nature genetics

Rare independent mutations in renal salt handling genes contribute to blood pressure variation

Weizhen Ji[1,2,8], Jia Nee Foo[1,2,8], Brian J O'Roak[1], Hongyu Zhao[1,3], Martin G Larson[4,5], David B Simon[1,2], Christopher Newton-Cheh[4], Matthew W State[1,6], Daniel Levy[4,7] & Richard P Lifton[1,2]

The effects of alleles in many genes are believed to contribute to common complex diseases such as hypertension. Whether risk alleles comprise a small number of common variants or many rare independent mutations at trait loci is largely unknown. We screened members of the Framingham Heart Study (FHS) for variation in three genes—*SLC12A3 (NCCT), SLC12A1 (NKCC2)* and *KCNJ1 (ROMK)*—causing rare recessive diseases featuring large reductions in blood pressure. Using comparative genomics, genetics and biochemistry, we identified subjects with mutations proven or inferred to be functional. These mutations, all heterozygous and rare, produce clinically significant blood pressure reduction and protect from development of hypertension. Our findings implicate many rare alleles that alter renal salt handling in blood pressure variation in the general population, and identify alleles with health benefit that are nonetheless under purifying selection. These findings have implications for the genetic architecture of hypertension and other common complex traits.

Hypertension affects 1 billion people world-wide and is a major contributor to death from stroke, myocardial infarction, end-stage renal disease and congestive heart failure. Although epidemiologic studies have demonstrated high heritability of blood pressure variation (h^2 of long-term blood pressure $\sim$0.6)[1], identification of the responsible genes has been difficult owing to the trait's complexity.

Whether common variants or, alternatively, many independent rare mutations will account for the contributions of specific genes is unknown. Recent large genome-wide association studies have identified a number of common variants that contribute to metabolic traits including diabetes[2–4], coronary artery disease[5,6], and obesity[7]; notably, these collectively explain only a small fraction of interindividual risk[2]. These studies have thus far failed to identify any statistically significant loci for blood pressure or hypertension[2,8], raising the possibility that rare independent variants account for a large fraction of blood pressure variation, as would be expected if trait alleles are under purifying selection[9].

The study of rare mendelian traits has identified more than 20 genes in which mutations impart large effects on blood pressure; these predominantly act by changing net renal salt reabsorption[10,11]. Several of these disorders, exemplified by Bartter's and Gitelman's syndromes, are recessive traits that lower blood pressure, raising the question of whether the more prevalent heterozygous mutations in these genes might commonly affect the trait. Bartter's syndrome features massive renal salt wasting and hypotension, often resulting in neonatal death. It is caused by recessive loss-of-function mutations in any of four genes required for normal renal NaCl reabsorption[12–15]. These include the Na-K-2Cl cotransporter gene *SLC12A1* and the inward rectifier K^+ channel gene *KCNJ1*. Gitelman's syndrome is a less severe salt-wasting disease caused by recessive loss-of-function mutations in the Na-Cl cotransporter gene *SLC12A3* (ref. 16).

The prevalences of Bartter's and Gitelman's syndromes (estimated at $\sim$1 per million and 1 per 40,000, respectively)[16,17] suggest that heterozygous disease alleles should be present in at least 1% of the population (likely to be an underestimate, owing to early lethality of Bartter's syndrome), motivating evaluation of these genes in a large, well characterized cohort. We examined *SLC12A1*, *SLC12A3* and *KCNJ1* in the Framingham Heart Study (FHS) offspring cohort. This cohort, comprising 5,124 subjects (3,125 with DNA), has been followed for up to 35 years with periodic evaluation of cardiovascular risk factors and other traits, permitting stable assessments of quantitative trait values over time.

RESULTS

Mutations in *SLC12A3, SLC12A1* and *KCNJ1* in FHS

We screened all coding exons and flanking intronic sequences of *SLC12A3*, *SLC12A1* and *KCNJ1* for DNA sequence variants in the 1,985 unrelated subjects and 1,140 relatives with available DNA

[1]Department of Genetics, Yale University School of Medicine, 333 Cedar Street, New Haven, Connecticut 06510, USA. [2]Howard Hughes Medical Institute, Yale University School of Medicine, 295 Congress Avenue, New Haven, Connecticut 06510, USA. [3]Department of Epidemiology and Public Health, Yale University School of Medicine, 60 College Street, New Haven, Connecticut 06510, USA. [4]National Heart, Lung, and Blood Institute's Framingham Heart Study, 73 Mount Wayte Avenue, Suite 2, Framingham, Massachusetts 01702, USA. [5]Department of Mathematics and Statistics, Boston University, 111 Cummington Street, Boston, Massachusetts 02215, USA. [6]Child Study Center, Yale University School of Medicine, 230 South Frontage Road, New Haven, Connecticut 06520, USA. [7]National Heart, Lung and Blood Institute, 31 Center Drive, Bethesda, Maryland 20892, USA. [8]These authors contributed equally to this work. Correspondence should be addressed to R.P.L. (richard.lifton@yale.edu).

Received 2 November 2007; accepted 15 February 2008; published online 6 April 2008; corrected online 13 April 2008 (details online); doi:10.1038/ng.118

samples. This constituted evaluation of ~24 Mb of diploid sequence. The sensitivity of variant detection was high, as we detected all 18 validated SNPs in these genes and their allele frequencies in FHS were similar to previous frequency estimates (**Supplementary Table 1** online). In addition, eight subjects with Gitelman's syndrome were blindly included in the screen and their 16 known mutations were correctly identified. Variants were confirmed by two independent methods (**Supplementary Fig. 1** online), and variants in the final set were further verified by independent amplification of the original sample. These variants recurred among siblings at expected mendelian frequencies (32 of 60 siblings), providing assurance of proper sample tracking and indicating that few variants are *de novo* mutations.

A total of 138 different coding sequence variants were found in 2,492 FHS subjects, whereas the remainder showed only wild-type sequence (**Supplementary Table 2** online). The variants included 46 different synonymous substitutions found in 2,235 subjects, 89 different missense variants found in 1,062 subjects, one nonsense mutation and two frameshift mutations.

Given the rarity of Bartter's and Gitelman's syndromes, the vast majority of these variants must not cause loss of function. We first looked for well documented disease-causing mutations that had been shown biochemically to cause loss of function of the encoded cotransporter or channel[18–22]. This identified ten different biochemically proven loss of function mutations found in 23 subjects (**Table 1**).

We next sought to identify additional functional variants using phylogenetic conservation as an indication of amino acid positions that are subject to purifying selection. We characterized orthologs and paralogs of each gene from diverse vertebrate and invertebrate species (10–12 orthologs and 6–14 paralogs for each gene; see Methods) and determined the conservation of each amino acid in each encoded protein. The proportion of amino acids that were completely conserved among orthologs ranged from 18% to 24% for each gene; nonconserved positions had an average of 3.5 ± 1.1 different variants, demonstrating substantial branch length in this set.

Similarly, variants that are under strong purifying selection are expected to be found at low frequency in the population; we used allele frequency <0.001 as a threshold (using a threshold of <0.01 would have yielded the identical set of variants). This criterion is supported by empiric data in our population, as the previously identified functional mutations that are found in FHS all have allele frequencies ≤ 0.0005.

As a critical test of the utility of these criteria, we determined their ability to identify disease-causing mutations in 148 unrelated subjects with Gitelman's syndrome and 144 with Bartter's syndrome (**Supplementary Table 3** online). These criteria identified 62 different disruptive mutations or missense mutations at highly conserved positions that accounted for 189 mutant alleles in subjects with Gitelman's syndrome. From these data we calculate that these criteria provide 77% sensitivity and 90% specificity for identification of functional mutations. Further, because 35 of the independent mutations in the Gitelman's cohort were only found once, saturation for disease-causing mutations has not been approached. Similar results were found for *SLC12A1* and *KCNJ1* (disruptive mutations or missense mutations at highly conserved positions were found in 86% and 65% of Bartter's alleles, respectively). It is worth noting that we obtained similar results if we confined the analysis to the 128 Gitelman's syndrome subjects of European ancestry (criteria showed 78% sensitivity and 90% specificity), excluding a strong effect of genetic background on the spectrum of functional mutations. These findings provide direct experimental validation that these criteria were sensitive and specific for identification of functional mutations

Table 1 Proven and inferred mutations in *SLC12A3*, *SLC12A1* and *KCNJ1* in FHS

Variant	Phylogenetic conservation of wild-type residue	Allele frequency in unrelated FHS subjects	Allele frequency in European-ancestry subjects with disease	Biochemical loss of function[18–22]	SIFT/PolyPhen
SLC12A3					
E112X[a]	Disruptive	0.00025	0.0039	Yes	NA
L155F	Complete	0.00025	0	Unknown	++/+
S188F	Complete	0.00025	0	Unknown	++/++
A232T	Complete	0.00050	0	Unknown	++/+
R261C[a]	Complete	0.00050	0	Unknown	++/++
R399L[a]	Vertebrate	0.00025	0.0078	Yes	++/++
G439S[a]	Complete	0.00050	0.0391	Yes	++/++
F495L	Complete	0.00025	0	Unknown	++/++
P560H[a]	Complete	0.00025	0	Unknown	++/++
G613S[a]	Complete	0.00025	0.0039	Yes	++/++
G741R[a]	Complete	0.00050	0.1211	Yes	++/++
G779E	Complete	0.00025	0	Unknown	++/++
R861C[a]	Vertebrate	0.00025	0.0195	Unknown	++/+
R964Q[a]	Vertebrate	0.00025	0.0117	Yes	++/++
G989R[a]	Vertebrate	0.00025	0.0156	Yes	–/+
SLC12A1					
T235M	Complete	0.00025	0	Unknown	++/++
P254A	Complete	0.00025	0	Unknown	++/++
R302FS[a]	Disruptive	0.00025	0.0069	Yes	NA
R302W	Complete	0.00025	0	Unknown	++/++
P348L	Complete	0.00025	0	Unknown	++/++
N399S	Complete	0.00050	0	Unknown	++/++
L505V	Complete	0.00025	0	Unknown	++/+
P569H	Complete	0.00025	0	Unknown	++/++
Y1070C	Complete	0.00025	0	Unknown	++/++
P1083A	Complete	0.00025	0	Unknown	++/++
KCNJ1					
P166S	Complete	0.00025	0	Unknown	++/+
R169H	Complete	0.00025	0	Unknown	++/++
R193P	Complete	0.00025	0	Unknown	++/++
H251Y	Complete	0.00050	0	Yes	++/++
T313FS[a]	Disruptive	0.00025	0.0104	Yes	NA

[a]Previously reported as disease-causing.
FS, frameshift; X, premature termination; ++, probably damaging (PolyPhen) or deleterious (SIFT); +, possibly damaging (PolyPhen) or deleterious with low confidence (SIFT); –, benign or tolerated; NA, not applicable.

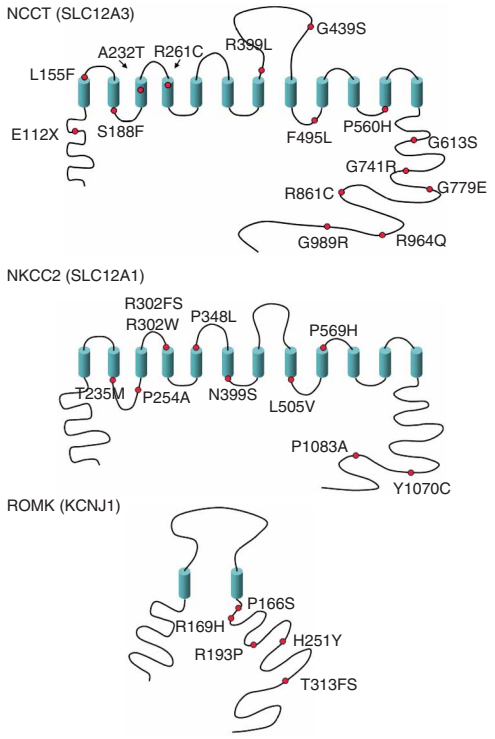

Figure 1 Proven and inferred mutations in *SLC12A3*, *SLC12A1* and *KCNJ1* in the Framingham Heart Study offspring cohort. The predicted structure of each protein is shown, with locations of putatively functional missense variants found in FHS, indicated in single letter code (X, premature termination; FS, frameshift).

in these genes (**Supplementary Fig. 2** online; **Supplementary Note** online; see Methods).

Applying these validated criteria—complete conservation and rare allele frequency—to identify mutations under purifying selection in FHS, we identified 19 additional missense variants found in 25 subjects; all were at completely conserved positions and had allele frequencies ≤ 0.0005. In addition, we included one further variant in *SLC12A3* found once in FHS, R861C, which had previously been reported as disease-causing[23,24] and was found in four of our European-ancestry subjects wtih Gitelman's syndrome (homozygous in one). This nonconservative substitution is at a position conserved among all vertebrates.

Reduced blood pressure among mutation carriers

The final set of functional variants included 30 different mutations in 49 subjects (**Table 1** and **Fig. 1**). Virtually all of these were nonconservative substitutions and were predicted to be damaging by PolyPhen[25], SIFT[26] and PANTHER[27] but represented a small subset of the variants predicted as functional by these programs. All of these mutations had allele frequencies ≤ 1 per 2,000 in FHS.

The standardized age- and sex-adjusted, long-term average systolic and diastolic blood pressure in the entire FHS offspring cohort was determined as defined previously[1]. Eighty percent of mutation carriers had long-term systolic blood pressure (SBP) below the mean of the cohort ($P = 0.001$; **Fig. 2**); results were similar when separately considering the biochemically defined (78% below mean) or inferred functional mutations (81% below mean; **Table 2**). There was an excess of mutation carriers among those in the lowest decile of SBP compared with the highest decile; this excess becomes increasingly significant as one compares the lowest to the highest 20%, 30%, 40% and 50% of the distribution (**Table 3**). This result is consistent with the expectation that the power to detect the impact of alleles of moderate effect is increased as one analyzes the complete distribution rather than simply the extremes[28].

Quantitatively, the mean long-term SBP among mutation carriers was 6.3 mm Hg lower than the mean of the cohort ($P = 0.0009$); similarly reduced blood pressures were seen in separate analyses of subjects with biochemically proven mutations (6.8 mm Hg reduction) and the inferred functional mutations (5.9 mm Hg reduction; **Table 2**). Similar results were found for diastolic blood pressure (DBP), with mean effect of –3.4 mm Hg ($P = 0.003$). Effects were not significantly different among the three genes studied, though there was a trend toward larger effect for *KCNJ1*. There were no significant differences in the magnitude of effects in males and females, and there was no significant difference in body mass index between carriers and noncarriers. Altered blood pressure was not found among carriers of other classes of rare or common synonymous or nonsynonymous variants (**Supplementary Table 4** online).

Mutation carriers showed reduced blood pressure from the earliest age measured to the last. Mean SBP was 5.7 mm Hg lower at age 40 ($P = 0.02$), 6.4 mm Hg at 50 ($P = 0.01$) and 9.0 mm Hg at 60 ($P = 0.0002$) (**Fig. 2c**). Blood pressure effects were significant and of similar magnitude for biochemically proven and inferred mutations (for example, –9.8 and –8.6 mm Hg, respectively, at age 60; **Table 2**). Although less than half the cohort had values measured after age 60, the magnitude of the effects remained large (SBP 11.5 mm Hg lower; $P = 0.03$). Similar effects were seen on DBP (**Fig. 2d**).

As an independent test of the effect of these mutations, we compared the blood pressures of siblings in FHS who were discordant for mutations. This within-family test of linkage is independent of the association tests. Under the null hypothesis, carrier and noncarrier siblings should have blood pressures that are not significantly different from one another. Among 43 sibling pairs discordant for mutations, mutation carriers had mean SBP 6.6 mm Hg lower than their noncarrier sibs ($P = 0.009$), providing further evidence of the impact of the carrier state on blood pressure. (Sibships with biochemically proven mutations accounted for 29 of these discordant sib pairs, and these mutation carriers showed SBP 9.2 mm Hg lower than noncarriers; $P = 0.002$). Similarly, the blood pressure variance among carrier-carrier pairs was lower than the variance among carrier-noncarrier pairs ($P = 0.01$). Finally, among the 15 sibling pairs concordant for mutations, in 12, both sibs had long-term blood pressures below the mean of the cohort (**Fig. 3**; $P = 0.002$), and their mean SBP was significantly lower than the mean value of noncarrier pairs (11.5 mm Hg reduction, $P = 0.0001$). Eleven of these pairs were concordant for biochemically proven mutations, and these showed a mean reduction of 11.6 mm Hg ($P = 0.001$).

Reduced prevalence of hypertension among mutation carriers

Mutation carriers were significantly protected from development of hypertension (SBP or DBP greater than 140 or 90 mm Hg, respectively, and/or treatment with medication for hypertension). Kaplan-Meier analysis of the likelihood of developing hypertension by age 60 showed a 59% reduction in the risk of hypertension among mutation carriers compared with noncarriers (log-rank $P < 0.003$; 95% confidence interval 23–71%) (**Fig. 4a**). Evidence of protection was seen at each age (**Fig. 4b**), and males and females showed no significant differences in the reduction in risk conferred by the carrier

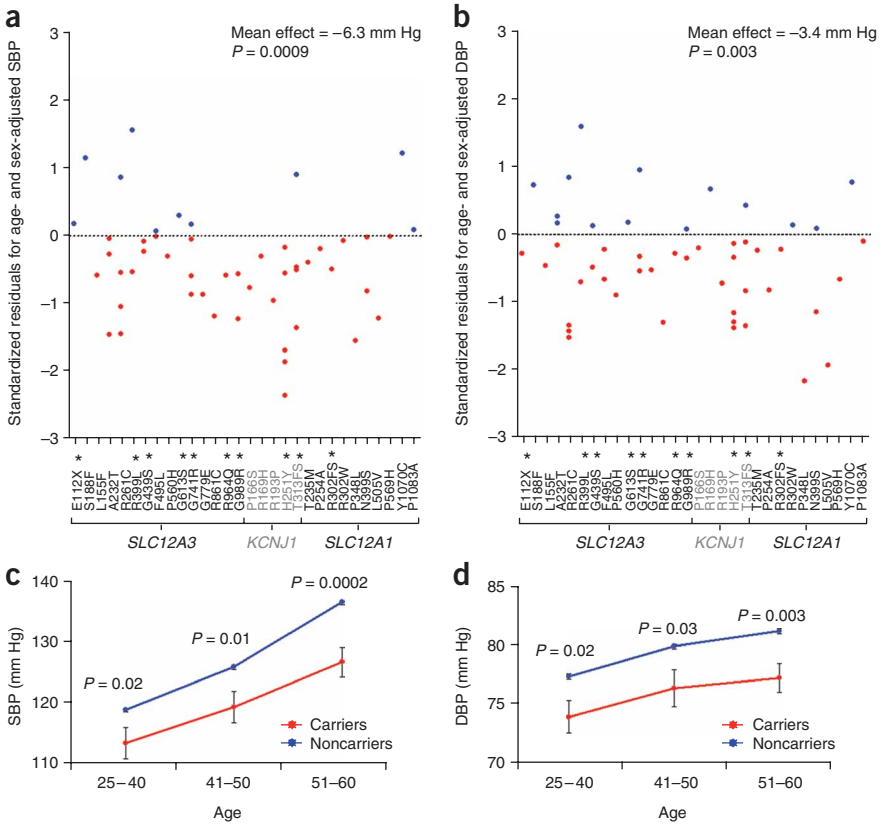

Figure 2 Heterozygous mutations in *SLC12A3*, *SLC12A1* and *KCNJ1* lower blood pressure. (**a**,**b**) Distributions of standardized residual values for age- and sex-adjusted, long-term average systolic (**a**) and diastolic (**b**) blood pressures among carriers of putative functional mutations. Adjusted values for the cohort have a mean value of zero and a s.d. of 1. Mean effect size and parametric *P* values are indicated. (**c**,**d**) Mean and s.e.m. of systolic (**c**) and diastolic (**d**) blood pressures among mutation carriers (*n* ranges from 31 to 43 in different decades) and noncarriers ($n = 2{,}094$–$2{,}753$) at the last examination in different age groups; *P* values compare values of mutation carriers and noncarriers by *t*-tests. Asterisks, biochemically proven loss-of-function mutations. Blue and red data points correspond to subjects with value above and below the population mean, respectively.

state (males $n = 25$, 60% reduction in risk; females $n = 24$, 57% reduction in risk). Carriers of biochemically proven and inferred mutations showed similar reductions in hypertension risk (45% and 68%, respectively; **Table 2**).

DISCUSSION

A large previous body of work has demonstrated that homozygous loss-of-function mutations in genes whose products mediate or regulate renal salt reabsorption result in reduced blood pressure in humans. These observations begged the question of whether the more prevalent heterozygous state for mutations in these same genes might have substantial effects in the population. We found that the carrier state for rare functional mutations in *SLC12A3*, *SLC12A1* and *KCNJ1* reduced blood pressure in FHS participants. Each of these genes governs renal salt handling, and homozygous loss of function mutations in them lower blood pressure by reducing salt reabsorption; heterozygous mutations in *SLC12A3* have also been shown to increase renal salt loss[29]. We infer that the heterozygous mutations we identified lower blood pressure through their effects on renal salt reabsorption. Notably, these results establish the role of altered renal salt handling in blood pressure variation in a general population.

The functional significance of identified mutations was strongly supported by comparative genomics, mendelian genetics, and biochemical evidence. Nearly half the mutation carriers (23 of 49) had mutations that have been proven by biochemical assay to cause loss of function, and the remainder were inferred to be functional by rigorous genetic and genomic criteria. Although it is possible that some of our identified mutations might not cause loss of function, our empiric

Table 2 Effects of mutations on blood pressure phenotypes

		All mutation carriers ($n = 49$)	*P* value	Carriers with biochemically proven mutations ($n = 23$)	*P* value	Carriers with genetic/genomically inferred mutations ($n = 26$)	*P* value
Percentage below long-term mean SBP		80	0.001	78	0.05	81	0.009
Mean effect of mutations on:	Long-term average SBP	−6.3 mm Hg	0.0009	−6.8 mm Hg	0.03	−5.9 mm Hg	0.01
	SBP at age 60	−9.0 mm Hg	0.0002	−9.8 mm Hg	0.02	−8.6 mm Hg	0.002
	SBP difference between carrier and noncarrier siblings	−6.6 mm Hg	0.009	−9.2 mm Hg ($n = 29$ pairs)	0.002	−1.3 mm Hg ($n = 14$ pairs)	0.79
	Mean difference in SBP from FHS mean for carrier-carrier sibling pairs	−11.5 mm Hg	0.0001	−11.6 mm Hg ($n = 11$ pairs)	0.001	−11.2 mm Hg ($n = 4$ pairs)	0.06
Kaplan-Meier risk of hypertension in mutation carriers compared with noncarriers		0.41	0.003	0.55	0.13	0.32	0.007

Table 3 Distribution of functional mutation carriers within the population distribution for longitudinal SBP

	Bottom 10%	Top 10%	Bottom 20%	Top 20%	Bottom 30%	Top 30%	Bottom 40%	Top 40%	Bottom 50%[a]	Top 50%
SLC12A3 carriers (%)	3 (11.1)	1 (3.7)	6 (22.2)	3 (11.1)	10 (37.0)	3 (11.1)	13 (48.2)	5 (18.5)	20 (74.1)	7 (25.9)
SLC12A1 carriers (%)	2 (18.2)	0 (0)	2 (18.2)	1 (9.1)	3 (27.3)	1 (9.1)	5 (45.4)	1 (9.1)	9 (81.8)	2 (18.3)
KCNJ1 carriers (%)	4 (33.3)	0 (0)	5 (41.7)	1 (8.3)	6 (50.0)	1 (8.3)	9 (75.0)	1 (8.3)	11 (91.7)	1 (8.3)
All carriers (%)	9 (18.4)	1 (2.0)	13 (26.5)	5 (10.2)	19 (38.8)	5 (10.2)	27 (55.1)	7 (10.2)	39 (79.6)	10 (20.4)
Noncarriers	304	312	613	621	920	934	1,225	1,245	1,633	1,443
Fisher's exact test	$P = 0.02$		$P = 0.09$		$P = 0.006$		$P = 0.0008$		$P = 0.0002$	

[a]Based on numbers of carriers and noncarriers above and below the population mean of 0.

data indicate that this should be an uncommon event; moreover, inclusion of such neutral or gain-of-function alleles should bias toward the null, rather than toward a false-positive, result. Finally, separate analysis of each of these two groups demonstrated significant effects on blood pressure in both, and the magnitudes of these effects were not significantly different from one another (**Table 2**). The phenotypic effects of these mutations were significant by tests of association with long-term and age-specific blood pressure for both systolic and diastolic blood pressure, as well as by Kaplan-Meier assessment for risk of development of hypertension. In addition, the within-family test of linkage, contrasting blood pressures of siblings, provided independent evidence of the significant impact of these mutations on blood pressure.

At least 1 in 64 FHS members carries a functional mutation in one of these three genes. This is likely to be an underestimate, given the stringent criteria applied and the expectation that ~15% of deleterious mutations lie outside the coding regions[30]. The effects of the carrier state are relatively large, similar in magnitude to those of clinically used antihypertensive agents[31]; these mutations reduced the risk of hypertension at age 60 by nearly 60% and, based on epidemiologic observations in the Framingham cohort, would be predicted to reduce the risk of stroke by 40% (refs. 32,33) and acute coronary syndrome by 15% (refs. 34,35). These effects are thus relevant to individual risk prediction and underscore the importance of these findings for public health.

The estimated prevalence of *SLC12A3* mutations in unrelated members of FHS was 0.48%, convergent to the estimated population frequency of Gitelman's syndrome of ~1 per 40,000. In contrast, the allele frequencies for *SLC12A1* and *KCNJ1* mutations were estimated to be 1/360 and 1/670, respectively, which leads to an estimated frequency of Bartter's syndrome due to these two genes of ~1/100,000, which is much higher than the observed prevalence of Bartter's syndrome in the population[17]. The likely explanation is early lethality of the homozygous state for these mutations. Unlike Gitelman's syndrome, Bartter's syndrome is well recognized to be an early lethal disease, with death during both fetal development and the neonatal period, and with few survivors beyond age 10 owing to myriad complications and development of renal failure[36–38].

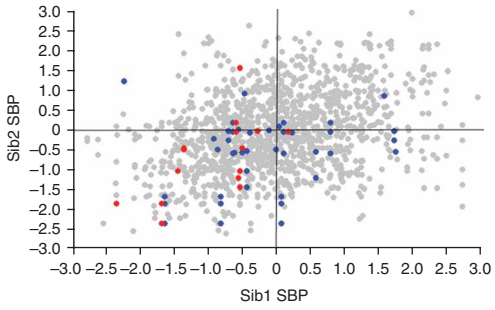

Figure 3 Correlation of blood pressures among sibling pairs in Framingham Heart Study. The distribution of standardized SBP residuals of all 1,369 sibling pairs in the FHS cohort is shown. The blood pressures of sibling pairs concordant for putative functional mutations (red), discordant sibling pairs (blue) and those with no mutations (gray) are indicated. For discordant pairs, the sibling carrying the mutation (sib2) is represented on the *y*-axis.

Similar findings of multiple rare mutations in genes for mendelian recessive forms of HDL level variation[39] and rare variants contributing to altered triglyceride levels[40] have been reported, suggesting that such findings will not be infrequent. It is axiomatic that for genes subject solely to purifying selection, rare independent mutations, and not common functional variants, will predominate; these considerations indicate practical limitations to genome-wide association studies. These findings suggest that the resequencing of genes validated by studies of rare mendelian traits or other methods and eventually the resequencing of all genes or entire genomes will make important contributions to understanding the inherited contributions to common traits. These findings do not, however, exclude potential effects of common variants on these or any other traits.

In contrast to the case of rare variants in HDL, although homozygous mutations in each of the genes we studied are likely to impair reproductive fitness, the heterozygous state is expected to confer health benefit in postreproductive ages. This expectation is supported both by epidemiologic studies of the relationship of blood pressure and mortality and by randomized controlled trials which document reduced mortality among patients treated with thiazide diuretics, pharmacologic inhibitors of NCCT (encoded by *SLC12A3*)[41]. These findings thus define a set of alleles that are inferred to reduce morbidity and mortality late in life, but are nonetheless under strong purifying selection because of adverse effects of the homozygous state. As this cohort ages, it will be of future interest to determine whether cardiovascular morbidity and mortality is in fact reduced among carriers.

We have shown that the combination of high conservation and low allele frequency provides both high sensitivity and high specificity for functional mutations in these genes (**Supplementary Fig. 2**). In contrast, prediction programs such as SIFT often overestimate functional alleles because they fail to take into account allele frequencies. In our dataset, SIFT predicted 43 different functional variants in 340 cohort members, and the effects of these variants were not significant ($P = 0.15$ for SBP); the lack of specificity of these prediction programs for functional mutations is well recognized[42]. The results of this study underscore the importance of recognizing functional variants within a large background of neutral variants. The combination of phylogenetic conservation and rare allele frequency we used was effective, and probably conservative—these criteria will not capture variants that are present at higher allele frequencies owing to balancing selection or nonequilibrium states. These results suggest that similar criteria may

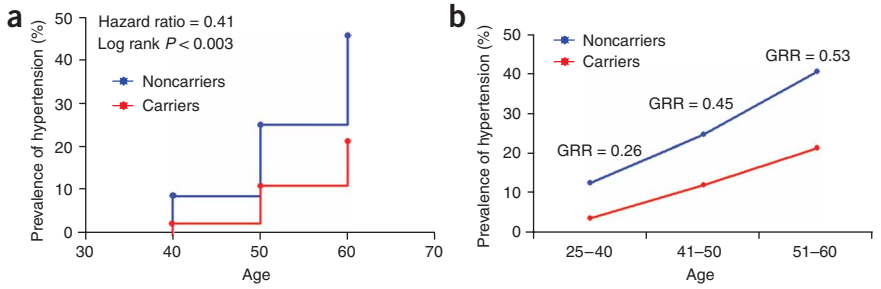

Figure 4 Reduced prevalence of hypertension among mutation carriers. (**a**) Kaplan-Meier plot of time to onset of hypertension in the 49 carriers and 3,076 noncarriers. (**b**) Prevalence of hypertension at the last exam within ages 25–40, 41–50 and 51–60 years for mutation carriers and noncarriers. The genotype relative risk (GRR) for mutation carriers is shown.

be more broadly useful in identifying functional mutations from raw sequence data.

It is worth noting that in our study, sequencing the entire cohort substantially increased the significance of our result compared with what would have been seen had we merely sequenced the tails of the distribution (**Table 3**). This is expected if the effect size attributable to mutations is moderate, rather than large[28]; this point should be carefully considered in the design of future studies. Similarly, as expected for variants with moderate effect size, the distribution of blood pressures remains quite variable among carriers, presumably reflecting influences of other genetic and environmental factors.

The three genes studied were selected in part because NCCT and NKCC2 (encoded by *SLC12A1*) are targets of the most commonly used diuretics, the thiazides and furosemide, respectively. Despite the much greater salt wasting seen in individuals homozygous for mutations in *SLC12A1*, heterozygous mutations in these two genes result in similar reductions in blood pressure. Dosage compensation at *SLC12A1* has been seen in heterozygous null mice, providing a potential explanation[43]. The study of ROMK (encoded by *KCNJ1*) was motivated in part by its potential as a new antihypertensive diuretic agent that might not produce the hypokalemia seen with the above diuretics[11]; the carrier state for *KCNJ1* shows effects at least as large as those seen among carriers of *SLC12A3* and *SLC12A1* mutations, providing encouragement for further investigation.

Finally, these findings have implications for the genetic architecture of blood pressure variation. We estimate that ~100 million subjects world-wide harbor loss-of-function mutations in these genes, and that these subjects are significantly protected from hypertension and its consequences; the overall prevalence of hypertension is inferred to be reduced by about 1% owing to their effects. Because these three genes comprise only a small fraction of those in which mutations are known to alter blood pressure[10], and because there are likely to be many more genes yet to be discovered, it seems probable that the combined effects of rare independent mutations will account for a substantial fraction of blood pressure variation in the population.

METHODS

Human subjects. Clinical data and DNA samples from 3,125 subjects in the offspring cohort of the Framingham Heart Study, as well as 292 subjects with Gitelman's and Bartter's syndrome, were used in this study (**Supplementary Note**). The study was approved by the Yale Human Investigation Committee and the Framingham DNA Committee.

Identification of genetic variation. We amplified the coding exons and their flanking splice sites of *SLC12A3*, *KCNJ1* and *SLC12A1* from DNA samples using specific primers to direct PCR (58 amplicons in total). Multiplex PCR was carried out using HotStart Taq polymerase (Qiagen) using 30 ng DNA in 5-μl reactions (**Supplementary Table 5** online). The PCR products were then diluted, heat denatured, and slowly cooled, permitting heteroduplex formation, as recommended by the manufacturer. We identified sequence variants by temperature-gradient capillary electrophoresis using a Spectrumedix instrument[44]. The PCR failure rate averaged ~1% across all amplicons. The DNA sequence of identified variants was determined by direct DNA sequencing of both strands using an ABI377 instrument; mutations in the final functional variant set were confirmed by independent amplification and sequencing from the original DNA sample.

Longitudinal blood pressure traits. We calculated long-term average SBP and DBP, including adjustment for the effects of age, sex and treatment for hypertension, as previously described[1]. In brief, we used SBP readings taken at exams between 25–75 years of age and DBP readings taken between 25–54 years of age (because DBP declines with age, beginning around age 55; refs. 1,45) in the calculation of the longitudinal readings. Treatment for hypertension was taken into account by calculating an adjusted residual[1]. These long-term average values were adjusted for mean age in a linear regression, done separately for males and females. The standardized residuals were used as quantitative traits. Additive heritability estimates were calculated using the program SOLAR. The heritability estimate (h^2) for long-term average SBP was 0.66 and for DBP was 0.60, each of which was higher than that for any single examination estimate. For analyses at ages 40, 50, 60 and 70, trait values at the last examination within each age range (25–40, 41–50, 51–60 and 61–70) were used. Numbers of individuals in each age group were, for ages 25–40, 31 carriers and 2,094 noncarriers; 41–50, 43 carriers and 2,753 noncarriers; 51–60, 38 carriers and 2,373 noncarriers; and 61–70, 22 carriers and 1,287 noncarriers. Mean ages of subjects in each age range were, for ages 25–40, 37.3 ± 2.6; 41–50, 47.5 ± 2.2; 51–60, 56.8 ± 2.6; and 61–70, 66.1 ± 2.8. Males and females were equally represented in each group.

Orthologs and paralogs. Full length orthologous protein sequences from a range of animal species (including primates, rodents, birds, fish, *Drosophila melanogaster*, sea urchin (*Strongylocentrotus purpuratus*) and *Caenorhabditis elegans*) were extracted from GenBank. We confirmed these as orthologs based on database annotation of identity and/or predicted function, as well as the requirement that the sequence be the top hit in a BLAST of the human sequence against the genome database for each organism. Conservation at selected positions was further checked and confirmed by BLAST searches of the sequence against the genomic and EST database of that organism. We aligned the human NCCT (SLC12A3) sequence to the following vertebrate orthologs: mouse, rat, rabbit, dog, cow, chicken, zebrafish and winter flounder (*Pseudopleuronectes americanus*). We aligned the NKCC2 (SLC12A1) human sequence to mouse, rat, rabbit, dog, chicken, zebrafish and green spotted pufferfish (*Tetraodon nigroviridis*). These two evolutionarily and functionally related proteins share a single common ancestor in *S. purpuratus*, *C. elegans* and *D. melanogaster*. Sequences were also compared among paralogs of similar function: we aligned NCCT and NKCC2 sequences to each other and to the close paralog human NKCC1 (SLC12A2), as well as to the more distant functional paralogs of the SLC12A family KCC1–KCC4 (SLC12A4–SLC12A7). ROMK (KCNJ1) orthologs included were human, mouse, rat, dog, chicken, cow, frog (*Xenopus laevis*), zebrafish, fugu (*Takifugu rubripes*), *C. elegans*, *S. purpuratus* and *D. melanogaster*. Human paralogs analyzed were KCNJ2–KCNJ6 and KCNJ8–KCNJ16. We used CLUSTALW for multiple and pairwise alignments. We inferred that mutations were functional if they occurred at residues completely conserved in orthologs; the only exception to this was when the mutant residue was found at the homologous position in distant paralogs, as that would imply that the new residue is tolerated at that position and that it is unlikely to cause loss of function.

Evaluation of criteria for functional mutations. We compared the set of 246 Gitelman's syndrome mutations from independent European-ancestry kindreds to the set of variants found in FHS unrelateds with minor allele frequency <1%. Using a disease allele frequency of 1/200 (refs. 16,46), we estimated the population frequency of each Gitelman's mutation. We then estimated the sensitivity and specificity of various criteria of evolutionary conservation for identification of loss of function mutations. Sensitivity was estimated based on the percentage of Gitelman's alleles captured by each criterion. The corresponding specificity of criteria was estimated using Bayes' theorem: $P(A|B) = P(B|A)P(A)/P(B)$, where the probability that an allele found in a FHS subject is disease causing (A) given that it is of type n (B) equals the proportion of alleles of type n among Gitelman's mutations, divided by the proportion of alleles of type n in FHS multiplied by the frequency of Gitelman's syndrome alleles in the population.

Statistical analyses. In statistical comparisons of adjusted blood pressure levels between carriers and noncarriers, we took into account correlations between relatives (siblings and first cousins). We used a two-tailed t-test to test the null hypothesis that the difference from the mean blood pressure level is 0 for carriers, and adjusted the standard error in the t-test by an inflation factor as given by

$$\sqrt{\frac{n_c + 2n_{sp}c_{sibs} + 2n_{fcp}c_{fc}}{n_c}}$$

where n_c, n_{sp}, and n_{fcp} represent the number of carriers, the number of sibling pairs who are both carriers, and the number of first cousin pairs who are both carriers, respectively, and c_{sibs} and c_{fc} are the covariances between the siblings and first cousins in the Framingham population in general.

The inflation factor is calculated as follows. Let X_i denote the trait value of the ith individual in the sample, and $\bar{X}$ the mean of their trait values. Then in the presence of sibling and first-cousin pairs, under the null hypothesis of no association between the mutations and trait values,

$$\begin{aligned}\mathrm{var}(\bar{X}) &= [\sum_{i=1}^{n}\mathrm{var}(X_i) + 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\mathrm{cov}(X_i, X_j)]/n^2 \\ &= [\sum_{i=1}^{n}\mathrm{var}(X_i) + 2n_{sp}c_{sibs} + 2n_{fcp}c_{fc}]/n^2\end{aligned}$$

where var and cov are variance and covariance, respectively. In standard t-test where the relatedness among sampled individuals is not considered, the variance of $\bar{X}$ would be estimated by

$$[\sum_{i=1}^{n}\mathrm{var}(X_i)]/n^2$$

so the ratio between the correct variance and this last expression is

$$[\sum_{i=1}^{n}\mathrm{var}(X_i) + 2n_{sp}c_{sibs} + 2n_{fcp}c_{fc}]/[\sum_{i=1}^{n}\mathrm{var}(X_i)]$$

Because the traits are normalized, we have

$$(n + 2n_{sp}c_{sibs} + 2n_{fcp}c_{fc})/n$$

The use of a t-test was justified because the distribution does not violate the assumption of a normal distribution using the Kolmogorov-Smirnov test ($P = 0.725$ and 0.917 for long term average SBP and DBP, respectively). An inflation factor of similar form was used for the binomial test, where the null hypothesis is that the carrier status does not change the probability of a subject having an adjusted blood pressure reading below or above the population mean.

To further control for family effects, we tested the null hypothesis that the variation in adjusted SBP values among mutation carriers was no different from the variation among all members in the same family. A statistically significant result would indicate that carriers have more similar adjusted SBP values. We used a two-tailed, paired t-test to compare mean SBP between carrier and noncarrier siblings. We considered the ratio of the variance of the carrier sibs versus the variance of all siblings in each family. Six sibships each having at least one carrier-carrier pair and one carrier-noncarrier pair were considered. The variance ratio was then averaged across families as the test statistic. To derive the empirical distribution of the test statistic under the null hypothesis, 10,000 permutations were performed by randomizing carrier and noncarrier status of siblings within each family and recalculating the averaged ratio each time. This enabled estimation of the probability of a ratio less than or equal to the observed ratio occurring by chance.

We performed Kaplan-Meier analysis of the time to onset of hypertension on the first age at which hypertension was diagnosed. A log-rank test and Cox proportional hazard analysis was used to compare carriers and noncarriers.

In all analyses, we used $\alpha = 0.05$ as the threshold for significance in two-tailed tests. Statistical analyses were done using R 2.4.1, GraphPad Prism 4 and SPSS 14.0.

Accession codes. GenBank mRNA: *SLC12A3* (*NCCT*), NM_000339; *SLC12A1* (*NKCC2*), NM_000338; *KCNJ1* (*ROMK*), NM_000220 and NM_153764-7. GenBank protein: SLC12A3 orthologs: human, NP_000330; mouse, NP_062288; rat, NP_062218; rabbit, AAC33139; dog, XP_535292; cow, XP_871112; chicken, XP_414059; zebrafish, NP_001038545; winter flounder, AAL26926. SLC12A1 orthologs: human, NP_000329; mouse, NP_899197; rat, NP_062007; rabbit, AAB03494; dog, XP_850426; chicken, XP_413814; zebrafish, NP_001002080; green spotted pufferfish, CAF99849. SLC12A1–SLC12A3 invertebrate orthologs: *S. purpuratus*, XP_783014; *C. elegans*, NP_502704; *D. melanogaster*, NP_648572. KCNJ1 orthologs: human, NP_722450; mouse, NP_062633; rat, NP_058719; dog, XP_546403; chicken, XP_425795; cow, XP_585917; frog, AAH79788; zebrafish, XP_684541; fugu, ABB87033; *C. elegans*, NP_509138; *S. purpuratus*, XP_789112; *D. melanogaster*, NP_651131. Other human paralogs: SLC12A2, NP_001037; SLC12A4, NP_005063; SLC12A5, NP_065759; SLC12A6, NP_005126; SLC12A7, NP_006589; KCNJ2, NP_000882; KCNJ3, NP_002230; KCNJ4, NP_004972; KCNJ5, NP_000881; KCNJ6, NP_002231; KCNJ8, NP_004973; KCNJ9, NP_004974; KCNJ10, NP_002232; KCNJ11, NP_000516; KCNJ12, NP_066292; KCNJ13, NP_002233; KCNJ14, NP_733838; KCNJ15, NP_733933; KCNJ16, NP_733938.

Note: Supplementary information is available on the Nature Genetics website.

ACKNOWLEDGMENTS

We thank C. Nelson-Williams for management of the DNA database. J.N.F. is supported by the Agency for Science, Technology and Research, Singapore. This work was supported in part by a US National Institutes of Health Specialized Center of Research in Hypertension grant (to R.P.L.) and the US National Heart, Lung and Blood Institute's Framingham Heart Study contract NO1-HC-25195.

Published online at http://www.nature.com/naturegenetics
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions

1. Levy, D. *et al.* Evidence for a gene influencing blood pressure on chromosome 17. Genome scan linkage results for longitudinal blood pressure phenotypes in subjects from the Framingham heart study. *Hypertension* **36**, 477–483 (2000).
2. Saxena, R. *et al.* Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science* **316**, 1331–1336 (2007).
3. Scott, L.J. *et al.* A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. *Science* **316**, 1341–1345 (2007).
4. Zeggini, E. *et al.* Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. *Science* **316**, 1336–1341 (2007).
5. Helgadottir, A. *et al.* A common variant on chromosome 9p21 affects the risk of myocardial infarction. *Science* **316**, 1491–1493 (2007).
6. McPherson, R. *et al.* A common allele on chromosome 9 associated with coronary heart disease. *Science* **316**, 1488–1491 (2007).
7. Frayling, T.M. *et al.* A common variant in the *FTO* gene is associated with body mass index and predisposes to childhood and adult obesity. *Science* **316**, 889–894 (2007).
8. The Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* **447**, 661–678 (2007).
9. Pritchard, J.K. Are rare variants responsible for susceptibility to complex diseases? *Am. J. Hum. Genet.* **69**, 124–137 (2001).
10. Lifton, R.P., Gharavi, A.G. & Geller, D.S. Molecular mechanisms of human hypertension. *Cell* **104**, 545–556 (2001).

11. Lifton, R.P. Genetic dissection of human blood pressure variation: common pathways from rare phenotypes. *Harvey Lect.* **100**, 71–101 (2004–2005).
12. Birkenhäger, R. *et al.* Mutation of *BSND* causes Bartter syndrome with sensorineural deafness and kidney failure. *Nat. Genet.* **29**, 310–314 (2001).
13. Simon, D.B. *et al.* Mutations in the chloride channel gene, *CLCNKB*, cause Bartter's syndrome type III. *Nat. Genet.* **17**, 171–178 (1997).
14. Simon, D.B. *et al.* Bartter's syndrome, hypokalaemic alkalosis with hypercalciuria, is caused by mutations in the Na-K-2Cl cotransporter *NKCC2*. *Nat. Genet.* **13**, 183–188 (1996).
15. Simon, D.B. *et al.* Genetic heterogeneity of Bartter's syndrome revealed by mutations in the K^+ channel, ROMK. *Nat. Genet.* **14**, 152–156 (1996).
16. Simon, D.B. *et al.* Gitelman's variant of Bartter's syndrome, inherited hypokalaemic alkalosis, is caused by mutations in the thiazide-sensitive Na-Cl cotransporter. *Nat. Genet.* **12**, 24–30 (1996).
17. Rudin, A. Bartter's syndrome. A review of 28 patients followed for 10 years. *Acta Med. Scand.* **224**, 165–171 (1988).
18. Chanchevalap, S. *et al.* Involvement of histidine residues in proton sensing of ROMK1 channel. *J. Biol. Chem.* **275**, 7811–7817 (2000).
19. Riveira-Munoz, E. *et al.* Transcriptional and functional analyses of *SLC12A3* mutations: new clues for the pathogenesis of Gitelman syndrome. *J. Am. Soc. Nephrol.* **18**, 1271–1283 (2007).
20. De Jong, J.C. *et al.* Functional expression of mutations in the human NaCl cotransporter: evidence for impaired routing mechanisms in Gitelman's syndrome. *J. Am. Soc. Nephrol.* **13**, 1442–1448 (2002).
21. Schwalbe, R.A., Bianchi, L., Accili, E.A. & Brown, A.M. Functional consequences of ROMK mutants linked to antenatal Bartter's syndrome and implications for treatment. *Hum. Mol. Genet.* **7**, 975–980 (1998).
22. Sabath, E. *et al.* Pathophysiology of functional mutations of the thiazide-sensitive Na-Cl cotransporter in Gitelman disease. *Am. J. Physiol. Renal Physiol.* **287**, F195–F203 (2004).
23. Syrén, M.-L. *et al.* Identification of fifteen novel mutations in the SLC12A3 gene encoding the Na-Cl co-transporter in Italian patients with Gitelman syndrome. *Hum. Mutat.* **20**, 78 (2002).
24. Lemmink, H.H. *et al.* Novel mutations in the thiazide-sensitive NaCl cotransporter gene in patients with Gitelman syndrome with predominant localization to the C-terminal domain. *Kidney Int.* **54**, 720–730 (1998).
25. Ramensky, V., Bork, P. & Sunyaev, S. Human non-synonymous SNPs: server and survey. *Nucleic Acids Res.* **30**, 3894–3900 (2002).
26. Ng, P.C. & Henikoff, S. SIFT: Predicting amino acid changes that affect protein function. *Nucleic Acids Res.* **31**, 3812–3814 (2003).
27. Thomas, P.D. *et al.* Applications for protein sequence–function evolution data: mRNA/protein expression analysis and coding SNP scoring tools. *Nucleic Acids Res.* **34**, W645–W650 (2006).
28. Van Gestel, S., Houwing-Duistermaat, J.J., Adolfsson, R., van Duijn, C.M. & Van Broeckhoven, C. Power of selective genotyping in genetic association analyses of quantitative traits. *Behav. Genet.* **30**, 141–146 (2000).
29. Cruz, D.N. *et al.* Mutations in the Na-Cl cotransporter reduce blood pressure in humans. *Hypertension* **37**, 1458–1464 (2001).
30. Cooper, D.N., Krawczak, M. & Antonorakis, S.E. The nature and mechanisms of human gene mutation. In *The Metabolic and Molecular Bases of Inherited Disease* (eds. Scriver, C., Beaudet, A.L., Sly, W.S. & Valle, D.) 259–291 (McGraw-Hill, New York, 1995).
31. The ALLHAT Officers and Coordinators for the ALLHAT Collaborative Research Group. Major outcomes in high-risk hypertensive patients randomized to angiotensin-converting enzyme inhibitor or calcium channel blocker vs diuretic: the Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack Trial (ALLHAT). *J. Am. Med. Assoc.* **288**, 2981–2997 (2002).
32. Kannel, W.B., Wolf, P.A., Verter, J. & McNamara, P.M. Epidemiologic assessment of the role of blood pressure in stroke. The Framingham Study. *J. Am. Med. Assoc.* **214**, 301–310 (1970).
33. He, J. & Whelton, P.K. Elevated systolic blood pressure and risk of cardiovascular and renal disease: overview of evidence from observational epidemiologic studies and randomized controlled trials. *Am. Heart J.* **138**, 211–219 (1999).
34. Kannel, W.B., Schwartz, M.J. & McNamara, P.M. Blood pressure and risk of coronary heart disease: the Framingham study. *Dis. Chest* **56**, 43–52 (1969).
35. Franklin, S.S. *et al.* Does the relation of blood pressure to coronary heart disease risk change with aging? The Framingham Heart Study. *Circulation* **103**, 1245–1249 (2001).
36. Sieck, U.V. & Ohlsson, A. Fetal polyuria and hydramnios associated with Bartter's syndrome. *Obstet. Gynecol.* **63**, 22S–24S (1984).
37. Ohlsson, A., Sieck, U., Cumming, W., Akhtar, M. & Serenius, F. A variant of Bartter's syndrome. Bartter's syndrome associated with hydramnios, prematurity, hypercalciuria and nephrocalcinosis. *Acta Paediatr. Scand.* **73**, 868–874 (1984).
38. Shaer, A.J. Inherited primary renal tubular hypokalemic alkalosis: a review of Gitelman and Bartter syndromes. *Am. J. Med. Sci.* **322**, 316–332 (2001).
39. Cohen, J.C. *et al.* Multiple rare alleles contribute to low plasma levels of HDL cholesterol. *Science* **305**, 869–872 (2004).
40. Romeo, S. *et al.* Population-based resequencing of *ANGPTL4* uncovers variations that reduce triglycerides and increase HDL. *Nat. Genet.* **39**, 513–516 (2007).
41. Veterans Administration Cooperative Study Group on Antihypertensive Agents. Effects of treatment on morbidity in hypertension. Results in patients with diastolic blood pressures averaging 115 through 129 mm Hg. *J. Am. Med. Assoc.* **202**, 1028–1034 (1967).
42. Tchernitchko, D., Goossens, M. & Wajcman, H. In silico prediction of the deleterious effect of a mutation: proceed with caution in clinical genetics. *Clin. Chem.* **50**, 1974–1978 (2004).
43. Takahashi, N. *et al.* Posttranscriptional compensation for heterozygous disruption of the kidney-specific NaK2Cl cotransporter gene. *J. Am. Soc. Nephrol.* **13**, 604–610 (2002).
44. Murphy, K.M. & Berg, K.D. Mutation and single nucleotide polymorphism detection using temperature gradient capillary electrophoresis. *Expert Rev. Mol. Diagn.* **3**, 811–818 (2003).
45. Franklin, S.S. *et al.* Hemodynamic patterns of age-related changes in blood pressure. The Framingham Heart Study. *Circulation* **96**, 308–315 (1997).
46. Cruz, D.N. The renal tubular Na-Cl co-transporter (NCCT): a potential genetic link between blood pressure and bone density? *Nephrol. Dial. Transplant.* **16**, 691–694 (2001).