HAL
open science

Novel polymorphic AluYb8 insertion in the WNK1 gene is associated with blood pressure variation in Europeans

Margus Putku, Katrin Kepp, Elin Org, Siim Sõber, David Comas, Margus Viigimaa, Gudrun Veldre, Peeter Juhanson, Pille Hallast, Neeme Tõnisson, et al.

To cite this version:

Margus Putku, Katrin Kepp, Elin Org, Siim Sõber, David Comas, et al.. Novel polymorphic AluYb8 insertion in the WNK1 gene is associated with blood pressure variation in Europeans. Human Mutation, 2011, 32 (7), pp.806. 10.1002/humu.21508 . hal-00639773

HAL Id: hal-00639773

https://hal.science/hal-00639773

Submitted on 10 Nov 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Human Mutation

Novel polymorphic AluYb8 insertion in the WNK1 gene is associated with blood pressure variation in Europeans

Journal:	*Human Mutation*
Manuscript ID:	humu-2010-0611.R1
Wiley - Manuscript type:	Research Article
Date Submitted by the	03-Mar-2011

For Peer Review

Complete List of Authors:	Putku, Margus; University of Tartu, Institute of Molecular and Cell Biology Kepp, Katrin; University of Tartu, Institute of Molecular and Cell Biology Org, Elin; University of Tartu, Institute of Molecular and Cell Biology Sõber, Siim; University of Tartu, Institute of Molecular and Cell Biology Comas, David; Universitat Pompeu Fabra, Institute of Evolutionary Biology (UPF-CSIC), CEXS-UPF-PRBB Viigimaa, Margus; North Estonia Medical Centre, Centre of Cardiology Veldre, Gudrun; University of Tartu, Institute of Molecular and Cell Biology; University of Tartu, Department of Cardiology Juhanson, Peeter; University of Tartu, Institute of Molecular and Cell Biology Hallast, Pille; University of Tartu, Institute of Molecular and Cell Biology Tõnisson, Neeme; University of Tartu, Institute of Molecular and Cell Biology Shaw-Hawkins, Sue; William Harvey Research Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, Clinical Pharmacology and The Genome Centre Caulfield, Mark; William Harvey Research Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, Clinical Pharmacology and The Genome Centre Khusnutdinova, Elza; Ufa Science Center, Russian Academy of Sciences, Institute of Biochemistry and Genetics Kožich, Viktor; Charles University-First Faculty of Medicine, Institute of Inherited Metabolic Disorders Munroe, Patricia; William Harvey Research Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, Clinical Pharmacology and The Genome Centre Laan, Maris; University of Tartu, Institute of Molecular and Cell Biology
Key Words:	WNK1, polymorphism screening, human-specific polymorphic AluYb8, blood pressure, association study, meta-analysis, HYPEST, BRIGHT , CADCZ, alternative splicing

SCHOLARONE™
Manuscripts

Humu-2010-0611

Research Article

Supporting Information for this preprint is available from the *Human Mutation* editorial office upon request (humu@wiley.com)

Novel polymorphic *AluYb8* insertion in the *WNK1* gene is associated with blood pressure variation in Europeans

Margus Putku[1], Katrin Kepp[1], Elin Org[1], Siim Sõber[1], David Comas[2], Margus Viigimaa[3,4], Gudrun Veldre[1,5], Peeter Juhanson[1], Pille Hallast[1], Neeme Tõnisson[1], HYPEST†, Sue Shaw-Hawkins[6], Mark J. Caulfield[6], BRIGHT†, Elza Khusnutdinova[7], Viktor Kožich[8], Patricia B. Munroe[6], Maris Laan[1]*

[1]Human Molecular Genetics Research Group, Institute of Molecular and Cell Biology, University of Tartu, Tartu, Estonia

[2]Institute of Evolutionary Biology (UPF-CSIC), CEXS-UPF-PRBB, Universitat Pompeu Fabra, Barcelona, Spain

[3]Centre of Cardiology, North Estonia Medical Centre, Tallinn, Estonia

[4]Tallinn University of Technology , Department of Biomedical Engineering, Chair of Medical Physics, Tallinn, Estonia

[5]Department of Cardiology, University of Tartu, Tartu, Estonia

For Peer Review

[6]Clinical Pharmacology and The Genome Centre, William Harvey Research Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London EC1M 6BQ

[7]Institute of Biochemistry and Genetics, Ufa Science Center, Russian Academy of Sciences, Ufa, Bashkortostan, Russia

[8]Institute of Inherited Metabolic Diseases, Charles University – First Faculty of Medicine, Prague, Czech Republic

†HYPertension in ESTonia (HYPEST), the British Genetics of Hypertension (BRIGHT); additional BRIGHT consortium members are given in Supp. Text S1.

* Correspondence: Maris Laan, Institute of Molecular and Cell Biology, University of Tartu, Riia 23, 51010 Tartu, Estonia; Telephone: +372-7375008, Fax: +372-7420286, Email: maris.laan@ut.ee

For Peer Review

Abstract

Mutations in *WNK1* and *WNK4* cause familial hypertension, the Gordon syndrome. *WNK1* and *WNK4* conserved non-coding regions were targeted to polymorphism screening using DHPLC and DGGE. The scan identified an undescribed polymorphic *AluYb8* insertion in *WNK1* intron 10. Screening in primates revealed that this *Alu*-insertion has probably occurred in human lineage. Genotyping in 18 populations from Europe, Asia and Africa (n=854) indicated an expansion of the *WNK1 AluYb8* bearing chromosomes out-of-Africa. The allele frequency in Sub-Saharan Africa was ~3.3 times lower than in other populations (4.8% versus 15.8%; P=9.7x10^{-9}). Meta-analysis across three European sample-sets (n=3494; HYPEST, Estonians; BRIGHT, the British; CADCZ, Czech) detected significant association of the *WNK1 AluYb8* insertion with blood pressure (BP; systolic BP, P=4.03x10^{-3}, effect 1.12; diastolic BP, P=1.21x10^{-2}, effect 0.67). Gender-stratified analysis revealed that this effect might be female-specific (n=2088; SBP, P=1.99x10^{-3}, effect 1.59; DBP P=3.64x10^{-4}, effect 1.23; resistant to Bonferroni correction), while no statistical support was identified for the association with male BP (n=1406). In leucocytes, the expressional proportions of the full-length *WNK1* transcript and the splice-form skipping exon 11 were significantly shifted in *AluYb8* carriers compared to non-carriers. The *WNK1 AluYb8* insertion might affect human BP via altering the profile of alternatively spliced transcripts.

Key words: *WNK1*; polymorphism screening; human-specific polymorphic *AluYb8*; blood pressure; association study; HYPEST; BRIGHT; CADCZ; meta-analysis; gene expression

For Peer Review

Introduction

Essential hypertension is a complex disease promoted by an unfavorable combination of person's life-style and heritable factors. It is a significant health risk leading to other cardiovascular and renal diseases. Genetic studies of monogenic, Mendelian forms of hypo- and hypertension have identified ~20 rare mutations in blood pressure regulating genes with a strong effect on the phenotype (Lifton et al., 2001; Vehaskari, 2007). Although these rare mutations do not explain blood pressure variation in general population, the identified genes are promising targets for functional, physiological and genetic studies of essential hypertension (Ji et al., 2008). Serine/threonine protein kinase family members WNK1 (MIM# 605232) and WNK4 (MIM# 601844) (Xu et al., 2000; Verissimo and Jordan, 2001) are involved in the development of a Mendelian form of hypertension, pseudohypoaldosteronism type II or the Gordon syndrome (Wilson et al., 2001, 2003). The syndrome is caused either by large deletions (two identified variants: 22 kb and 42 kb) in the first intron of *WNK1* or by non-synonymous substitutions in *WNK4* (four described mutations). Although *WNK1* and *WNK4* are expressed in multiple tissues, their major role is to regulate the transport of sodium and potassium ions in distal convoluted tubule and cortical collecting duct of nephrons, and thereby to contribute to blood pressure determination (Verissimo and Jordan, 2001; Wilson et al., 2001). The human *WNK4* gene (19 exons) spans ~16 kb on chromosome 17q21.31. The human *WNK1* gene (29 exons) covers ~160 kb on chromosome 12p13 and codes for multiple transcripts initiated by alternative promoters (Xu et al., 2000; Wilson et al., 2001; Delaloy et al., 2003). Two major WNK1 isoforms have been described, a long isoform (L-WNK1) with complete kinase domain and a short kidney-specific isoform (KS-WNK1), which is kinase-deficient (Xu et al., 2000). Although multiple alternative splice-forms of *WNK1* have been identified, the function of individual transcripts is

For Peer Review

yet to be determined. In addition to the identification of rare variants in *WNK1* and *WNK4* responsible for the Gordon syndrome, common single nucleotide polymorphisms (SNPs) in these genes have been associated with blood pressure variation and susceptibility to hypertension in general population among adults as well as children (Kokubo et al., 2004; Newhouse et al., 2005; Tobin et al., 2005; Tobin et al., 2008; Newhouse et al., 2009; Osada et al., 2009). SNPs in *WNK1* also affect the response of thiazide diuretics treatment on patient's blood pressure (Turner et al., 2005).

While monogenic diseases are usually caused by rare variants located in the coding sequence of a gene, common diseases are rather considered to result from genetic variation in gene regulatory elements altering the expressional profile of the locus (Pastinen and Hudson, 2004; Visel et al., 2009). As gene regulatory elements tend to map within evolutionarily conserved segments of the genome (Hardison, 2000; Elgar and Vavouri, 2008), these regions have a potential to harbour polymorphisms contributing to the susceptibility to common traits including essential hypertension.

The aim of the current study was to screen the evolutionarily conserved non-coding regions of *WNK1* and *WNK4* to identify novel polymorphisms potentially affecting blood pressure in general population. Variant screening resulted in the identification of a novel human-specific polymorphic *AluYb8* insertion in *WNK1* intron 10. This *Alu*-insertion was targeted to further evolutionary and population genetic analysis, as well as was also explored for association with blood pressure and its effect on the transcriptional profile of the *WNK1* gene in leucocytes.

For Peer Review

Materials and Methods

In silico analysis of conserved non-coding regions in *WNK1* and *WNK4*

Conserved Non-coding Regions (CNRs) in *WNK1* and *WNK4* were screened using the web-based VISTA software (http://genome.lbl.gov/vista/index.shtml) with the proposed default parameters (cut-off criteria: 100 bp sliding window; sequence identity ≥70%; comparison with rat and mouse). The analyzed loci spanned from 10 kb upstream to 10 kb downstream of *WNK1* (12p13.3; coordinates 722,486-900,879, NCBI Build 36.1, hg18) and *WNK4* (17q21.31; coordinates 38,176,222- 38,212,587, NCBI Build 36.1, hg18). All VISTA regions that had any overlap with annotated genes track at UCSC Genome Browser (http://genome.ucsc.edu/) were excluded as potential coding regions. Polymorphism discovery was targeted to CNRs with sequence identity ≥70% between human and rodents, length of the region 50-300 bp, and location >200bp from the nearest exon (Supp. Table S1).

Screening for novel polymorphisms in *WNK1* and *WNK4* conserved non-coding regions

In total 40 CNRs (n=29 in *WNK1*; n=11 in *WNK4*; definition of CNR is given above) were selected for polymorphism screening, which was conducted either by *Denaturing Gradient Gel Electrophoresis* (DGGE; INGENYphorU-2x2 system, Ingeny International BV, Goes, The Netherlands) and/or *Denaturing High-Performance Liquid Chromatography* method (DHPLC; Wave Technologies Inc. USA). In the design of the DGGE and DHPLC assays and in establishing the experimental conditions, the manufacturers' recommendations were followed. Details of the assays are given in Supp. Text S1. The design of both DGGE and DHPLC assays was unsuccessful for 7 CNRs in *WNK1* and 2 CNRs in *WNK4* due to failure in primer design

For Peer Review

(inappropriate primer Tm or more than two Tm melting points for the region of interest) or a negative result in the genome test (Supp. Table S1). The genome test was applied to confirm the unique binding of a tested primer in the genome. The rest of the 31 selected CNRs were screened for polymorphisms either by DHPLC (7 regions in *WNK1*; 5 in *WNK4*), by DGGE (7 in *WNK1*; 1 in *WNK4*) or by both assays (8 in *WNK1*; 3 in *WNK4*; Supp. Table S1). Primers for DHPLC and DGGE assays are given in Supp. Table S2 and S3, respectively.

The average length of the CNR segments selected for polymorphism screening was 145 bp (range: 68-291 bp) and PCR fragments was 360 bp (range: 245-487 bp). Genomic DNAs of essential hypertension patients from two Eastern European studies (n=22 from HYPEST; n=24 from CADCZ; detailed description below) were targeted to polymorphism screening by DGGE (individual DNAs) and/or DHPLC (pools of DNA from three patients). PCR products exhibiting evidence for the presence of a polymorphism were sequenced on both forward and reverse orientations. Polymorphisms were identified using BioEdit Sequence Alignment Editor (Hall T., Department of Microbiology, North Carolina State University).

Genotyping of the *WNK1 AluYb8* in general human population samples

For large scale genotyping of *WNK1 AluYb8* in humans PCR followed by standard agarose gel electrophoresis was used. The primer design (WNK1_Alu_F: 5`-GGGTAACCAACCCTTGAAGTAGG-3`; WNK1_Alu_R: 5`-GGGTACTTCTCAAGTGATTAGGAGGA-3`) was carried out using the web-based program Primer3 (Rozen and Skaletsky, 2000). Quality control of the genotyping by agarose gel electrophoresis was assured by including previously resequenced positive controls representing alternative genotype carriers on each gel: wild-type (PCR product 353 bp); heterozygous (PCR

For Peer Review

products 353 and 660 bp) and homozygous (PCR product 660 bp) individuals for the *AluYb8* insertion (Figure 1; Supp. Figures S1, S2). The distribution of the *WNK1 AluYb8* insertion was studied in six European (Estonians, n=100; Czech, n=50; CEPH, n=30; the Basque, n=50; Catalans, n=41; Spanish Gypsies, n=50), four Asian (Koreans, n=43; Chinese Han, n=25; Tatars, n=47; Bashkir, n=47) and eight African populations (Tunisians, n=48; Algerians, n=48; Moroccans, n=84; Mandenkalu, n=24; Saharawi, n=50; Gabon Bantus, n=50; Gabon Pygmies, n=50; Tanzanians, n=17).

Conservation of the *WNK1 AluYb8* insertion in primates

The presence of the *WNK1 AluYb8* insertion was ascertained for a gorilla (*Gorilla gorilla;* primary cell line AG05251B, purchased from ECACC), for an orangutan (*Pongo pygmaeus*; primary cell line AG12256, purchased from ECACC) and for eleven western chimpanzees (*Pan troglodytes verus*) using identical PCR setup as in human genotyping. DNA sample of one chimpanzee originates from a wild-born male specimen (Pino) from Tallinn Zoo, Estonia. Ten samples of wild caught and unrelated animals (Annaclara, Frits, Hilko, Louise, Marco, Oscar, Regina, Socrates, Sonja and Yoran) are from the collection stored at the Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany and were kindly shared by Dr. Svante Pääbo. This sample collection is described in detail elsewhere (Ptak et al., 2004; Becquet et al., 2007). Ancestral sequence of the targeted genomic region (*WNK1* exon10-intron10-exon11) was assessed by the comparative sequencing of the genomic DNA from *WNK1 AluYb8* insertion non-carrier (–/–) and carrier (Alu/Alu) human homozygotes as well as from a chimpanzee (Pino). Sequencing primers are listed in Supp. Table S4. PCR cycling conditions, product purification and sequencing have been described elsewhere (Hallast et al., 2005). Sequences were aligned

For Peer Review

using web-based global alignment program ClustalW2 (http://www.ebi.ac.uk/Tools/clustalw2/). Human *WNK1* alternative sequences of the region including exon 10/intron 10/exon 11 (without and with *AluYb8* insertion) were compared with available genome sequences from multiple species using the BLAST tool blastn (http://blast.ncbi.nlm.nih.gov/Blast.cgi). The searches were performed against the following sequence databases: NCBI Genomes, Whole-Genome-Shotgun Sequences. Nucleotide substitution rates between human and chimpanzee were calculated as the percentage of the number of substitutions divided with the total number of aligned nucleotides in the specific genomic region. The number of substitutions and the total number of aligned nucleotides were calculated using EMBOSS stretcher (http://emboss.sourceforge.net/; Rice et al., 2000).

Stage 1 association analysis in HYPEST

In Stage 1, the association of the *WNK1 AluYb8* insertion with blood pressure (BP) was addressed using HYPEST (HYPertension in ESTonia) case-cohort sample collection (Table 1; recruitment details in Supp. Text S1). The HYPEST study has been approved by the Ethics Committee on Human Research of University of Tartu (no. 122/13, 22.12.2003; 137/20, 25.04.2005) and it was carried out in compliance with the Helsinki Declaration. All the participants have given their written informed consent. HYPEST subjects were recruited across Estonia during 2004-2007 (1823 individuals, age range 18-85 years) with the aim to analyze genetic-epidemiological risk factors for essential hypertension and related cardiovascular disease in Estonian population. In the current study, the total number of genotyped HYPEST subjects was n=1747. At the recruitment, the resting BP of each participant has been measured by trained clinicians using a standard mercury column sphygmomanometer and size-adjusted cuffs.

For Peer Review

HYPEST individuals possessed a documented history of multiple systolic and diastolic BP readings (on average 4.31 readings per individual during mean 3.17 years). For the analysis, the median across the longitudinal BP readings as well as the median of the subject's age during the readings were used.

Association analysis with SBP and DBP was performed using 1,211 individuals (803 women, 408 men) derived from the population-based cohort across Estonia consisting of long-term blood donors not receiving any antihypertensive medication (Table 1). For binary analysis with essential hypertension, cases (n=673) were defined as untreated subjects with BP readings ≥ 160/100 mmHg based on the median of several measurements or patients receiving anti-hypertensive therapy. Normotensive controls (n=601; SBP ≤140 mmHg/ DBP ≤90 mmHg) were selected from the population-based HYPEST cohort among the subjects that have never been prescribed antihypertensive treatment.

Stage 2 replication

In Stage 2, association testing of the *WNK1 AluYb8* insertion and blood pressure (BP) was performed in two European samples, the BRIGHT (BRItish Genetics of HyperTension) and the CADCZ (Coronary Artery Disease in Czech), and the results were combined in meta-analysis with Stage 1 study samples. The final sample size in meta-analysis with SBP and DBP was 3494 subjects (2088 women, 1406 men; none treated with antihypertensive medication), and with hypertension 3312 cases/3047 controls (women, 1924/1911; men 1388/1136).

CADCZ study has been approved by the Ethics Committee of Charles University—1st Faculty of Medicine (December 1996) and the BRIGHT study was approved by the Ethics Committee from local research committees of all partner institutes. All BRIGHT and CADCZ participants

For Peer Review

have given their written informed consent. The MRC British Genetics of Hypertension case–control samples have been recruited across UK (http://www.brightstudy.ac.uk). Case ascertainment and phenotyping has been described elsewhere (Caulfield et al., 2003). Briefly, cases originated from severely hypertensive families (1700 sib-pairs and 800 families collected for transmission disequilibrium test) were defined as patients under antihypertensive treatment and with BP readings ≥150/100 mmHg based on one reading or ≥145/95 mmHg based on the mean of three readings. Healthy normotensive controls (n=2000; BP <140/90 mmHg, no antihypertensive medication and no diagnosed diseases) were recruited by matching age, sex and geographical distribution across UK. Following the study design, the association analysis with SBP and DBP included healthy untreated BRIGHT controls (n=1639, 1002 women, 637 men; Table 1). In case-control analysis hypertensives (n=2242) and normotensives (n=1639) were classified as defined at the recruitment. CADCZ subjects were recruited by the Cardiology Department of the 2nd Clinic of Internal Medicine, Faculty Hospital Královské Vinohrady in Prague Czech Republic and details of the recruitment are published elsewhere (Janosikova et al., 2003). Trained clinicians documented three measurements of resting BP and the median value was recorded. Association testing between SBP, DBP and the carrier status of *WNK1 AluYb8* insertion included subjects not receiving anti-hypertensive treatment (n=644, 283 women, 361 men; Table 1). Hypertensives (n=266) and normotensives (n=480) in case-control analysis were defined as in HYPEST.

RNA extraction and cDNA synthesis

EDTA-blood (9 ml) was collected from nine female subjects from HYPEST study selected based on their alternative genotypes: three heterozygotes (Alu/–) and three homozygotes (Alu/Alu) for

For Peer Review

the *WNK1 AluYb8* insertion; and three subjects with wild-type (–/–) sequence (Supp. Table S5). Total RNA from leucocytes was extracted using LeukoLOCK™ Total RNA Isolation System (Ambion Inc, Austin, Texas, USA) including an optional TURBO™ DNase treatment to degrade the genomic DNA. Quantity and quality of extracted RNA was assessed with NanoDrop® ND-1000 UV-Vis Spectrophotometer (NanoDrop Technologies, LLC, Wilmington, Delaware, USA). RNA was reverse transcribed using SuperScript™ III First-Strand Synthesis SuperMix for qRT-PCR (Life Technologies Corporation, Carlsbad, California, USA) according to the manufacturer's instructions (details in Supp. Text S1).

Quantification of *WNK1* transcripts by real-time PCR

Relative expression analysis of three *WNK1* splice-forms (ex+11+12, ex–11+12 and ex–11–12; Fig. 2a) was performed with real-time PCR. Primer-probe mix of the *WNK1* transcript including exon 11 (ex+11+12; Hs01018312_m1, amplicon size 78bp) and selected reference gene *HPRT1* [Human HPRT1 (HGPRT) Endogenous Control (VIC / MGB Probe, Primer Limited, amplicon size 100 bp)] were purchased from Applied Biosystems, Inc. Primers and probes for the *WNK1* transcripts lacking exon 11 (ex–11+12) and both exons 11 & 12 (ex–11–12) were designed using Primer Express version 3.0 (Applied Biosystems Inc, Foster City, CA, USA). Oligonucleotide sequences are given in Supp. Table S4.

The real-time RT-PCRs were performed using Applied Biosystems 7900HT Fast Real-time PCR system in 96 micro-well plates. Target region and endogenous control were amplified in the same well. The experimental conditions for the real-time PCR are given in detail in Supp. Text S1. In total six replicate analyses of each of the nine extracted RNA samples were conducted: two independently synthesized cDNAs were assayed by RT-PCR reactions in triplicate.

For Peer Review

Statistical analysis

Statistical differences in allele frequencies between populations were calculated using the web-based Fisher´s Exact Test calculator (http://www.langsrud.com/fisher.htm). The significance of the associations between the *WNK1 AluYb8* insertion and BP (SBP; DBP) as a quantitative trait was tested using linear regression (additive genetic model) with age and gender as covariates. Additive genetic model assumes a trend per copy of the minor allele to contribute to the trait or disease susceptibility on genotype categories. Association with the diagnosis of hypertension as a binary trait was assessed by logistic regression adjusted for age and sex. Association tests and calculation of LD between SNP pairs (r^2) were implemented in the PLINK software, ver. 1.04 (http://pngu.mgh.harvard.edu/~purcell/plink/). The Bonferroni threshold for multiple testing correction was estimated $0.05/9=5.56\times10^{-3}$, taking into account the number of tested phenotypes (three) and tested study samples (three). Results were combined in a meta-analysis using the inverse-variance method under fixed-effects model using R, ver. 2.7.2 (R Development Core Team 2008, http://www.r-project.org/).

Normalized expression values of target regions were calculated using Microsoft® Excel®-based software Q-Gene (Muller et al., 2002). Q-Gene calculates the normalized expression values of the target gene based on the Ct values and the reaction efficiencies of the target and the reference gene (here *HPRT*). For every study subject six replicate values of relative expression per each alternative *WNK1* splice-form (ex+11+12, ex–11+12 and ex–11–12) was calculated. As each of the genotypes (Alu/Alu, Alu/–; –/–) was represented by three individuals, in total 18 data-points were collected per transcript within a genotype group. The most outlier Ct value within the respective genotype group was excluded from the statistical testing. Differences of normalized

For Peer Review

expression values between alternative genotype groups were estimated by Wilcoxon rank sum test implemented in R software.

Results

Polymorphism screening in *WNK1* and *WNK4* conserved non-coding regions

DHPLC and/or DGGE assays were designed for screening novel polymorphisms in conserved non-coding regions (CNRs) of the *WNK1* (29 targeted CNRs based on criteria outlined in Materials and Methods) and the *WNK4* (11 CNRs) genes (Supp. Table S1). Based on the *in silico* quality control criteria for assay design, nine regions were excluded from the wet-lab analysis. Finally, 31 CNRs entered variant detection in Eastern European essential hypertension patients (from HYPEST and CADCZ studies). Among the screened 31 CNRs, one single nucleotide polymorphism (SNP) was identified in the *WNK4* and six SNPs in the *WNK1* gene (Supp. Table S6). All but one (rs36052085) of the detected SNPs were rare (minor allele frequency <10%), including three singletons (two novel). The functional effect of these SNPs was addressed neither by association study due to large sample size requirements nor by gene expression analysis due to unavailability of minor allele homozygotes.

In addition, in one of the *WNK1* CNRs, a novel unreported common indel (~ 300 bp) was detected (Fig. 1; Supp. Table S6). Sequence analysis of this variant revealed a polymorphic insertion of an *AluYb8* element (288 bp without flanking T nucleotides) into a poly-T tract within *WNK1* intron 10, ~780 bp upstream from exon 11 (Supp. Figure S2). This *Alu*-insertion was targeted to further evolutionary and population genetic analysis as well as was explored for the association with cardiovascular disease and the effect on the gene expression profile.

For Peer Review

Alu distribution among general human populations

The *WNK1 AluYb8* insertion was genotyped in 18 population samples from Europe, Asia and Africa (Supp. Table S7, Supp. Figure S3). The genotyped variant was in Hardy-Weinberg equilibrium in all but one (Saharawi, n=50) studied population samples. Frequency of the *AluYb8* insertion in human populations differed based on their geographic affiliation (Table 2, Supp. Figure S3). The proportion of *WNK1 AluYb8* carriers in Sub-Saharan Africa was significantly lower (average allele frequency 4.8%; range 2.1–7.0%) compared to North-African (mean 16.4%, range 10.4–25.0%; Fisher`s Exact Test, P=2.2x10^{-6}), European (mean 15.1%, range 12.0–16.5%; P=8.7x10^{-9}) and Asian (mean 15.9%, range 11.6%–22.0%; P=9.4x10^{-6}). On average, the allele frequency of the *WNK1 AluYb8* in Sub-Saharan Africa was ~3.3 times lower than in other studied populations (P=9.7x10^{-9}).

Alu insertion in primates and conservation around insertion site

The analysis of the *WNK1* intron 10 in eleven chimpanzees, one gorilla and one orangutan revealed that the *WNK1 AluYb8* insertion has most probably occurred in human lineage. No *AluYb8* insertion was detected in the *WNK1* intron 10 of the studied primate genomes (Fig. 1, Supp. Figure S1). Comparative sequencing of the *WNK1* genomic fragment (exon10/ intron10/ exon11) amplified from a chimpanzee and from human wild-type as well as *AluYb8* insertion carrying chromosomes revealed high conservation of intron 10 (Supp. Figure S1). The substitution divergence between human wild-type and chimpanzee *WNK1* was 0%, 0.2% and 1.1% for exon 10 (150 bp), exon 11 (459 bp) and intron 10 (1211 bp), respectively. Overall substitution rate was 0.7% and 1.3% between *WNK1* exons and introns, respectively. The

For Peer Review

uniqueness of the *WNK1 AluYb8* insertion in human was supported by a negative result of the BLAST search among available genome sequences.

Stage 1 association testing of the *WNK1 AluYb8* insertion with blood pressure and hypertension in the HYPEST study

The identified *WNK1 AluYb8* insertion was tested for association with blood pressure (BP) in the Estonian HYPEST cohort subjects (n=1211). The analysis detected significantly higher SBP (P=1.26x10^{-2}, effect 2.23 mmHg; linear regression, additive model) and DBP (P=3.04x10^{-2}, effect 1.22 mmHg; Table 3) among *Alu*-insertion carriers. Analysis in men and women separately revealed that the effect of the *AluYb8* insertion on BP might be female-specific (SBP: P=1.32x10^{-2}, effect 2.72 mmHg; DBP: P=6.20x10^{-3}, effect 1.84 mmHg), while no statistical support was found for the association in men (SBP, DBP, P>5.9x10^{-1}). We also observed higher *WNK1 AluYb8* frequency among HYPEST essential hypertension patients (n=673; 17.7%) compared to normotensive controls (n=601; 14.5%; Supp. Table S8).

Stage 2 association testing in BRIGHT and CADCZ, and meta-analysis across three European sample sets

To confirm the discovery association between the *WNK1 AluYb8* insertion and blood pressure (BP) identified in the HYPEST study (Estonians), we performed Stage 2 replication testing in two independent European samples, the BRIGHT (the British) and the CADCZ (Czech) (Table 1). Stage 1 and Stage 2 results were combined in a joint meta-analysis (Table 3). Meta-analysis across all studies (n=3494) improved significantly the support for the association with blood pressure (SBP, P=4.03x10^{-3}, effect 1.12 mmHg; DBP P=1.21x10^{-2} effect 0.67 mmHg). The

For Peer Review

pronounced effect of this *Alu*-insertion on BP in women was confirmed (n=2088; SBP, P=1.99x10^{-3}, effect 1.59 mmHg; DBP P=3.64x10^{-4} effect 1.23 mmHg). Detected associations of the *WNK1 AluYb8* insertion with SBP in the full sample and with female SBP and DBP remained significant after stringent correction for multiple testing (Bonferroni threshold α=0.05/9=5.56x10^{-3}; Table 3). Consistent with the discovery sample, no support was detected to the effect of *AluYb8* insertion on male BP (n=1406; SBP, DBP, P>4.7x10^{-1}).

Consistent with HYPEST, the trend for higher frequency of *WNK1 AluYb8* was observed in CADCZ hypertensive patients compared to controls (17.1% *vs* 15.3%), but not in the BRIGHT cases representing extreme family based hypertension (Supp. Table S8).

Linkage disequilibrium (LD) landscape between the novel *AluYb8* insertion and reported blood pressure associated *WNK1* SNPs in HYPEST and BRIGHT

When allelic association was assessed between the *AluYb8* insertion and rs765250 the associated SNP with SBP variation in a meta-analyses including the BRIGHT study and the HYPEST subjects in a previous report (Newhouse et al., 2009), low LD was estimated in both samples (HYPEST: r^2=0.069; BRIGHT: r^2=0.080). However, *AluYb8* was in strong LD with three previously genotyped *WNK1* SNPs in the BRIGHT resource: rs11064527 (r^2=0.821; MAF=0.16, intron 1), rs12816718 (r^2=0.921; MAF=0.15, intron 6) and rs956868 (r^2=0.975; MAF=0.14, exon 13) (Supp. Table S9). Amino acid alignment of the *WNK1* exon 13 among vertebrates exhibited low evolutionary conservation at the position of rs956868 coding for two alternative amino acids Proline or Threonine in human (Supp. Figure S4). When association testing with BP was performed for the BRIGHT individuals genotyped for the *AluYb8* insertion as well as the three SNPs in LD (n=1421), the strongest association was detected with *AluYb8* (Supp. Table S10).

For Peer Review

The impact of the *AluYb8* insertion on the expression profile of *WNK1* in leucocytes

The human *WNK1* gene codes for a high number of mRNA transcripts and extensive alternative splicing has been described for exons 9, 11 and 12 (Verissimo and Jordan, 2001). The major human *WNK1* transcript completely lacks exons 11 and 12 (Delaloy et al., 2003). In order to explore the functional effect of the *AluYb8* insertion on the expression profile of *WNK1* alternative transcripts, we quantified the gene transcripts in mRNA extracted from the leucocytes of nine women with alternative genotypes (Supp. Table S5). The study subjects included three individuals heterozygous and three homozygous for the *AluYb8* insertion, as well as three wild-type genotype carriers. The expression of three *WNK1* splice-forms was addressed using relative quantification method based on real-time RT-PCR assays. The studied splice-forms differed by alternative inclusion/exclusion of exon 11 and exon 12 (Fig. 2a). Compared to the subjects with the wild-type genotype, the heterozygous *AluYb8* carriers had significantly lower expression of all three splice-forms (Wilcoxon rank sum test; ex+11+12: P=6.08x10^{-3}; ex–11+12: P=6.78x10^{-7}; ex–11–12: P=5.39x10^{-3}; Fig. 2b). Consistently, homozygous *AluYb8* carriers showed lower expression level of splice forms ex–11+12 (P=1.62x10^{-3}) and ex–11–12 (P=1.82x10^{-1}; Fig. 2c). However, the expression of the full-length *WNK1* transcript, which includes both exon 11 and 12 (ex+11+12) was up-regulated among *AluYb8* homozygous carriers (P=3.56x10^{-2}). We conclude that the carrier-status of the *AluYb8* insertion may have an impact on the profile in *WNK1* transcript in leucocytes.

For Peer Review

Discussion

We targeted conserved non-coding regions in hypertension candidate genes *WNK1* and *WNK4* to polymorphism screening in order to identify functional variants potentially contributing to BP determination. We identified a novel human-specific polymorphic *AluYb8* insertion in *WNK1* intron 10. The *AluYb8* insertion belongs to a young *Alu* subfamily represented with ~2200 copies in the human genome compared to only nine insertions detected in chimpanzee (Gibbons et al., 2004). Consistently, we were unable to detect the studied *WNK1 AluYb8* insertion in chimpanzee, gorilla and orangutan (Fig. 1). As *AluYb8* elements are relatively mobile *Alu*-repeats, they represent together with *AluYa5* subfamily ~58% of the polymorphic *Alu*-s in the human genome (Bennett et al., 2008). The increased carrier frequency of *WNK1 AluYb8* insertion out of Africa is consistent with recent studies showing that the allele frequencies of polymorphic *Alu*-insertions tend to be lowest in Sub-Saharan populations (Table 2) (Watkins et al., 2001, 2003). In Africa, the fraction of carriers of polymorphic *Alu*-s increase with sharp cline in the north of Sahara compared to the populations living south of the desert (Comas et al., 2000).

Our study identified a significant association between *WNK1 AluYb8* insertion and blood pressure (BP) in the Estonian HYPEST cohort and confirmed this finding in the meta-analysis across three independent European study samples (n=3494; HYPEST, BRIGHT, CADCZ). The carriers of the *WNK1 Alu*-insertion had a consistent tendency for higher blood pressure (Table 3). Notably, when the analysis was performed using samples stratified by gender, the *WNK1 AluYb8* insertion was associated with BP only among women (meta-analysis: SBP, $P = 1.99 \times 10^{-3}$; DBP, $P = 3.64 \times 10^{-4}$) and no association was detected in men. Consistent with our findings, a sex-specific effect on BP determination was recently shown for a *WNK1* SNP in intron 1

For Peer Review

(rs10774461), which was also associated with BP only in females (Padmanabhan et al., 2010). Similar gender-specific effects have been reported for the polymorphic *Alu*-insertion (rs4646994) located in intron 16 of the *ACE* (*angiotensin converting enzyme*) gene (Rigat et al., 1990). Three independent studies showed that *ACE Alu* I/D variant is associated with the hypertension risk only in men and not in women (O'Donnell et al., 1998; Higaki et al., 2000; Stankovic et al., 2002). These differential effects on BP may reflect the differences in male and female physiology. Sex hormones have an important role in regulating a variety of renal transport functions and may contribute to gender differences in several kidney-related traits. Clinical observations in humans and experimental animals have shown that renal structure and functions under various physiological, pharmacological, and toxicological conditions are different in males and females, and that these differences may be related to the sex-hormone regulated expression and action of transporters in epithelial cells of nephrons (reviewed by Sabolic et al., 2007). Previously, BP variation in the BRIGHT and the HYPEST subjects has been associated with the *WNK1* SNP rs765250 (Newhouse et al., 2009). We detected low allelic association between *AluYb8* and rs765250 in both samples, which may indicate independent effects of *WNK1* intron 1 and intron 10 polymorphisms on BP. Interestingly, strong LD (r^2>0.8) between *AluYb8* and three previously studied *WNK1* SNPs in the BRIGHT study alone (rs11064527, rs12816718, rs956868) positioned *AluYb8* on the *WNK1* haplotype, was reported to show borderline associations with SBP and DBP ($P \leq 9\text{x}10^{-2}$) and a strong association with 24h urine potassium ($P<1\text{x}10^{-4}$; Newhouse et al., 2009). Although this haplotype includes a non-synonymous change (rs956868; Exon 13, Thr1316Pro, NP_001171914), *P*-values for the association with BP were the lowest for the *AluYb8* insertion compared to SNPs in LD. Previously, rs956868 has been shown to exhibit suggestive effect on ambulatory SBP in Europeans ($P<9\text{x}10^{-2}$; Tobin et al.,

For Peer Review

2005), and a significant association with SBP in Japanese ($P<5x10^{-2}$; Osada et al., 2009). The current and previous studies consistently report the highest BP levels in homozygotes for the minor allele of these polymorphisms (rs956868: Thr/Thr in LD with *AluYb8*: +/+). Functional assays should bring understanding whether the detected association with BP is driven by one primary variant or by a combinatory effect of the haplotype-forming alleles.

The design of the current study did not allow us to draw any conclusion about the contribution of the *WNK1 AluYb8* insertion to the risk for developing hypertension. The initially observed but not confirmed higher proportion of *AluYb8* carriers among hypertensives may have resulted from non-optimal selection of HYPEST controls (biased to too young), and/or different recruitment strategies of hypertensive patients among studies (HYPEST, essential hypertension; BRIGHT, extreme family-based hypertension; CADCZ, hypertension in CAD patients).

Although a majority of *Alu* elements are considered to be neutral residents of the human genome, an inserted copy of an *Alu*-repeat could interrupt structurally or functionally important genomic regions and consequently affect the expression of a locus (Batzer and Deininger, 2002; Callinan and Batzer, 2006). *Alu* elements may alter gene expression through modulating alternative splicing, RNA editing, epigenetic regulation and translation regulation (Hasler and Strub, 2006; Cordaux and Batzer, 2009). So far, thirty-three diseases directly caused by novel *Alu* insertions have been identified (Belancio et al., 2008). Our study using mRNA extracted from human leucocytes indicated a potential effect of the presence of the *AluYb8* insertion in *WNK1* intron 10 on the expressional profile of *WNK1* alternative transcripts. Splicing is an incompletely understood process carried out by large macromolecular complex spliceosome and directed by numerous regulatory elements located within exonic and intronic sequence (Black, 2003). The size of the *WNK1* intron 10 (human wild-type 1211 bp) is remarkably increased by the ~300 bp

For Peer Review

AluYb8 insertion (human variant >1500 bp). Thus, we hypothesize that the presence of the *AluYb8* insertion may disrupt the spatial intronic structure and/or disarrange the possible splicing regulatory sequences within *WNK* intron 10. Consequently, it may affect the splicing efficiency of the down-stream exons 11 and 12. As alternative splicing tends to be a tissue and developmental stage specific process (Xu et al., 2002), the impact of *AluYb8* insertion on the expressional profile of *WNK1* may vary in different tissues. The current study design was limited to addressing the effect of *AluYb8* insertion on *WNK1* expressional profile in leucocytes using a small number of samples. Further *in vitro* and *in vivo* studies using renal tissues would reveal the potential effect of this *Alu*-insertion on *WNK1* expressional profile in kidneys, where it plays an important role in contributing to the regulation of ion transport.

In summary, we identified a novel human-specific polymorphic *AluYb8* insertion in *WNK1*. This *AluYb8* insertion showed significant replicated association with blood pressure and a potential effect on the expressional profile of alternative *WNK1* transcripts in leucocytes.

For Peer Review

Acknowledgments

We express our appreciation to all the HYPEST, the BRIGHT and the CADCZ participants. We thank Dr. Bohumila Janosikova for technical assistance in recruitment of the CADCZ sample collection and the personnel of the Centre of Cardiology, North Estonia Medical Centre, Tallinn and the Department of Cardiology, University of Tartu for assistance in HYPEST subject recruitment and data collection. Dr Stephen Newhouse is acknowledged for kindly sharing previously published genotyping data of *WNK1* SNPs for the BRIGHT subjects. Piret Kelgo is thanked for laboratory assistance. We also thank Lolke van der Veen, DDL Laboratory, Institut des Sciences de l'Homme, Lyon, for sharing Gabonese samples that he and his team collected during the LCGB project. The study was supported by Wellcome Trust International Senior Research Fellowships (grant no. 070191/Z/03/Z, to ML; no. 070255/Z/03/Z to VK) in Biomedical Science in Central Europe. The BRIGHT study was supported by the Medical Research Council of Great Britain (grant number; G9521010D) and the British Heart Foundation (grant number PG02/128), and The Barts and The London Charity. This work forms part of the research themes contributing to the translational research portfolio for Barts and the London Cardiovascular Biomedical Research Unit which is supported and funded by the National Institute of Health Research. Additionally, the study has been supported by Estonian Ministry of Education and Science core grant no. 0182721s06, HHMI International Scholarship #55005617 (to ML), Estonian Science Foundation grants no ETF7471 (to ML; stipend for MP), and by a grant from the Ministry of Health of the Czech Republic NS10036-4/2008 (to VK). KK has been supported by personal travel-stipends from Kristjan-Jaak foundation and EC ECOGENE project.

For Peer Review

Disclosures: The authors declare that they have no conflict of interest.

References

Batzer MA, Deininger PL. 2002. Alu repeats and human genomic diversity. Nat Rev Genet 3(5):370-9.

Becquet C, Patterson N, Stone AC, Przeworski M, Reich D. 2007. Genetic structure of chimpanzee populations. PLoS Genet 3(4):e66.

Belancio VP, Hedges DJ, Deininger P. 2008. Mammalian non-LTR retrotransposons: for better or worse, in sickness and in health. Genome Res 18(3):343-58.

Bennett EA, Keller H, Mills RE, Schmidt S, Moran JV, Weichenrieder O, Devine SE. 2008. Active Alu retrotransposons in the human genome. Genome Res 18(12):1875-83.

Black DL. 2003. Mechanisms of alternative pre-messenger RNA splicing. Annu Rev Biochem 72:291-336.

Callinan PA, Batzer MA. 2006. Retrotransposable Elements and Human Disease. Genome and Disease 1:104-115.

Caulfield M, Munroe P, Pembroke J, Samani N, Dominiczak A, Brown M, Benjamin N, Webster J, Ratcliffe P, O'Shea S, Papp J, Taylor E, Dobson R, Knight J, Newhouse S, Hooper J, Lee W, Brain N, Clayton D, Lathrop GM, Farrall M, Connell J. 2003. Genome-wide mapping of human loci for essential hypertension. Lancet 361(9375):2118-23.

Comas D, Calafell F, Benchemsi N, Helal A, Lefranc G, Stoneking M, Batzer MA, Bertranpetit J, Sajantila A. 2000. Alu insertion polymorphisms in NW Africa and the Iberian Peninsula: evidence for a strong genetic boundary through the Gibraltar Straits. Hum Genet 107(4):312-9.

For Peer Review

Cordaux R, Batzer MA. 2009. The impact of retrotransposons on human genome evolution. Nat Rev Genet 10(10):691-703.

Delaloy C, Lu J, Houot AM, Disse-Nicodeme S, Gasc JM, Corvol P, Jeunemaitre X. 2003. Multiple promoters in the WNK1 gene: one controls expression of a kidney-specific kinase-defective isoform. Mol Cell Biol 23(24):9208-21.

Elgar G, Vavouri T. 2008. Tuning in to the signals: noncoding sequence conservation in vertebrate genomes. Trends Genet 24(7):344-52.

Gibbons R, Dugaiczyk LJ, Girke T, Duistermars B, Zielinski R, Dugaiczyk A. 2004. Distinguishing humans from great apes with AluYb8 repeats. J Mol Biol 339(4):721-9.

Hallast P, Nagirnaja L, Margus T, Laan M. 2005. Segmental duplications and gene conversion: Human luteinizing hormone/chorionic gonadotropin beta gene cluster. Genome Res 15(11):1535-46.

Hardison RC. 2000. Conserved noncoding sequences are reliable guides to regulatory elements. Trends Genet 16(9):369-72.

Hasler J, Strub K. 2006. Alu elements as regulators of gene expression. Nucleic Acids Res 34(19):5491-7.

Higaki J, Baba S, Katsuya T, Sato N, Ishikawa K, Mannami T, Ogata J, Ogihara T. 2000. Deletion allele of angiotensin-converting enzyme gene increases risk of essential hypertension in Japanese men : the Suita Study. Circulation 101(17):2060-5.

Janosikova B, Pavlikova M, Kocmanova D, Vitova A, Vesela K, Krupkova L, Kahleova R, Krijt J, Kraml P, Hyanek J, Zvarova J, Andel M, Kozich V. 2003. Genetic variants of homocysteine metabolizing enzymes and the risk of coronary artery disease. Mol Genet Metab 79(3):167-75.

For Peer Review

Ji W, Foo JN, O'Roak BJ, Zhao H, Larson MG, Simon DB, Newton-Cheh C, State MW, Levy D, Lifton RP. 2008. Rare independent mutations in renal salt handling genes contribute to blood pressure variation. Nat Genet 40(5):592-9.

Kokubo Y, Kamide K, Inamoto N, Tanaka C, Banno M, Takiuchi S, Kawano Y, Tomoike H, Miyata T. 2004. Identification of 108 SNPs in TSC, WNK1, and WNK4 and their association with hypertension in a Japanese general population. J Hum Genet 49(9):507-15.

Lifton RP, Gharavi AG, Geller DS. 2001. Molecular mechanisms of human hypertension. Cell 104(4):545-56.

Muller PY, Janovjak H, Miserez AR, Dobbie Z. 2002. Processing of gene expression data generated by quantitative real-time RT-PCR. Biotechniques 32(6):1372-4, 1376, 1378-9.

Newhouse S, Farrall M, Wallace C, Hoti M, Burke B, Howard P, Onipinla A, Lee K, Shaw-Hawkins S, Dobson R, Brown M, Samani NJ, Dominiczak AF, Connell JM, Lathrop GM, Kooner J, Chambers J, Elliott P, Clarke R, Collins R, Laan M, Org E, Juhanson P, Veldre G, Viigimaa M, Eyheramendy S, Cappuccio FP, Ji C, Iacone R, Strazzullo P, Kumari M, Marmot M, Brunner E, Caulfield M, Munroe PB. 2009. Polymorphisms in the WNK1 gene are associated with blood pressure variation and urinary potassium excretion. PLoS One 4(4):e5003.

Newhouse SJ, Wallace C, Dobson R, Mein C, Pembroke J, Farrall M, Clayton D, Brown M, Samani N, Dominiczak A, Connell JM, Webster J, Lathrop GM, Caulfield M, Munroe PB. 2005. Haplotypes of the WNK1 gene associate with blood pressure variation in a severely hypertensive population from the British Genetics of Hypertension study. Hum Mol Genet 14(13):1805-14.

For Peer Review

O'Donnell CJ, Lindpaintner K, Larson MG, Rao VS, Ordovas JM, Schaefer EJ, Myers RH, Levy D. 1998. Evidence for association and genetic linkage of the angiotensin-converting enzyme locus with hypertension and blood pressure in men but not women in the Framingham Heart Study. Circulation 97(18):1766-72.

Osada Y, Miyauchi R, Goda T, Kasezawa N, Horiike H, Iida M, Sasaki S, Yamakawa-Kobayashi K. 2009. Variations in the WNK1 gene modulates the effect of dietary intake of sodium and potassium on blood pressure determination. J Hum Genet 54(8):474-8.

Padmanabhan S, Menni C, Lee WK, Laing S, Brambilla P, Sega R, Perego R, Grassi G, Cesana G, Delles C, Mancia G, Dominiczak AF. 2010. The effects of sex and method of blood pressure measurement on genetic associations with blood pressure in the PAMELA study. J Hypertens 28(3):465-77.

Pastinen T, Hudson TJ. 2004. Cis-acting regulatory variation in the human genome. Science 306(5696):647-50.

Ptak SE, Roeder AD, Stephens M, Gilad Y, Paabo S, Przeworski M. 2004. Absence of the TAP2 human recombination hotspot in chimpanzees. PLoS Biol 2(6):e155.

Rice P, Longden I, Bleasby A. 2000. EMBOSS: the European Molecular Biology Open Software Suite. Trends Genet 16(6):276-7.

Rigat B, Hubert C, Alhenc-Gelas F, Cambien F, Corvol P, Soubrier F. 1990. An insertion/deletion polymorphism in the angiotensin I-converting enzyme gene accounting for half the variance of serum enzyme levels. J Clin Invest 86(4):1343-6.

Rozen S, Skaletsky H. 2000. Primer3 on the WWW for general users and for biologist programmers. Methods Mol Biol 132:365-86.

For Peer Review

Sabolic I, Asif AR, Budach WE, Wanke C, Bahn A, Burckhardt G. 2007. Gender differences in kidney function. Pflugers Arch 455(3):397-429.

Stankovic A, Zivkovic M, Alavantic D. 2002. Angiotensin I-converting enzyme gene polymorphism in a Serbian population: a gender-specific association with hypertension. Scand J Clin Lab Invest 62(6):469-75.

Tobin MD, Raleigh SM, Newhouse S, Braund P, Bodycote C, Ogleby J, Cross D, Gracey J, Hayes S, Smith T, Ridge C, Caulfield M, Sheehan NA, Munroe PB, Burton PR, Samani NJ. 2005. Association of WNK1 gene polymorphisms and haplotypes with ambulatory blood pressure in the general population. Circulation 112(22):3423-9.

Tobin MD, Timpson NJ, Wain LV, Ring S, Jones LR, Emmett PM, Palmer TM, Ness AR, Samani NJ, Smith GD, Burton PR. 2008. Common variation in the WNK1 gene and blood pressure in childhood: the Avon Longitudinal Study of Parents and Children. Hypertension 52(5):974-9.

Turner ST, Schwartz GL, Chapman AB, Boerwinkle E. 2005. WNK1 kinase polymorphism and blood pressure response to a thiazide diuretic. Hypertension 46(4):758-65.

Watkins WS, Ricker CE, Bamshad MJ, Carroll ML, Nguyen SV, Batzer MA, Harpending HC, Rogers AR, Jorde LB. 2001. Patterns of ancestral human diversity: an analysis of Alu-insertion and restriction-site polymorphisms. Am J Hum Genet 68(3):738-52.

Watkins WS, Rogers AR, Ostler CT, Wooding S, Bamshad MJ, Brassington AM, Carroll ML, Nguyen SV, Walker JA, Prasad BV, Reddy PG, Das PK, Batzer MA, Jorde LB. 2003. Genetic variation among world populations: inferences from 100 Alu insertion polymorphisms. Genome Res 13(7):1607-18.

Vehaskari VM. 2007. Heritable forms of hypertension. Pediatr Nephrol 24(10):1929-1937.

For Peer Review

Verissimo F, Jordan P. 2001. WNK kinases, a novel protein kinase subfamily in multi-cellular organisms. Oncogene 20(39):5562-9.

Wilson FH, Disse-Nicodeme S, Choate KA, Ishikawa K, Nelson-Williams C, Desitter I, Gunel M, Milford DV, Lipkin GW, Achard JM, Feely MP, Dussol B, Berland Y, Unwin RJ, Mayan H, Simon DB, Farfel Z, Jeunemaitre X, Lifton RP. 2001. Human hypertension caused by mutations in WNK kinases. Science 293(5532):1107-12.

Wilson FH, Kahle KT, Sabath E, Lalioti MD, Rapson AK, Hoover RS, Hebert SC, Gamba G, Lifton RP. 2003. Molecular pathogenesis of inherited hypertension with hyperkalemia: the Na-Cl cotransporter is inhibited by wild-type but not mutant WNK4. Proc Natl Acad Sci U S A 100(2):680-4.

Visel A, Rubin EM, Pennacchio LA. 2009. Genomic views of distant-acting enhancers. Nature 461(7261):199-205.

Xu B, English JM, Wilsbacher JL, Stippec S, Goldsmith EJ, Cobb MH. 2000. WNK1, a novel mammalian serine/threonine protein kinase lacking the catalytic lysine in subdomain II. J Biol Chem 275(22):16795-801.

Xu Q, Modrek B, Lee C. 2002. Genome-wide detection of tissue-specific alternative splicing in the human transcriptome. Nucleic Acids Res 30(17):3754-66.

For Peer Review

Figure Legends

Fig. 1. Detection of the presence of *WNK1* intron 10 *AluYb8* insertion in primates. Agarose gel (3%) electrophoresis of *WNK1* intron 10 PCR products amplified from human, chimpanzee, gorilla and orangutan genomic DNAs. In humans, alternative genotype carriers are shown: wild-

type homozygote without *AluYb8* insertion (–/–, PCR product 353 bp); heterozygous (A/–) and homozygous (A/A, PCR product 660 bp) carriers of the insertion

Fig. 2. Expression of **(a)** three *WNK1* alternative splice-forms in blood leucocytes obtained from **(b)** heterozygous (Alu/–) and **(c)** homozygous (Alu/Alu) carriers of the *WNK1 AluYb8* insertion in comparison with the wild-type homozygote without the insertion.

(a) Alternative splicing of *WNK1* exons 10-13 is presented schematically according to (Delaloy et al., 2003). Black numbered boxes and horizontal lines represent exons and introns, respectively and dotted lines indicate splicing events. **(b, c)** Relative mRNA quantification of the targeted *WNK1* splice-forms in leucocytes was performed with real-time RT-PCR (Taqman assay, *HPRT* as a reference gene). Relative expression of each targeted *WNK1* splice-form in subjects with the *AluYb8* insertion (Alu/Alu homozygotes, Alu/– heterozygotes) is shown using the quantity of the transcript in wild-type homozygotes (–/–) as a reference value (wt=1). The presented relative expression levels represent the mean values of the three study subjects within the genotype group (each individual represented by six data-points from replicate experiments). Bars represent standard error of the relative expression. *P*-values reflecting the differences between groups were estimated by Wilcoxon rank sum test.

For Peer Review

Table 1. Phenotypic parameters of study subjects in the analysis with systolic (SBP) and diastolic (DBP) blood pressure and hypertension (HYP)

	HYPEST			CADCZ			BRIGHT	
	Population	Essential hypertension [b]		Population	Essential hypertension [b]		Family based hypertension [c]	
Parameter (mean±SD)	sample [a]	Cases	Controls	sample [a]	Cases	Controls	Cases	Controls
No. of individuals	1211	673	601	644	266	480	2242	1639
(male/female)	(408/803)	(228/445)	(162/439)	(361/283)	(180/86)	(229/251)	(922/1320)	(637/1002)
Age at recruitment (y)	44.8±12.5	56.0 ±9.5	38.9±9.0	47.8±11.0	55.5±6.5	45.9±11.1	57.2±10.8	58.8±9.0
Age at onset of disease (y)	na	44.0±12.7	na	na	46.5±9.6	na	46.8±10.3	na
BMI (kg/m^2)	26.7±4.8	30.3±5.1	24.6±4.3	26.1±4.0	29.3±4.5	25.6±3.9	27.4±3.9	25.3±3.2
SBP (mmHg)	141.0±19.0	144.2±17.6	127.8±8.0	126.0±14.4	146.8±17.7	122.0±10.2	154.2±20.8	123.1±10.5
DBP (mmHg)	87.2±11.0	88.3±10.6	80.8±6.4	80.4±9.4	90.6±9.3	78.2±8.0	94.1±11.1	76.5±7.1
Antihypertensive treatment (% of subjects)	0.0%	78.5%	0.0%	0.0%	85.0%	0.0%	100%	0.0%

[a] Subjects from Estonian (HYPEST) and Czech (CADCZ) populations not receiving antihypertensive medication and used in association analysis with SBP and DBP

[b] Cases: subjects under antihypertensive treatment or untreated subjects SBP ≥160 mmHg and/or DBP ≥100 mmHg; Controls: subjects with SBP ≤140 mmHg and DBP≤90 mmHg, receiving no antihypertensive medication.

[c] Cases: patients from severely hypertensive families, under antihypertensive treatment and with BP ≥150/100 mmHg based on one reading or ≥145/95 mmHg based on the mean of three readings; Controls: subjects with BP <140/90 mmHg, receiving no antihypertensive medication

For Peer Review

SD, standard deviation; BMI, body mass index (kg/m^2); SBP, DBP, systolic and diastolic blood pressure; y, age in years; na, not applicable

For Peer Review

Table 2. *WNK1* intron 10 *AluYb8* allele frequencies in population groups

Group	No of subjects	Allele frequency	Population composition
Eastern Europe	150	16.3%	Estonians, Czech
Western Europe	121	14.1%	CEPH/Utah families, Basques, Catalans
Gypsies	50	12.0%	Spanish Gypsies
Volga-Ural	94	17.0%	Tatars, Bashkirs
Eastern Asia	68	16.8%	Chinese Han, Koreans
North Africa	230	16.4%	Moroccans, Saharawi, Algerians, Tunisians
Sub-Saharan Africa	141	4.8%	Mandenkalu, Tanzanians, Gabon Bantus, Gabon Pygmies

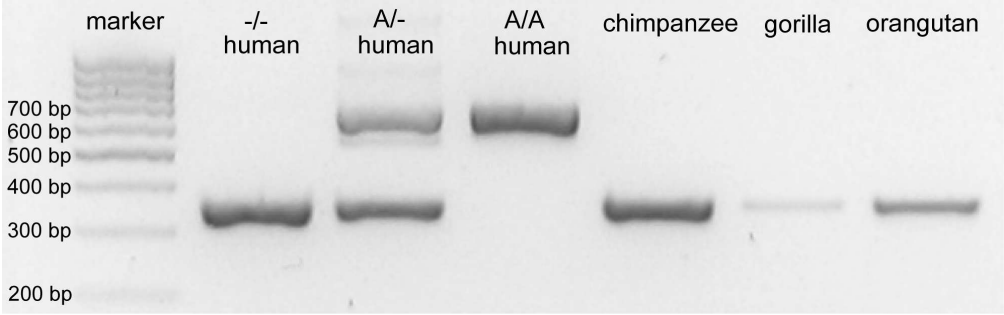
For Peer Review

Table 3. Association of *AluYb8* insertion with systolic (SBP) and diastolic (DBP) blood pressure

		HYPEST[a] (*N*=1211/803/408)[d]		CADCZ[a] (*N*=644/283/361)		BRIGHT[b] (*N*=1639/1002/637)		Joint meta-analysis[c] (*N*=3494/2088/1406)	
		Beta (SE)[e]	*P*-value	Beta (SE)	*P*-value	Beta (SE)	*P*-value	Beta (SE)	*P*-value
SBP									
	all	2.23 (0.89)	**1.26×10^{-2}**	0.05 (1.09)	9.65×10^{-1}	1.01 (0.47)	**3.23×10^{-2}**	1.12 (0.39)	**4.03×10^{-3}***
	women	2.72 (1.10)	**1.32×10^{-2}**	2.45 (1.60)	1.26×10^{-1}	1.09 (0.63)	8.24×10^{-2}	1.59 (0.52)	**1.99×10^{-3}***
	men	0.78 (1.48)	5.97×10^{-1}	-2.10 (1.48)	1.58×10^{-1}	0.74 (0.68)	2.71×10^{-1}	0.33 (0.57)	5.58×10^{-1}
DBP									
	all	1.22 (0.56)	**3.04×10^{-2}**	0.10 (0.71)	8.95×10^{-1}	0.60 (0.33)	7.22×10^{-2}	0.67 (0.27)	**1.21×10^{-2}**
	women	1.84 (0.67)	**6.19×10^{-3}**	1.00 (1.07)	3.51×10^{-1}	1.02 (0.44)	**2.00×10^{-2}**	1.23 (0.35)	**3.64×10^{-4}***
	men	-0.17 (1.03)	8.67×10^{-1}	-0.72 (0.94)	4.42×10^{-1}	-0.20 (0.51)	6.95×10^{-1}	-0.30 (0.41)	4.71×10^{-1}

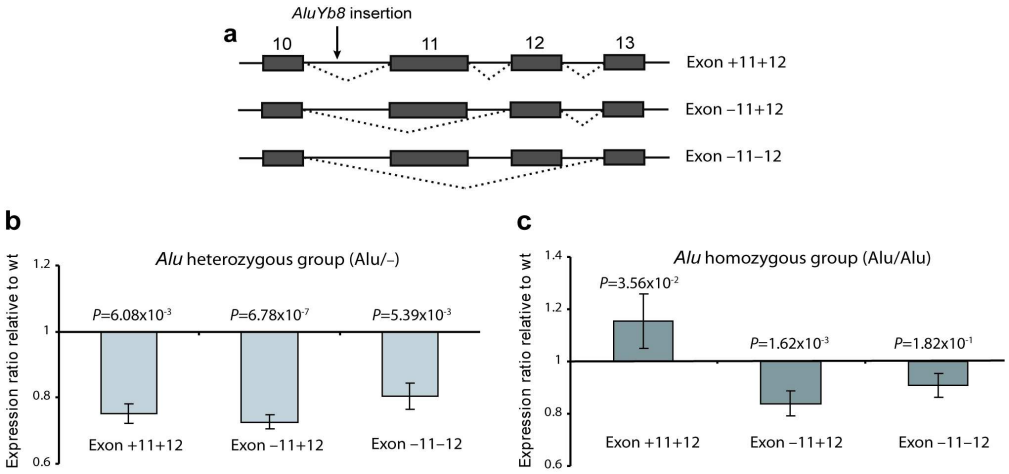
[a] Population-based subjects not receiving blood pressure lowering medication
[b] Normotensive controls across UK not receiving blood pressure lowering medication
[c] Meta-analysis of HYPEST, BRIGHT and CADCZ; inverse-variance method under fixed-effect model
[d] *N*=all/women/men
[e] Linear regression (additive model, age and gender as covariates) was used to test association with SBP and DBP (effect given as beta, SE). $P < 0.05$ is given **in bold** and *P*-values resistant to Bonferroni correction for multiple testing are indicated with the asterisk (*). Bonferroni significance level was estimated $\alpha = 0.05/9 = 5.56 \times 10^{-3}$ (3 phenotypes x 3 study samples).
SBP, DBP, systolic and diastolic blood pressure; N, number of subjects; SE, standard error

For Peer Review



Detection of the presence of WNK1 intron 10 AluYb8 insertion in primates. Agarose gel (3%) electrophoresis of WNK1 intron 10 PCR products amplified from human, chimpanzee, gorilla and orangutan genomic DNAs. In humans, alternative genotype carriers are shown: wild-type homozygote without AluYb8 insertion (–/–, PCR product 353 bp); heterozygous (A/–) and homozygous (A/A, PCR product 660 bp) carriers of the insertion
168x53mm (300 x 300 DPI)

Peer Review



Expression of (a) three WNK1 alternative splice-forms in blood leucocytes obtained from (b) heterozygous (Alu/–) and (c) homozygous (Alu/Alu) carriers of the WNK1 AluYb8 insertion in comparison with the wild-type homozygote without the insertion.
(a) Alternative splicing of WNK1 exons 10-13 is presented schematically according to (Delaloy et al., 2003). Black numbered boxes and horizontal lines represent exons and introns, respectively and dotted lines indicate splicing events. (b, c) Relative mRNA quantification of the targeted WNK1 splice-forms in leucocytes was performed with real-time RT-PCR (Taqman assay, HPRT as a reference gene). Relative expression of each targeted WNK1 splice-form in subjects with the AluYb8 insertion (Alu/Alu homozygotes, Alu/– heterozygotes) is shown using the quantity of the transcript in wild-type homozygotes (–/–) as a reference value (wt=1). The presented relative expression levels represent the mean values of the three study subjects within the genotype group (each individual represented by six data-points from replicate experiments). Bars represent standard error of the relative expression. P-values reflecting the differences between groups were estimated by Wilcoxon rank sum test.
171x80mm (300 x 300 DPI)

SUPPORTING INFORMATION

Novel polymorphic *AluYb8* insertion in the *WNK1* gene is associated with blood pressure variation in Europeans

Margus Putku[1], Katrin Kepp[1], Elin Org[1], Siim Sõber[1], David Comas[2], Margus Viigimaa[3,4], Gudrun Veldre[1,5], Peeter Juhanson[1], Pille Hallast[1], Neeme Tõnisson[1], HYPEST†, Sue Shaw-Hawkins[6], Mark J. Caulfield[6], BRIGHT†, Elza Khusnutdinova[7], Viktor Kožich[8], Patricia B. Munroe[6], Maris Laan[1]*

[1]Human Molecular Genetics Research Group, Institute of Molecular and Cell Biology, University of Tartu, Tartu, Estonia

[2]Institute of Evolutionary Biology (UPF-CSIC), CEXS-UPF-PRBB, Universitat Pompeu Fabra, Barcelona, Spain

[3]Centre of Cardiology, North Estonia Medical Centre, Tallinn, Estonia

[4]Tallinn University of Technology , Department of Biomedical Engineering, Chair of Medical Physics, Tallinn, Estonia

[5]Department of Cardiology, University of Tartu, Tartu, Estonia

[6]Clinical Pharmacology and The Genome Centre, William Harvey Research Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London EC1M 6BQ

[7]Institute of Biochemistry and Genetics, Ufa Science Center, Russian Academy of Sciences, Ufa, Bashkortostan, Russia

For Peer Review

[8]Institute of Inherited Metabolic Diseases, Charles University – First Faculty of Medicine, Prague, Czech Republic

†HYPertension in ESTonia (HYPEST), the British Genetics of Hypertension (BRIGHT); additional BRIGHT consortium contributors are given in Supp. Text S1.

* Correspondence and reprint requests to: Maris Laan, Institute of Molecular and Cell Biology, University of Tartu, Riia 23, 51010 Tartu, Estonia; Telephone: +372-7375008, Fax: +372-7420286, Email: maris.laan@ut.ee

For Peer Review

LIST OF SUPPORTING INFORMATION:

Supp. Text S1. List of additional BRIGHT consortium contributors, additional information for Materials and Methods

Supp. Figure S1. PCR amplification of the locus for *WNK1* intron 10 *AluYb8* insertion in eleven chimpanzees as resolved by agarose gel (3%) electrophoresis.

Supp. Figure S2. Sequence alignment of the *WNK1* genomic fragment spanning from exon 10 to exon 11.

Supp. Figure S3. Global map of the allele frequencies of the *WNK1 AluYb8* insertion in studied human population samples from Eurasia and Africa.

Supp. Figure S4. Amino acid alignment of the *WNK1* exon 13 among vertebrates.

Supp. Table S1. Characteristics of the *WNK1* (12p13.33; Ensembl ID: ENSG00000060237) and *WNK4* (17q21.31; Ensembl ID: ENSG00000126562) non-coding regions selected for polymorphism screening.

Supp. Table S2. Polymorphism screening primers for DHPLC

Supp. Table S3. Polymorphism screening primers for DGGE

Supp. Table S4. Sequences of primers and probes used in PCR and RT-PCR

Supp. Table S5. Characteristics of the individuals used for *WNK1* expression profiling in leucocytes

Supp. Table S6. Polymorphisms detected in screened conserved non-coding regions of human *WNK1* and *WNK4* genes

Supp. Table S7. Allele frequencies of the *WNK1 AluYb8* insertion in general human population samples

Supp. Table S8. Association testing of the *WNK1* intron 10 *AluYb8* with hypertension

Supp. Table S9. Linkage disequilibrium (r^2) between the *WNK1 AluYb8* and common SNPs in the *WNK1* gene previously reported in the BRIGHT study

Supp. Table S10. Association of the *WNK1 AluYb8* and SNPs rs11064527, rs12816718, rs956868 with blood pressure in the BRIGHT

For Peer Review

Supp. Text S1. List of additional BRIGHT consortium contributors, additional information for Materials and Methods

Additional MRC the British Genetics of Hypertension (BRIGHT) study consortium members:

Toby Johnson, Stephen J. Newhouse, Philip J. Howard, Abiodun Onipinla, Anna Dominiczak, Nilesh J. Samani, Martin Farrall, Morris Brown, Mark Lathrop, John Connell

Affiliations:

Clinical Pharmacology and Barts and the London Genome Centre, William Harvey Research Institute, Barts and the London School of Medicine, Queen Mary University of London, London, UK (T.J., S.J.N., P.J.H, A. O.); BHF Glasgow Cardiovascular Research Centre, Division of Cardiovascular and Medical Sciences, University of Glasgow, Western Infirmary, Glasgow, UK (A.D.); Department of Cardiovascular Sciences, University of Leicester, Glenfield Hospital, Leicester, UK (N.S.); Department of Cardiovascular Medicine, University of Oxford, Wellcome Trust Centre for Human Genetics, Oxford, UK (M.F.); Clinical Pharmacology and the Cambridge Institute of Medical Research, University of Cambridge, Addenbrooke's Hospital, Cambridge, UK (M.B.); Centre National de Genotypage, Evry, France (M.L.); University of Dundee, Medical School, Ninewells Hospital and Medical School Dundee, UK (J.C.)

Details of the design and experimental condition of the DGGE and DHLPC assays

DGGE primers with GC rich tail at 5´ end of one primer of the primer pair were designed with program MELTingeny 1.0.1 using primer design criteria recommended by Ingeny International. Supplementing the 5' end of the reverse primer with a GC clamp (a sequence of 40 G and C deoxynucleotides) allowed the detection of single-base substitutions present in the fragment.

For Peer Review

DHPLC primers were designed according to manufacturer recommendations (DHPLC; Wave Technologies Inc. USA) using web-based program Primer3 (3). The calculations of the appropriate melting conditions for each of the selected DHPLC PCR fragments we used web based DHPLC Melt Program (http://insertion.stanford.edu/melt.html). The uniqueness of the primer sequence in the human genome were tested and verifying using GenomeTester1.3 (http://bioinfo.ut.ee/genometester/). The DGGE and DHPLC primers are given in Supp. Table S1.

Regions targeted for polymorphism screening were amplified by touchdown method using 100 ng genomic DNA, Smart-Taq Hot DNA polymerase (Naxo OÜ, Tartu, Estonia), GeneAmp PCR System 2700 thermal cycler (Applied Biosystems Inc., USA) and amplification conditions described elsewhere (1). DGGE electrophoresis by INGENY PhorU2 system was performed according to the manufacturer's instructions (Ingeny, Goes, Netherlands). Electrophoresis (16-17 hours) was carried out at 58ºC, amperage of 40 mA and voltage of 100 V. The DHPLC analysis was performed according to the manufacturer's recommendations (Wave Technologies Inc. USA) using pools consisting of the genomic DNA from three individuals. Individual DNA probes exhibiting unusual pattern in the DGGE analysis or forming a DNA-pool indicative to the presence of a polymorphism in the DHPLC analysis were resequenced to determine the exact changes in the genomic sequence.

Recruitment of individuals to the HYPEST study

The HYPEST (n=1823) study has been approved by the Ethics Committee on Human Research of University of Tartu, Estonia. All of the study participants are of Eastern European ancestry,

For Peer Review

have filled a self-administrated epidemiological questionnaire recording to their past and present health and life-style and have given their written informed consent.

The HYPEST sample collection represents a case-cohort study. Essential hypertension patients were recruited by blood pressure specialists during the patients' ambulatory visits or hospitalization at the North Estonia Medical Center, Centre of Cardiology, or at the Cardiology Clinic, Tartu University Hospital, Estonia. The HYPEST healthy control cohort was recruited from among the long-term blood donors across Estonia. All the included donors had no personal history of cardiovascular diseases and had also never been prescribed any relevant medications.

Details of the cDNA synthesis protocol

Total RNA from leucocytes was extracted using LeukoLOCK™ Total RNA Isolation System (Ambion Inc, Austin, Texas, USA) including an optional TURBO™ DNase treatment to degrade genomic DNA. Concentration of extracted RNA was measured with NanoDrop® ND-1000 UV-Vis Spectrophotometer (NanoDrop Technologies, LLC, Wilmington, Delaware, USA).

RNA was reverse transcribed using SuperScript™ III First-Strand Synthesis SuperMix for qRT-PCR (Life Technologies Corporation, Carlsbad, California, USA) in the final volume of 20 µl containing 0,4 µg of RNA, 10 µl of 2X RT Reaction Mix [oligo$(dT)_{20}$ (2.5 µM), random hexamers (2.5 ng/µl), 10 mM MgCl2, and dNTPs], 2µl of RT Enzyme Mix (includes SuperScript™ III RT and RNaseOUT™) and DEPC-treated water. The reaction mixture was incubated at 25ºC for 10 minutes and at 50ºC for 30 minutes. Reaction was terminated at 85ºC at 5 minutes and chilled on ice. cDNA was treated with 2 U of *E.coli* Rnase H at 37ºC for 20 minutes to remove the RNA template from the cDNA:RNA hybrid molecule.

For Peer Review

Detailed conditions of the real-time RT-PCR reactions

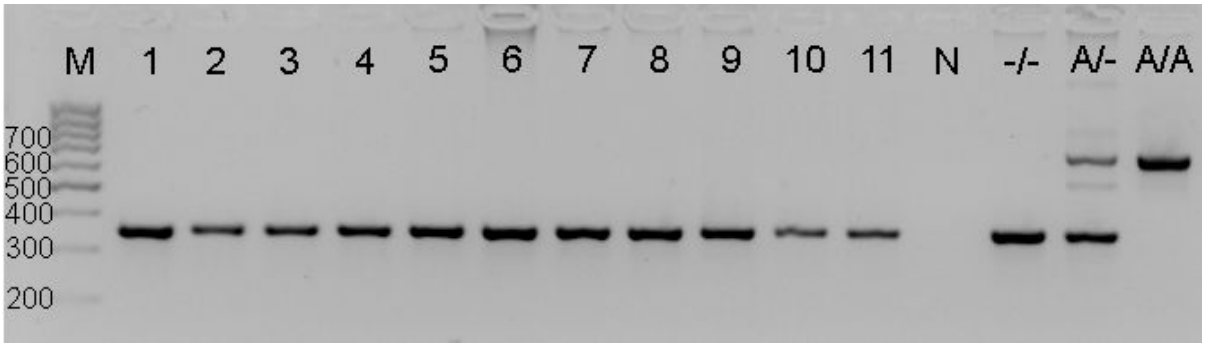
The 20 µl PCR reactions consisted of 2 µl cDNA product 4X dilution, 4 µl of 5X HOT FIREPol® Probe qPCR Mix (Solis BioDyne, Tartu, Estonia) containing HOT FIREPol® DNA Polymerase, 5X Probe qPCR buffer, 15 mM $MgCl_2$, dNTPs and 2,5 µM ROX dye, as well as primers and a probe in following concentration: 400 nM forward and reverse primer, 250 nM probe for spliced transcripts or 1 µl of 20X primer-probe mix for transcript including exon 11 (Hs01018312_m1) and 0.9 µl (ex+11+12 and ex−11−12) or 0.8 µl (ex−11+12) of 20X ready-to-use primer-probe mix for endogenous control *HPRT1*. The PCR protocol was identical in all runs: incubation step at 95ºC for 15' followed by 40 two-step amplification cycles (15'' at 95ºC; 1' at 60ºC).

References

Hallast P, Nagirnaja L, Margus T, and Laan M. 2005. Segmental duplications and gene conversion: Human luteinizing hormone/chorionic gonadotropin beta gene cluster. *Genome Res* 15: 1535-1546.

Rozen S, and Skaletsky H. 2000. Primer3 on the WWW for general users and for biologist programmers. *Methods Mol Biol* 132: 365-386.

For Peer Review



Supp. Figure S1. PCR amplification of the locus for *WNK1* intron 10 *AluYb8* insertion in eleven chimpanzees as resolved by agarose gel (3%) electrophoresis. In all analyzed individuals (numbered 1-11) the resulting PCR product corresponded to the wild-type genotype lacking the *WNK1* intron 10 *AluYb8* insertion. Human genomic DNAs representing alternative genotypes were used as positive controls: wild-type genotype without *AluYb8* insertion (–/–, PCR product 353 bp); heterozygous (A/–) and homozygous (A/A, PCR product 660 bp) carriers of the insertion.
N, negative control; M, marker (Fermentas MassRuler™ DNA Ladder, Low Range)

For Peer Review

Supp. Figure S2. Sequence alignment of the *WNK1* genomic fragment spanning from exon 10 to exon 11. Human chromosomes carrying (Alu +) and not carrying (Alu –) the *AluYb8* insertion are shown in comparison with the respective genomic region in common chimpanzee (*Pan troglodytes*). The sequences of the *WNK1* exons 10 and 11 are shown in blue and the *AluYb8* insertion to human *WNK1* intron 10 in red font. The substitution-based divergence between human wild-type (Alu –) and chimpanzee sequences (excluding species-specific insertion/deletions) is 0%, 0.2% and 1.1% for exon 10 (150 bp), exon 11 (459 bp) and intron 10 (1211 bp), respectively. Sequence alignment was performed using the web-based global alignment program ClustalW2 (http://www.ebi.ac.uk/Tools/clustalw2/).

EXON 10:

Human Alu (+)
CAGCAGGGAATACAGCAGACAGCCCCTCCTCAACAGACAGTGCAGTATTCACTTTCACAG 60
Human Alu (–)
CAGCAGGGAATACAGCAGACAGCCCCTCCTCAACAGACAGTGCAGTATTCACTTTCACAG 60
Chimpanzee
CAGCAGGGAATACAGCAGACAGCCCCTCCTCAACAGACAGTGCAGTATTCACTTTCACAG 60

Human Alu (+)
ACATCAACCTCCAGTGAGGCCACTACTGCACAGCCAGTGAGTCAACCTCAAGCTCCACAA 120
Human Alu (–)
ACATCAACCTCCAGTGAGGCCACTACTGCACAGCCAGTGAGTCAACCTCAAGCTCCACAA 120
Chimpanzee
ACATCAACCTCCAGTGAGGCCACTACTGCACAGCCAGTGAGTCAACCTCAAGCTCCACAA 120

Human Alu (+) **INTRON 10**
GTCTTGCCTCAAGTATCAGCTGGAAAACAG GTAAACTTTTTTTTTTTTTTTAAACAGGTA 180
Human Alu (–)
GTCTTGCCTCAAGTATCAGCTGGAAAACAG GTAAACTTTTTTTTTTTTTTTAAACAGGTA 180
Chimpanzee
GTCTTGCCTCAAGTATCAGCTGGAAAACAG GTAAACTTTTTTTTTTTTTTTAAACAGGTA 180

Human Alu (+)
AACTCTTAATTTCTGAAAGGGTGCTAAAAGGGATTTCCATGTAACTTGTCCTTTCATGTG 240
Human Alu (–)
AACTCTTAATTTCTGAAAGGGTGCTAAAAGGGATTTCCATGTAACTTGTCCTTTCATGTG 240
Chimpanzee
AACTCTTAATTTCTGAAAGGGTGCTAAAAGGGATTTCCATGTAACTTGTCCTTTCATGTG 240

Human Alu (+)
GATAGACTTCTACCTTTTCTTCTAAGGGTAACCAACCCTTGAAGTAGGTTAATCTCATTG 300
Human Alu (–)
GATAGACTTCTACCTTTTCTTCTAAGGGTAACCAACCCTTGAAGTAGGTTAATCTCATTG 300
Chimpanzee
GATAGACTTCTACCTTTTCTTCTAAGGGTAACCAACCCTTGAAGTAGGTTAATCTCATTG 300

Human Alu (+)
CAGAAATGAAGTGAAGATCACCTCATTGGTGCATATGCATTATTTAATGTAAATGGGTAC 360
Human Alu (–)
CAGAAATGAAGTGAAGATCACCTCATTGGTGCATATGCATTATTTAATGTAAATGGGTAC 360
Chimpanzee
CAGAAATGAAGTGAAGATCACCTCATTGGTGCATATGCATTATTTAATGTAAATGGGTAC 360

For Peer Review

Human Alu (+)
GTTACTGACAGCAGTGACAATCCAAAGTTTCACTTCTAGTCTTACCAGTCCAAAATTGAT 420
Human Alu (–)
GTTACTGACAGCAGTGACAATCCAAAGTTTCACTTCTAGTCTTACCAGTCCAAAATTGAT 420
Chimpanzee
GTTACTGACAGCAGTGACAATCCAAAGTTTCACTTCTAGTCTTACCAGTCCAAAATCGAT 420

Human Alu (+)
GTAACAGGTATTAGAAATAACAGGTTAAATAAAGGTGCTCTGGTAACTGAACCTATGACA 480
Human Alu (–)
GTAACAGGTATTAGAAATAACAGGTTAAATAAAGGTGCTCTGGTAACTGAACCTATGACA 480
Chimpanzee
TTAACAGGTATTAGAAATAACAGGTTAAATAAAGGTGCTCTGGTAACTGAACCTATGACA 480

Human Alu (+)
TTCCCTAGAATTGTGTCAGCTATGTCTGATATTCTAGGTTTGAAACTATGCTTTATTTTC 540
Human Alu (–)
TTCCCTAGAATTGTGTCAGCTATGTCTGATATTCTAGGTTTGAAACTATGCTTTATTTTC 540
Chimpanzee
TTCCCTAGAATTGTGTCAGCTATGTCTGATATTCTAGGTTTGAAACTATGCTTTATTTTC 540

Human Alu (+) ***AluYb8* insertion**
ATGTAGATTTCTGCACTGTTTACTGTGTGGAATGTTTTTTTTTTTTTTTTTTT GAGACG 600
Human Alu (–)
ATGTAGATTTCTGCACTGTTTACTGTGTGGAATGTTTTTTTTT- - - - - - - - - - - - - - - - - 583
Chimpanzee
ATGTAGATTTCTGCACTGTTTACTGTGTGGAATGTTTTTTTTTT- - - - - - - - - - - - - - - - - 584

Human Alu (+)
GAGTCTCGCTCTGTCGCCCAGGTCGGACTGCGGACTGCAGTGGCGCAATCTCGGCTCACT 660
Human Alu (–)
- -
Chimpanzee
- -

Human Alu (+)
GCAAGCTCCGCTTCCCGGGTTCACGCCATTCTCCTGCCTCAGCCTCCCGAGTAGCTGGGA 720
Human Alu (–)
- -
Chimpanzee
- -

Human Alu (+)
CTACAGGCGCCCGCCACCGCGCCCGGCTAATTTTTTGTATTTTTAGTAGAGACGGGGTTT 780
Human Alu (–)
- -
Chimpanzee
- -

For Peer Review

Human Alu (+)
CACCTTGTTAGCCAGGATGGTCTCGATCTCCTGACCTCATGATCCACCCGCCTCGGCCTC 840
Human Alu (–)
- -
Chimpanzee
- -

Human Alu (+)
CCAAAGTGCTGGGATTACAGGCGTGAGCCACCGCGCCCGGCC TTTTTTTTCTTAAGTAGT 900
Human Alu (–)
- CTTAAGTAGT 593
Chimpanzee
- CTTAAGTAGT 594

Human Alu (+)
CCTCCTAATCACTTGAGAAGTACCCTGTAGAAGGTATGCTTTTCTGAATGACCTTTTTTC 960
Human Alu (–)
CCTCCTAATCACTTGAGAAGTACCCTGTAGAAGGTATGCTTTTCTGAATGACCTTTTTTC 653
Chimpanzee
CCTCCTAATCACTTGAGAAGTACCCTGTAGAAGTTATGCTTTTCTGAATGACCTTTGTTC 654

Human Alu (+)
AACTTTATTTTTGAGTACTGTTGTGCATGTATGTGCACAGTGTATTGTTACCCTAATATT 1020
Human Alu (–)
AACTTTATTTTTGAGTACTGTTGTGCATGTATGTGCACAGTGTATTGTTACCCTAATATT 713
Chimpanzee
AACTTTATTTTTGAGCACTGTTGTGCATGTATGTGCACAGTGTATTGTTACCCTAATATT 714

Human Alu (+)
TCATCACATCTGAGATGCTATCAACTTTAAAATGTACCATATTTTATGTACCACATTATC 1080
Human Alu (–)
TCATCACATCTGAGATGCTATCAACTTTAAAATGTACCATATTTTATGTACCACATTATC 773
Chimpanzee
TCATCACATCTGAGATGCTATCAACTTTAAAATGTACCA- - - - - - - - - - - - - - CATTATC 760

Human Alu (+)
AGTTATAAAGTAGATCCCAATTTCTCACATGTTAAAATGATCAGGTAGTGGAGAGGGATG 1140
Human Alu (–)
AGTTATAAAGTAGATCCCAATTTCTCACATGTTAAAATGATCAGGTAGTGGAGAGGGATG 833
Chimpanzee
AGTTATAAGGTAGATCCCAATTTCTCACATGTTAAAATGGTCAGGTAGTGGAGAGGGATG 820

Human Alu (+)
ATTGCATCCTAGAATGAGTGAAATATGTGAATTATACCAATTTTTATTGAGCCATGTTGG 1200
Human Alu (–)
ATTGCATCCTAGAATGAGTGAAATATGTGAATTATACCAATTTTTATTGAGCCATGTTGG 893
Chimpanzee
ATTGCATCCTAGAATGAGTGAAATATGTGAATTATACCAATTTTTATTGAGCCATGTTGG 880

Human Alu (+)
ACTTAAGAATTTTAGAAATAATACGAATGAAATCTTAATTCCACTGGGAAGGAATTCATT 1260
Human Alu (–)
ACTTAAGAATTTTAGAAATAATACGAATGAAATCTTAATTCCACTGGGAAGGAATTCATT 953
Chimpanzee
ACTTAAGAATTTTAGAAATAATATGAATGAAATCTTAATTCCACTGGGAAGGAATTCATT 940

For Peer Review

Human Alu (+)
TGAATTTCAGTAACTACTGTAACAGCAGTCGTAACTTTAGTGATAGCATTATGGTATATA 1320
Human Alu (–)
TGAATTTCAGTAACTACTGTAACAGCAGTCGTAACTTTAGTGATAGCATTATGGTATATA 1013
Chimpanzee
TGAATTTCAGTAACTACTGTAATAGCAGTCGTAACTTTAGTGATAGCATTATGGTATATA 1000

Human Alu (+)
AAATATATTCCAATAAAGCTGTTAAAAAAAAAA-GCCAACCCCTTGTCTATGGAAGGGTC 1379
Human Alu (–)
AAATATATTCCAATAAAGCTGTTAAAAAAAAAA-GCCAACCCCTTGTCTATGGAAGGGTC 1072
Chimpanzee
AAATATATTCCAATAAGGCTGTAAAAAAAAAAAAGCCAACCCCTTGTCTATGGAAGGGTC 1060

Human Alu (+)
CTCTTCTATTGCCAAATGCTGAAGCATTAGCAAATATTTCTATGACAAAAGGTGTAGAAC 1439
Human Alu (–)
CTCTTCTATTGCCAAATGCTGAAGCATTAGCAAATATTTCTATGACAAAAGGTGTAGAAC 1132
Chimpanzee
CTCTTCTATTGCCAAATGCTGAAGCATTAGCAAATATTTCTATGACAAAAGGTGTAGAAC 1120

Human Alu (+)
AGTAATAGTCTATTTAGCCTCTTTCTCTCCTGCTCTCCTTTCCATATTTTTATGTGGCAT 1499
Human Alu (–)
AGTAATAGTCTATTTAGCCTCTTTCTCTCCTGCTCTCCTTTCCATATTCTTATGTGGCAT 1192
Chimpanzee AGTAATAGTCTATTTAGCCTCTTTCTCTCCTGCTCTCCTTTCCATATTCTTATGTGGCAT 1180

Human Alu (+)
ATTAACTTAACACTAATGTATGCAGGGTTTTGTTGGTTTGGTGTTTTTTTTTTTTTGTTT 1559
Human Alu (–)
ATTAACTTAACACTAATGTATGCAGGGTTTTGTTGGTTTGGTGTTTTTTTTTTT- - GTTT 1250
Chimpanzee
ATTAACTTAACACTAATGTATGCAGGGTTTTGTTGGTTTGGTGTTTTTTTTGTT- - - - TT 1236

Human Alu (+)
GTTTTTTCCTTCTTTTTGGCTAATACATAAATCTTGCTTTTGGCAGCCTTGCTTTTTTTT 1619
Human Alu (–)
GTTTTTTCCTTCTTTTTGGCTAATACATAAATCTTGCTTTTGGCAGCCTTGCTTTTTTTT 1310
Chimpanzee GTTTTTTCCTTCTTTTTGGCTAATACATAAATCTTGCTTTTGGCAGCCTTGCTATTTTTT 1296

Human Alu (+) **EXON 11**
TTTTTTTTTTTTTTTTAA-GCCTGTCTGTTTTGTTTTTCTTTACCTTCCCAG CTTCCAGT 1678
Human Alu (–)
TTTTTTTTTTTTTTTTAA-GCCTGTCTGTTTTGTTTTTCTTTACCTTCCCAG CTTCCAGT 1369
Chimpanzee
TTTTTTTTTTTTTTTTAAAGCCTGTCTGTTTTGTTTTTCTTTACCTTCCCAG CTTCCAGT 1356

Human Alu (+)
TTCCCAGCCAGTACCAACTATCCAAGGCGAACCTCAGATCCCAGTTGCGACACAACCCTC 1738
Human Alu (–)
TTCCCAGCCAGTACCAACTATCCAAGGCGAACCTCAGATCCCAGTTGCGACACAACCCTC 1429
Chimpanzee
TTCCCAGCCAGTACCAACTATCCAAGGCGAACCTCAGATCCCAGTTGCGACACAACCCTC 1416

For Peer Review

Human Alu (+)
GGTTGTTCCAGTCCACTCTGGTGCTCATTTCCTTCCAGTGGGACAGCCGCTCCCTACTCC 1798
Human Alu (–)
GGTTGTTCCAGTCCACTCTGGTGCTCATTTCCTTCCAGTGGGACAGCCGCTCCCTACTCC 1489
Chimpanzee
GGTTGTTCCAGTCCACTCTGGTGCTCATTTCCTTCCAGTGGGACAGCCGCTCCCTACTCC 1476

Human Alu (+)
CTTGCTCCCTCAGTACCCTGTCTCTCAGATTCCCATATCAACTCCTCATGTGTCTACGGC 1858
Human Alu (–)
CTTGCTCCCTCAGTACCCTGTCTCTCAGATTCCCATATCAACTCCTCATGTGTCTACGGC 1549
Chimpanzee
CTTGCTCCCTCAGTACCCTGTCTCTCAAATTCCCATATCAACTCCTCATGTGTCTACGGC 1536

Human Alu (+)
TCAGACAGGTTTCTCATCCCTTCCCATCACAATGGCAGCTGGCATTACTCAGCCTCTGCT 1918
Human Alu (–)
TCAGACAGGTTTCTCATCCCTTCCCATCACAATGGCAGCTGGCATTACTCAGCCTCTGCT 1609
Chimpanzee
TCAGACAGGTTTCTCATCCCTTCCCATCACAATGGCAGCTGGCATTACTCAGCCTCTGCT 1596

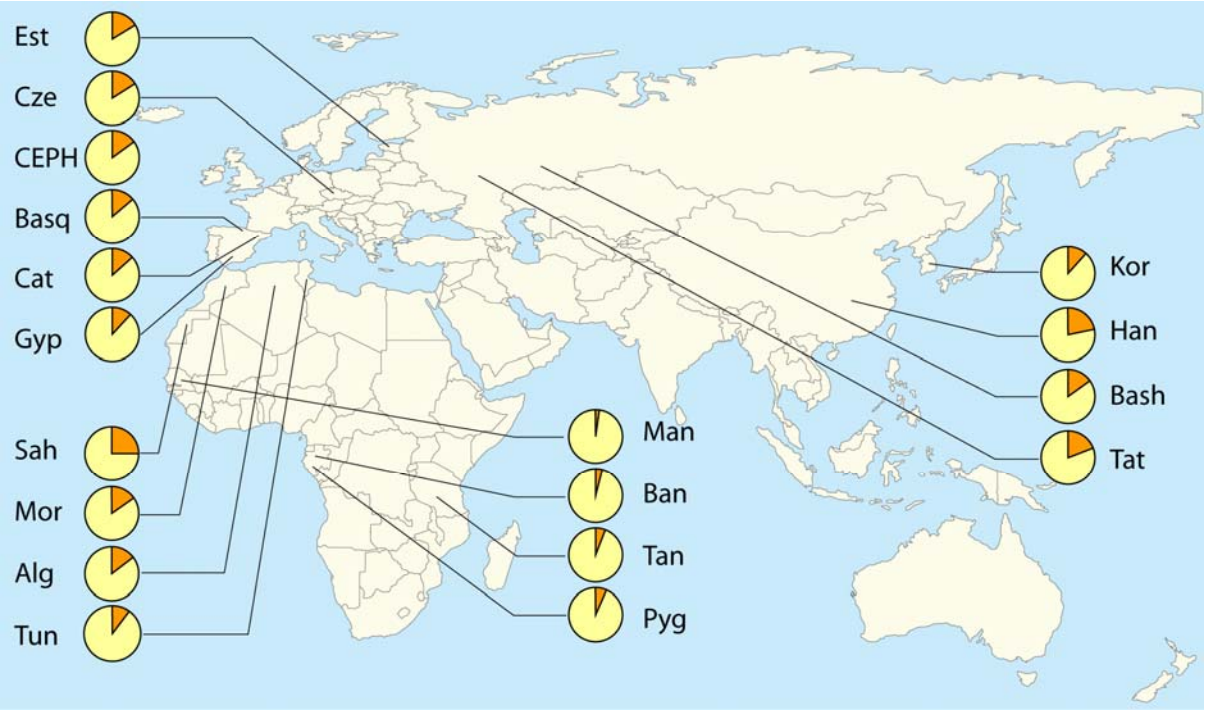
Human Alu (+)
CACGTTGGCTTCATCTGCTACAACAGCTGCGATCCCGGGGGTATCAACTGTGGTTCCTAG 1978
Human Alu (–)
CACGTTGGCTTCATCTGCTACAACAGCTGCGATCCCGGGGGTATCAACTGTGGTTCCTAG 1669
Chimpanzee
CACGTTGGCTTCATCTGCTACAACAGCTGCGATCCCGGGGGTATCAACTGTGGTTCCTAG 1656

Human Alu (+)
TCAGCTTCCAACCCTTCTGCAGCCTGTGACTCAGCTGCCAAGTCAGGTTCACCCACAGCT 2038
Human Alu (–)
TCAGCTTCCAACCCTTCTGCAGCCTGTGACTCAGCTGCCAAGTCAGGTTCACCCACAGCT 1729
Chimpanzee
TCAGCTTCCAACCCTTCTGCAGCCTGTGACTCAGCTGCCAAGTCAGGTTCACCCACAGCT 1716

Human Alu (+)
CCTACAACCAGCAGTTCAGTCCATGGGAATACCAGCTAACCTTGGACAAGCTGCTGAGGT 2098
Human Alu (–)
CCTACAACCAGCAGTTCAGTCCATGGGAATACCAGCTAACCTTGGACAAGCTGCTGAGGT 1789
Chimpanzee
CCTACAACCAGCAGTTCAGTCCATGGGAATACCAGCTAACCTTGGACAAGCTGCTGAGGT 1776

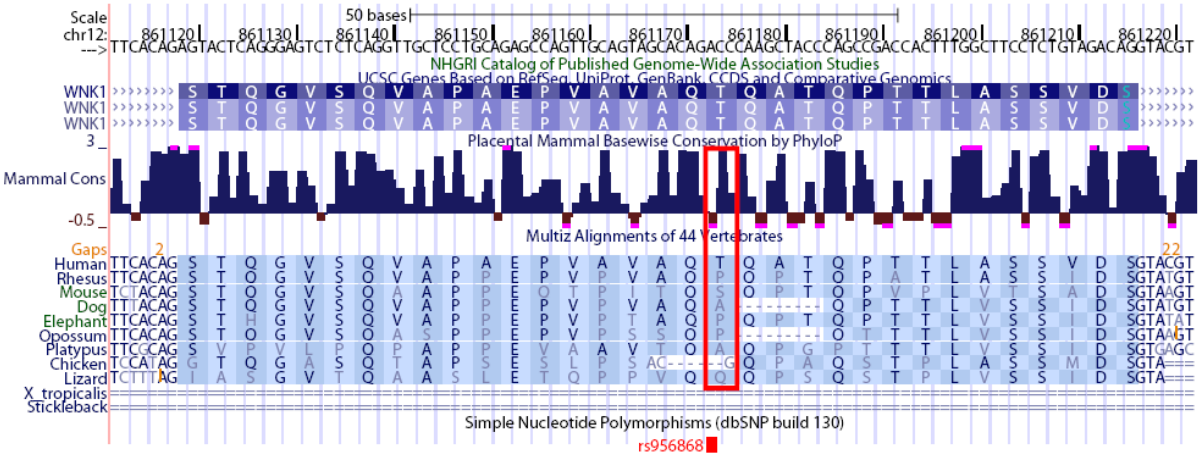
Human Alu (+)
TCCACTTTCCTCTGGAGATGTTCTGTACCAG 2129
Human Alu (–)
TCCACTTTCCTCTGGAGATGTTCTGTACCAG 1820
Chimpanzee
TCCACTTTCCTCTGGAGATGTTCTGTACCAG 1807

For Peer Review



Supp Figure S3. Global map of the allele frequencies of the *WNK1 AluYb8* insertion in studied human population samples from Eurasia and Africa. Pie charts represent the allele frequency of *AluYb8* insertion (orange sector). Lines drawn from the pies point to the geographic affiliation of the respective study population.
Est, Estonians; Cze, Czech; CEPH, CEPH (Utah) families with ancestry from Northern and Western Europe; Basq, Basque; Cat, Catalans; Gyp, Spanish Gypsies; Sah, Saharawi; Mor, Moroccans; Alg, Algerians; Tun, Tunisians; Man, Mandenkalu; Ban, Gabon Bantus; Tan, Tanzanians; Pyg, Gabon Pygmies; Kor, Koreans; Han, Chinese Han; Bash, Bashkir; Tat, Tatars.

eer Review



Supp. Figure S4. Amino acid alignment of the *WNK1* exon 13 among vertebrates. Snapshot from UCSC Genome Browser (http://genome.ucsc.edu/; version NCBI36/hg18, Human March 2006) showing the human peptide sequence coded by the *WNK1* exon 13 in comparison to the orthologous region in other species. Red rectangle highlights the human variable position rs956868 coding for two alternative amino acids Proline or Threonine, and exhibiting low evolutionary conservation among vertebrates. The reference sequence of the hg18 (March 2006) carries the minor allele coding for the Threonine (T).
P – proline, S – Serine, A – Alanine, Q – Glutamine, G – Glycine

For Peer Review

Supp. Table S1. Characteristics of the *WNK1* (12p13.33; Ensembl ID: ENSG00000060237) and *WNK4* (17q21.31; Ensembl ID: ENSG00000126562) non-coding regions selected for polymorphism screening

No.	*WNK4/WNK1* CNS regions	Start[a] position	End[a] position	Sequence type	CNR length (bp)	Rn%[b]	Mm%[c]	DHPLC detected variations; rs	DGGE detected variations; rs	Variation position (dbSNP build130)	Reason for exclusion
WNK4											
1	33_WNK4_44	38188209	38188301	non-coding	93	70.3	69.8	ND	none	-	two different Tm melting points for DHPLC
2	33_WNK4_62	38193583	38193673	non-coding	91	68	70.3	none	none	-	
3	33_WNK4_92	38202314	38202452	UTR	139	72.6	72.7	none	none	-	
4	33_WNK4_97	38202700	38202870	non-coding	171	80.7	75.4	A/c; novel	none	38202858	
5	33_WNK4_32	38186084	38186188	non-coding	105	77.5	74.8	ND	ND	-	inappropriate primer Tm
6	33_WNK4_89	38202123	38202229	non-coding	107	82.4	80.6	ND	ND	-	inappropriate primer Tm
7	33_WNK4_91	38202233	38202312	UTR	80	80.5	79.5	none	ND	-	two different Tm melting points for DGGE
8	33_WNK4_100	38202944	38203178	non-coding	235	79.8	85.8	none	ND	-	two different Tm melting points for DGGE
9	33_WNK4_120	38210504	38210693	non-coding	190	82.8	74.9	none	ND	-	two different Tm melting points for DGGE
10	33_WNK4_122	38210713	38210829	non-coding	117	74.5	74.9	none	ND	-	two different Tm melting points for DGGE
11	33_WNK4_127	38211778	38211946	non-coding	169	75.6	74.3	none	ND	-	inappropriate primer Tm for DGGE
WNK1											
12	34_WNK1_90	831113	831231	non-coding	119	85.8	93.4	none	none	-	
13	34_WNK1_139	846970	847070	non-coding	101	87.6	71.6	none	none	-	
14	34_WNK1_155	858084	858176	non-coding	93	69.9	80.6	*AluYb8*; novel	*AluYb8;* novel	858212	
15	34_WNK1_167	860427	860561	non-coding	135	79.5	79.4	ND	none	-	two different Tm melting points for DHPLC
16	34_WNK1_168	860562	860692	non-coding	131	79.5	79.4	none	none	-	
17	34_WNK1_208	877345	877471	non-coding	127	75	73.3	none	none	-	
18	34_WNK1_222	883173	883364	non-coding	192	72.7	76.1	A/g; rs10849582	none	883437	
19	34_WNK1_227	883843	883984	non-coding	142	78.5	76.6	none	none	-	
20	34_WNK1_250	889267	889379	UTR	113	77	75.7	A/g; novel	none	889252	
21	34_WNK1_12	731433	731594	non-coding	162	77.3	79	ND	ND	-	two different Tm melting points for DHPLC and DGGE
22	34_WNK1_37	756329	756526	non-coding	198	75.1	75.5	ND	none	-	genome test negative for DHPLC
23	34_WNK1_51	789691	789890	non-coding	200	74.1	75.4	ND	ND	-	genome test negative for DHPLC

For Peer Review

No.	*WNK4/WNK1* CNS regions	Start[a] position	End[a] position	Sequence type	CNR length (bp)	Rn%[b]	Mm%[c]	DHPLC detected variations; rs	DGGE detected variations; rs	Variation position (dbSNP build130)	Reason for exclusion
24	34_WNK1_53	790344	790469	non-coding	126	92.6	96.2	G/a; rs36092349	ND	790218	two different Tm melting points for DGGE
25	34_WNK1_63	808723	808848	non-coding	126	72.6	0	C/t; rs72648627	ND	808848	two different Tm melting points for DGGE
26	34_WNK1_67	809338	809429	non-coding	92	0	70.8	ND	ND	-	two different Tm melting points for DHPLC and DGGE
27	34_WNK1_81	826028	826198	non-coding	171	82.8	82.8	ND	none	-	genome test negative for DHPLC
28	34_WNK1_97	832069	832178	non-coding	110	77	73.7	ND	C/g; rs36052085	832130	genome test negative for DHPLC
29	34_WNK1_117	841698	841796	non-coding	99	73.7	75.8	G/a; rs12825084	ND	841868	two different Tm melting points for DGGE
30	34_WNK1_120	842050	842208	non-coding	159	72.6	75	ND	none	-	genome test negative for DHPLC
31	34_WNK1_146	848285	848553	non-coding	269	87.6	89	ND	ND	-	genome test negative for DHPLC
32	34_WNK1_151	850776	850900	non-coding	125	66.9	69	ND	none	-	genome test negative for DHPLC
33	34_WNK1_161	858897	858999	non-coding	103	75	70.2	ND	ND	-	two different Tm melting points for DHPLC and DGGE
34	34_WNK1_165	860037	860147	non-coding	111	73.9	73	none	ND	-	two different Tm melting points for DGGE
35	34_WNK1_174	861501	861617	non-coding	117	67.9	73.1	ND	ND	-	two different Tm melting points for DHPLC and DGGE
36	34_WNK1_214	880862	881152	non-coding	291	86.8	85.9	none	ND	-	two different Tm melting points for DGGE
37	34_WNK1_218	882302	882503	non-coding	202	75.2	74.1	none	ND	-	two different Tm melting points for DGGE
38	34_WNK1_224	883491	883608	non-coding	118	74.6	76.6	ND	none	-	genome test negative for DHPLC
39	34_WNK1_246	888231	888518	UTR	288	80.5	80.1	none	ND	-	inappropriate primer Tm
40	34_WNK1_259	890715	890782	UTR	68	89.4	85.3	ND	ND	-	two different Tm melting points for DHPLC and DGGE

[a] Start and end positions of the non-coding regions according to Human Genome March 2006 (NCBI Build 36/hg18) assembly.
[b] Sequence identity with *Rattus norvegicus*.
[c] Sequence identity with *Mus musculus*.
CNR - conserved non-coding region; ND - not done

For Peer Review

Supp. Table S2. Polymorphism screening primers for DHPLC

No.	*WNK4*/*WNK1* CNS regions	DHPLC forward primer 5´-3´	DHPLC reverse primer 5´-3´	PCR product (bp)	DHPLC Tm
WNK4					
2	33_WNK4_62	TGACCTATCTCCCTCCTTGTGAA	ATGACTGTCAGAGCAGACACTGG	323	60°
3	33_WNK4_92	CACCTTGGAGGTTTCTTCATCAC	GCTGCTAATGTGGTGTGAGCATA	395	60°
4	33_WNK4_97	CACCACATTAGCAGCAACCAA	TCTGTCCTCTACCCACAAGAACA	398	60°
7	33_WNK4_91	GATGTTGGCAGGATGGTGAG	ATGAAGATGGTAGGCAGCAGAAC	354	61°
8	33_WNK4_100	CTTGTCTTTCACCCTTCATCCAC	CTGTGCAGTGTTTGAGCCTACAG	383	60°
9	33_WNK4_120	CCTAGGTTAGGTGATACAACTCATGG	GCAGCATCTGGCCATAGAATAA	385	58°
10	33_WNK4_122	CCTGAATTGCCAGCATACCTATC	TCTCTTAGGATCCTCCCATCTCC	378	56°
11	33_WNK4_127	TTGCTTGAACCTGGAAGGTGTA	GGTAGAGGTAGAGAAACCTCCATCTG	372	60°
WNK1					
12	34_WNK1_90	GTCTGACTTTGATGTCTGCCTTC	CCTCCTCACACAGAATACACCAT	381	55°
13	34_WNK1_139	GGCCTCTAACATATCACAGGGTAA	ATGCACAACTGGGTAGTTCTGTG	429	54°
14	34_WNK1_155	GGGTAACCAACCCTTGAAGTAGG	GGGTACTTCTCAAGTGATTAGGAGGA	353	56°
16	34_WNK1_168	CCAGGAGGGAGTTTAGCACAAG	AATGATTGCTGAGAGCACAGGAG	399	58°
17	34_WNK1_208	TGTCTCCAGCTGTATGTGACCTG	CACCGTGCCTAACCAATGACT	486	54°
18	34_WNK1_222	GGCCCTACATCCTTAGAAGTCACA	GATCCTGTCCCACTTCCTGTTTC	413	57°
19	34_WNK1_227	TGTCTCACCTCCTGCTTTCTCTT	TGTCTGTCACAAGCTAGGCTCAC	368	57°
20	34_WNK1_250	CATGTGGAAGGAAATCGTAGGTC	ACGTCTGTTCAAGCCTAGCTCAA	339	58°
24	34_WNK1_53	AAGGGACGGAAGATGTGCAA	CTCTTCACCATTATGGAGCTGGA	398	59°
25	34_WNK1_63	GCCTGTTTCTTGGAAATGCCTA	CAGCTACACTCTTGGCATTAACACTC	380	57°
29	34_WNK1_117	CATGAACAGGCACATTCTACAGG	GCTTCATGTAGGCAGTTCAGCTA	296	59°
34	34_WNK1_165	GTCCTGTCTCTTCTGTTGGAACC	TGTGGGAAAGTACCCAGGTGTAG	348	60°
36	34_WNK1_214	CTTGATTGGTAAGGAGGGCAAAC	GCCCAAACTTCTGAGGGTAACA	487	59°
37	34_WNK1_218	TCTGATGAGTCAGTGCCTACCTG	GGAGCCAATCCATTTCTAATGC	373	58°
39	34_WNK1_246	GGAACTGCCTCCTTGCAGAAT	AGAGCTCCACCCACTCGAAATTA	452	60°

For Peer Review

Supp. Table S3. Polymorphism screening primers for DGGE

No.	*WNK4/WNK1* CNS regions	DGGE forward primer 5´-3´	DGGE reverse primer 5´-3´	PCR product (bp)
WNK4				
1	33_WNK4_44	CGCCCGCCGCGCCCCGCGCCCGGCCCGCCGCCCCCGCCC GCGCCCCCGGCGGCAAGTAATCCCATTAAC CCCA	ATAGAAGCGGACGATGTTGGGGT	278
2	33_WNK4_62	CGCCCGCCGCGCCCCGCGCCCGGCCCGCCGCCCCCGCCC GTTTTTTTTTTGACCTATCTCCCTCCTTGTGAAA	GCCGACCCCAACTTGTACCAGAACAT	322
3	33_WNK4_92	GGGCCCCGGGCCACCTTGGAGGTTTCTTCATCAC	CGCCCGCCGCGCCCCGCGCCCGGCCCGCCGCCCCCGCCC GTTTTTTTTTGATTGCTTAGTGGTTGACTCTTGGC	374
4	33_WNK4_97	AACGCCGTGAATCCATATGCCAT	CGCCCGCCGCGCCCCGCGCCCGGCCCGCCGCCCCCGCCC GTTTTGAACATACTGTA ACTCTACTGC	384
WNK1				
12	34_WNK1_90	TTCCTTATGCTGTGGGCTGATGTA	CGCCCGCCGCGCCCCGCGCCCGGCCCGCCGCCCCCGCCC GGCATTCTCAAGTTAACAA	341
13	34_WNK1_139	CGCCCGCCGCGCCCCGCGCCCGGCCCGCCGCCCCCGCCC GTTTTTTTTTTTTTTTTAGGGAAAGTTAGTTCAATAGTT	CCGAATGAAATCAATCTGTGGGCTCT	329
14	34_WNK1_155	ATTGCAGAAATGAAGATC	CGCCCGCCGCGCCCCGCGCCCGGCCCGCCGCCCCCGCCC GTTTTTTTTTTTTTTACAGG GTACTTCTCAAGTGATTAGG	379
15	34_WNK1_167	CGCCCGCCGCGCCCCGCGCCCGGCCCGCCGCCCCCGCCC GTTTTTTTTTTTTTTTTTTT CACTACATCCTCCCAGCAAG	AGGACCCCACATTGGCATTTCATTC	263
16	34_WNK1_168	CGCCCGCCGCGCCCCGCGCCCGGCCCGCCGCCCCCGCCC GCGTATATATGTAATTGGACAAC	AGCCAGCATCTTACCAAGAGTGT	433
17	34_WNK1_208	CGCCCGCCGCGCCCCGCGCCCGGCCCGCCGCCCCCGCCC GTTTTTTTTTTTTAGACTTAGTGAATACAATGGC	GGCCGGTTGCAATTATGACTAAGTCA	248
18	34_WNK1_222	CGGGCGTGGTATTACCCCTGTTGTTCAA	CGCCCGCCGCGCCCCGCGCCCGGCCCGCCGCCCCCGCCC GGGCTATTATATTTCCACATCCC	415
19	34_WNK1_227	TGTCTCACCTCCTGCTTTCTCTT	CGCCCGCCGCGCCCCGCGCCCGGCCCGCCGCCCCCGCCC GTTTTTTTTTTGAATAGGATACAGAGATAGATGG	328
20	34_WNK1_250	CGCCCGCCGCGCCCCGCGCCCGGCCCGCCGCCCCCGCCC GTTTTTTTTTTTCAGTGGTTATGCCAAAGGGAAAT	GGGCCCACCATTCTTTTAAGTTGTCCAA	245
22	34_WNK1_37	GTGAACTATAGAGAATTGACC	CGCCCGCCGCGCCCCGCGCCCGGCCCGCCGCCCCCGCCC GTTTTTTTTTTTTTTTTTTTGCACAGATGCTGACTAATTTT	324

For Peer Review

No.	*WNK4*/*WNK1* CNS regions	DGGE forward primer 5´-3´	DGGE reverse primer 5´-3´	PCR product (bp)
27	34_WNK1_81	GGTTTCCAATGGGATCATAGGGT	CGCCCGCCGCGCCCCGCGCCCGGCCCGCCGCCCCCGCCC GTTTTTTTTTTTTTTCAGTAAAAAGGTAGGGTGAATT	334
28	34_WNK1_97	AGTGGAATGAGAGGGCTGTAT	CGCCCGCCGCGCCCCGCGCCCGGCCCGCCGCCCCCGCCC GGGGACATGTAATAGAAAGCTG	340
30	34_WNK1_120	CGCCCGCCGCGCCCCGCGCCCGGCCCGCCGCCCCCGCCC GTTTTTTTTTTTTTTTGGCGAGAAGAGTTTTTCAAAA	TCCAACACAGACTTTAAAAGGCA	283
32	34_WNK1_151	TCATCTTCCCAACCCATACAA	CGCCCGCCGCGCCCCGCGCCCGGCCCGCCGCCCCCGCCC GTTTTTTTTTGAGATTTTTCTGAAACACAG	264
38	34_WNK1_224	CCGGCCCTACTGTGGTGGGGTTGCTAATGCAT	CGCCCGCCGCGCCCCGCGCCCGGCCCGCCGCCCCCGCCC GCAGATATAGTTCTACAGAATG	342

For Peer Review

Supp. Table S4. Sequences of oligonucleotide primers and probes used in PCR and RT-PCR reactions

Primer ID	Sequence 5´- 3´	5´modification	3´modification
A. Amplification of genomic area covering exons 10 and 11			
WNK1_Seq1_F	TGGGGTGAGGGAGATAATTGGGTG		
WNK1_Seq_R	ACTCTGTGGTGCCCTCTTTTGCT		
B. Sequencing primers			
WNK1_Seq3_F	TCAGGTAGTGGAGAGGGATGATTG		
WNK1_Seq4_F	ACCCCTTGTCTATGGAAGGGTCC		
WNK1_Seq5_F	CCTTCCAGTGGGAGAGCCGC		
WNK1_Alu_F	CGGCCTCCCAAAGTGCTGGG		
C. RT-PCR primers and probes			
WNK1_Exon10_F	CCACAAGTCTTGCCTCAAGTATCA		
WNK1_Exon12_R	GGAAGAGGGAGCAATATTTGAATC		
WNK1_Exon13_R	TGCAGGAGCAACCTGAGAGA		
Ex–11+12_probe	TGGAAAACAGGGCTTC	[a]6-FAM™	[b]MGBNFQ
Ex–11–12_probe	CTGGAAAACAGAGTACTC	6-FAM™	MGBNFQ

[a] Phosphoramidite
[b] Minor groove binding non-fluorescent quencher

For Peer Review

Supp. Table S5. Characteristics of the individuals used for *WNK1* expression profiling in leucocytes

Genotype	Gender	Age	SBP	DBP	BMI
–/–	female	40	100	60	23.0
	female	36	105	70	21.3
	female	46	118	83	23.0
Alu/–	female	51	118	75	27.0
	female	43	113	73	25.6
	female	47	127	82	20.8
Alu/Alu	female	35	156	100	21.8
	female	64	190	110	31.6
	female	57	165	100	26.3

SBP – Systolic Blood Pressure; DBP – Diastolic Blood Pressure;
BMI – Body Mass Index

For Peer Review

Supp. Table S6. Polymorphisms detected in screened conserved non-coding regions of human *WNK1* and *WNK4* genes

Analyzed regions[a]			Detected variants				
Gene	Location	Length of fragment (bp)	Position of the variant[b]	Alleles[c]	HYPEST (hom/het)[d]	CADCZ (hom/het)[d]	Validation/ rs-number[b]
WNK1	Intron 1	398bp	Chr.12; 790218	G/a	1/3	0/1	rs36092349
WNK1	Intron 3	380bp	Chr.12; 808848	C/t	0/1	0/0	rs72648627
WNK1	Intron 4	340bp	Chr.12; 832130	C/g	na/6	na/6	rs36052085
WNK1	Intron 8	296bp	Chr.12; 841868	G/a	0/0	0/2	rs12825084
WNK1	Intron 10	353bp	Chr.12; 858212	D/i	1/7	3/9	novel
WNK1	Intron 26	413bp	Chr.12; 883437	A/g	0/0	0/4	rs10849582
WNK1	3´UTR	339bp	Chr.12; 889252	A/g	0/1	0/0	novel
WNK4	5´UTR	398bp	Chr.17; 38202858	A/c	0/0	0/1	novel

[a] Details of the analyzed regions and conditions of the DHPLC/DGGE assays are given in Supp. Table S1.
[b] Positions of the variants and rs-numbers of the validated SNPs according to the dbSNP build 130; Human Genome March 2006.
[c] Major (capital letter) and minor (lower case letter) alleles according to the coding strand.
[d] Number of homozygotes/ heterozygotes (hom/het) of each identified polymorphism among the screened HYPEST (n=22) and CADCZ (n=24) patients.
D/i – deletion/insertion variant of *WNK1 AluYb8*
na – not applicable as DGGE method was incapable to distinguish CC and GG homozygotes

For Peer Review

Supp. Table S7. Allele frequencies of the *WNK1 AluYb8* insertion in general human population samples

Region	Population	Number of individuals				*AluYb8* allele frequency
		Alu/Alu[a]	Alu/–[b]	–/–[c]	total	
Europe	Estonians	4	25	71	100	16.5%
	Czechs	1	14	35	50	16.0%
	CEPH (Utah) families	2	5	23	30	15.0%
	Basques	0	14	36	50	14.0%
	Catalans	0	11	30	41	13.4%
	Spain Gypsies	0	12	38	50	12.0%
Asia	Chinese Han	2	7	16	25	22.0%
	Koreans	1	8	34	43	11.6%
	Tatars	1	16	30	47	19.1%
	Bashkirs	2	10	35	47	14.9%
North-Africa	Moroccans[d]	2	22	60	84	15.5%
	Saharawi	6	13	31	50	25.0%
	Tunisians	1	8	39	48	10.4%
	Algerians	0	14	34	48	14.6%
Sub-Saharan Africa	Mandenkalu	0	1	23	24	2.1%
	Tanzanians	0	2	15	17	5.9%
	Bantus	0	4	46	50	4.0%
	Pygmies	0	7	43	50	7.0%

[a] Homozygotes and [b] Heterozygotes for the *WNK1 AluYb8* insertion.
[c] Wild-type genotype without the *WNK1 AluYb8* insertion.
[d] Samples were collected from five different cities in Morocco (Oujda, Casablanca, Rabat, Chefchaouen, Nador).
Mandenkalu and Chinese Han samples were purchased from the HGDP-CEPH Human Genome Diversity Cell Line Panel (http://www.cephb.fr/HGDP-CEPH-Panel/). Korean samples were kindly shared by Dr. Woo Chul Moon (Good-Gene Inc. Seoul, Korea).

For Peer Review

Supp. Table S8. Association testing of the *WNK1* intron 10 *AluYb8* with hypertension

	Sample size: total and stratified by genotype (Alu/Alu; Alu/–; –/–)		*AluYb8* allele frequency (%)		Association testing[a]	
Sample	Cases	Controls	Cases	Controls	OR (CI 95%)	*P*-value
HYPEST[b]: essential hypertension/ normotensive controls						
all	673 (21/197/455)	601 (13/148/440)	17.70	14.51	1.24 (0.91, 1.69)	1.66×10^{-1}
women	445 (16/140/289)	439 (11/109/319)	19.28	14.97	1.17 (0.81, 1.69)	4.03×10^{-1}
men	228 (5/57/166)	162 (2/39/121)	14.63	13.27	1.42 (0.79, 2.52)	2.38×10^{-1}
CADCZ[b]: CAD patients with hypertension/ normotensive controls						
all	266 (10/71/185)	480 (9/129/342)	17.11	15.31	1.23 (0.89, 1.71)	2.14×10^{-1}
women	86 (5/21/60)	251 (5/61/185)	18.02	14.14	1.37 (0.83, 2.29)	2.22×10^{-1}
men	180 (5/50/125)	229 (4/68/157)	16.67	16.60	1.14 (0.74, 1.75)	5.57×10^{-1}
BRIGHT[c]: extreme family based hypertension/ normotensive controls						
all	2242 (54/544/1644)	1639 (43/391/1205)	14.54	14.55	1.00 (0.88, 1.14)	9.95×10^{-1}
women	1320 (40/319/961)	1002 (29/249/724)	15.11	15.32	0.99 (0.84, 1.16)	8.85×10^{-1}
men	922 (14/225/683)	637 (14/142/481)	13.72	13.34	1.02 (0.83, 1.26)	8.66×10^{-1}

[a] Logistic regression (additive model, age and gender as covariates) were used to test association with HYP.
[b] Cases: subjects under antihypertensive treatment or untreated subjects SBP ≥160 mmHg and/or DBP ≥100 mmHg; Controls: subjects with SBP ≤140 mmHg and DBP≤90 mmHg, receiving no antihypertensive medication.
[c] Cases: patients originating from severely hypertensive families, under antihypertensive treatment and with BP readings ≥150/100 mmHg based on one reading or ≥145/95 mmHg based on the mean of three readings; Controls: normotensive subjects with BP <140/90 mmHg, receiving no antihypertensive medication.
Alu/–, Alu/Alu *WNK1 AluYb8* hetero- and homozygote carriers; –/– wild-type homozygotes; CAD, coronary artery disease; OR, odds ratio; CI, confidence interval

For Peer Review

Supp. Table S9. Linkage disequilibrium (r^2) between the *WNK1 AluYb8* and common SNPs in the *WNK1* gene previously reported in the BRIGHT study (data from Newhouse et al., 2009)

SNP	Location	Position[a]	MAF[b]	r^2
rs7137188	5´	724592	0.48	0.0003
rs1468326	5´	727762	0.11	0.0029
rs3088353	5´	732901	0.46	0.1908
rs11064519	Intron 1	741472	0.34	0.0044
rs2369402	Intron 1	748925	0.21	0.0476
rs2107612	Intron 1	758581	0.27	0.0723
rs2107613	Intron 1	758689	0.23	0.0543
rs11608756	Intron 1	760094	0.39	0.2561
rs11064524	Intron 1	760163	0.25	0.0423
rs11064527	Intron 1	761286	0.16	**0.8216***
rs11064536	Intron 1	775843	0.17	0.0345
rs765250	Intron 1	778544	0.31	0.0800
rs12314329	Intron 1	779992	0.08	0.0176
rs11611246	Intron 4	809741	0.22	0.0294
rs6489756	Intron 4	834767	0.47	0.1800
rs12816718	Intron 6	840061	0.15	**0.9213***
rs2286007	Exon 8	841552	0.07	0.0140
rs11611231	Intron 8	844587	0.10	0.0095
rs4980973	Intron 9	853307	0.13	0.0248
rs4980974	Intron 9	855289	0.44	0.1277
AluYb8	Intron 10	858213	0.15	1
rs880054	Intron 10	858819	0.41	0.1286
rs956868	Exon 13	861173	0.14	**0.9753***
rs953361	Intron 22	872068	0.40	0.0878
rs2301880	Intron 23	874098	0.24	0.0602
rs2286026	Intron 26	885434	0.41	0.1247
rs2286028	Intron 26	885730	0.20	0.0371
rs2277869	Intron 26	887171	0.16	0.0233
rs11571461	3´	896580	0.07	0.0150

[a] Positions of the SNPs according to the dbSNP build 130; Human Genome March 2006.
[b] MAF, minor allele frequency
*** r^2>0.8**

Table reference:

Newhouse S, Farrall M, Wallace C, Hoti M, Burke B, Howard P, Onipinla A, Lee K, Shaw-Hawkins S, Dobson R, Brown M, Samani NJ, Dominiczak AF, Connell JM, Lathrop GM, Kooner J, Chambers J, Elliott P, Clarke R, Collins R, Laan M, Org E, Juhanson P, Veldre G,

For Peer Review

Viigimaa M, Eyheramendy S, Cappuccio FP, Ji C, Iacone R, Strazzullo P, Kumari M, Marmot M, Brunner E, Caulfield M, Munroe PB. 2009. Polymorphisms in the WNK1 gene are associated with blood pressure variation and urinary potassium excretion. *PLoS One 4*: e5003.

For Peer Review

Supp. Table S10. Association of the *WNK1 AluYb8* and SNPs rs11064527, rs12816718, rs956868 with blood pressure in the BRIGHT

Trait		rs11064527 (Intron 1)		rs12816718 (Intron 6)		*AluYb8* indel (Intron 10)		rs956868 (Exon 13) Thr1316Pro (NP_001171914)	
SBP		Beta (SE)[b]	*P*-value	Beta (SE)	*P*-value	Beta (SE)	*P*-value	Beta (SE)	*P*-value
	all	1.16 (0.49)	1.67×10^{-2}	1.15 (0.50)	2.18×10^{-2}	1.50 (0.51)	3.01×10^{-3}	1.47 (0.52)	4.90×10^{-3}
	women	1.26 (0.67)	5.83×10^{-2}	1.22 (0.68)	7.55×10^{-2}	1.60 (0.68)	1.91×10^{-2}	1.48 (0.70)	3.46×10^{-2}
	men	1.00 (0.67)	1.40×10^{-1}	1.07 (0.70)	1.27×10^{-1}	1.27 (0.73)	8.12×10^{-2}	1.41 (0.75)	6.11×10^{-2}
DBP									
	all	0.61 (0.34)	7.64×10^{-2}	0.52 (0.35)	1.40×10^{-1}	0.78 (0.36)	2.99×10^{-2}	0.72 (0.37)	5.15×10^{-2}
	women	1.07 (0.46)	1.89×10^{-2}	0.93 (0.47)	4.98×10^{-2}	1.22 (0.47)	9.48×10^{-3}	1.06 (0.48)	2.88×10^{-2}
	men	-0.14 (0.51)	7.85×10^{-1}	-0.10 (0.52)	8.43×10^{-1}	-0.03 (0.55)	9.49×10^{-1}	0.09 (0.56)	8.70×10^{-1}

The association testing was performed using the BRIGHT normotensive individuals (n=1421) genotyped in the current study for the *AluYb8* insertion, and in the previous study for rs11064527, rs12816718, rs956868 (Newhouse et al., 2009). Statistical analysis was identical to the conditions described in Materials and Methods for the association testing with quantitative variables, SBP and DBP (linear regression, additive model, age and gender as covariates).
SBP, DBP, systolic and diastolic blood pressure; SE, standard error

For Peer Review