0.476	Yes	rs2075612	0.476	No
15	Exon 2	ch5-169744976	C>T	S Cys18Cys	0.017	0.025	0.000	0.025	0.025		No			
16	Intron	ch5-169744881	C>T		0.169	0.079	0.200	0.225	0.225	0.233	Yes	rs2071156	0.233	Yes
17	Intron	ch5-169744866	C>G		0.009	0.000	0.025	0.000	0.000		No			
18	Intron	ch5-169744864	A>G		0.121	0.000	0.075	0.289	0.289		No			
19	Exon 3	ch5-169743374	G>A	NS Glu65Lys	0.117	0.175	0.075	0.100	0.100	NR	Yes	rs11739136	NR	No
20	Intron	ch5-169738579	A>G		0.339	0.175	0.500	0.342	0.342	0.390	Yes	rs314157	0.390	Yes
21	Exon 4	ch5-169738534	G>C	NS Val110Leu	0.093	0.075	0.125	0.079	0.079	0.114	Yes	rs2301149	0.114	Yes
22	3'UTR	ch5-169738208	C>A		0.325	0.150	0.500	0.325	0.325	0.354	Yes	rs2656842	0.354	Yes
23	3'UTR	ch5-169738207	A>C		0.317	0.125	0.500	0.325	0.325	0.371	Yes	rs2656841	0.371	Yes
24	3'UTR	ch5-169738172	G>C		0.017	0.000	0.050	0.000	0.000		No			
25	3'UTR	ch5-169737989	A>G		0.172	0.105	0.325	0.079	0.079	0.109	Yes	rs314156	0.109	No

Pharmacogenet Genomics. Author manuscript; available in PMC 2009 July 21.

The alleles are of all minus strand. All chromosome positions are as of 4 November 2005. AA, African-Americans; Caut, Caucasians; dbSNP, database single nucleotide polymorphism; NA, native Americans; NR, not reported; NS, nonsynonymous; S, synonymous; SNP, single nucleotide polymorphism.

^aMinor allele frequency information available in dbSNP.

^bValidated by multiple independent submissions to dbSNP (<http://www.ncbi.nlm.nih.gov/projects/SNP/>).

Table 2Neutral selection tests in *KCNMB1* among three populations

Population	Tajima's D	Fu's Fs
AA	1.734 (0.90)	-12.038 (< 0.0001)
NA	1.175 (0.83)	-7.912 (0.003)
Caucasians	0.756 (0.74)	-5.151 (0.011)

Nominal *P* values obtained from permutation tests are given in parentheses for each test.
AA, African-Americans; NA, native Americans.

Table 3
Final list of common *KCNMB1* SNPs for further studies

SNP	rs no.	MAF ^a	In conserved region?	NS	In-silico functional prediction
1	rs314127	0.314			Y
3	rs314129	0.388			Y
11	rs827778	0.092	Y		
12	NA	0.042	Y		
14	rs2075612	0.375			Y
19	rs11739136	0.117	Y	Y (Glu65Lys)	Y
21	rs2301149	0.093	Y	Y (Val110Leu)	
22	rs2656842	0.325	Y		

MAF, minor allele frequency; NS, nonsynonymous cSNPs; SNP, single nucleotide polymorphism; SNP, single nucleotide polymorphism; Y, yes.

^aMAF in the overall population.