

ARTICLE

# Functional identification of the promoter of *SLC4A5*, a gene associated with cardiovascular and metabolic phenotypes in the HERITAGE Family Study

Adrian M Stütz<sup>1</sup>, Margarita Teran-Garcia<sup>1</sup>, DC Rao<sup>2,3</sup>, Treva Rice<sup>2</sup>, Claude Bouchard<sup>\*,1</sup> and Tuomo Rankinen<sup>1</sup>

<sup>1</sup>Human Genomics Laboratory, Pennington Biomedical Research Center, Louisiana State University System, Baton Rouge, LA, USA; <sup>2</sup>Division of Biostatistics, Washington University School of Medicine, St Louis, MO, USA; <sup>3</sup>Department of Genetics and Psychiatry, Washington University School of Medicine, St Louis, MO, USA

The sodium bicarbonate cotransporter gene *SLC4A5*, associated earlier with cardiovascular phenotypes, was tested for associations in the HERITAGE Family Study, and possible mechanisms were investigated. Twelve tag-single nucleotide polymorphisms (SNPs) covering the *SLC4A5* gene were analyzed in 276 Black and 503 White healthy, sedentary subjects. Associations were tested using a variance components-based (QTD) method with data adjusted for age, sex and body size. In Whites, rs6731545 and rs7571842 were significantly associated with resting and submaximal exercise pulse pressure (PP) ( $0.0004 < P < 0.0007$  and  $0.002 < P < 0.003$  respectively). Additionally, rs6731545 was associated with submaximal-exercise systolic blood pressure (SBP) and rate pressure product ( $P = 0.002$ , both). New associations between rs6731545 and submaximal-exercise  $VO_2$  ( $P = 0.003$ ), rs7587117 and rs7571842 and  $VCO_2$  ( $0.0005 < P < 0.0009$ ) and rs828853 and VE ( $P = 0.002$ ) were found. All these associations had a FDR  $< 0.05$ . Single-marker associations were confirmed in haplotype analyses. Using in silico analysis, evidence was found for a main and an alternative promoter for *SLC4A5*. Specific promoter activity was experimentally confirmed using reporter constructs targeting both promoters in three physiologically relevant cell lines. Re-sequencing of 32 individuals having opposite homozygotes for rs7571842 and rs6731545 and exhibiting significantly different phenotypes showed no SNPs in the alternative promoter and no differences between the groups with SNPs in the main promoter. Also, of all known *SLC4A5*-coding SNPs, only one synonymous SNP was detected. Summarizing, the observed associations with resting and submaximal-exercise cardiovascular and metabolic traits in the HERITAGE Family Study are likely due to neither variation in the promoter nor known coding SNPs of *SLC4A5*.

*European Journal of Human Genetics* (2009) 17, 1481–1489; doi:10.1038/ejhg.2009.64; published online 22 April 2009

**Keywords:** NBC4; hypertension; gene structure; SNP; exercise

\*Correspondence: Dr C Bouchard, Pennington Biomedical Research Center, 6400 Perkins Road, Baton Rouge, LA 70808-4124, USA.  
Tel: +1 225 763 2513; Fax: +1 225 763 0935;

E-mail: bouchac@pbrc.edu

Received 22 July 2008; revised 4 February 2009; accepted 17 March 2009; published online 22 April 2009

## Introduction

Hypertension and other cardiovascular diseases are believed to result from the influence of both a multitude of genes and one's lifestyle. Complex interactions may explain in part why the replication of candidate gene findings among studies investigating multi-factorial diseases has been difficult. However, when a candidate

gene is found to be associated with a given disease or phenotype in different populations, the hypothesis of a functional effect of sequence variation in this gene needs to be considered seriously.

In 2002, a study based on the Family Blood Pressure Program (FBPP)<sup>1</sup> associated the Solute Carrier Family 4 member 5 (*SLC4A5*) gene with blood pressure phenotypes. The authors tested associations of single nucleotide polymorphisms (SNPs) in eight candidate genes in the study of a quantitative trait locus (QTL) for hypertension in African Americans. Five SNPs in the *SLC4A5* gene showed significant associations with either pulse pressure (PP) in African Americans or pulse rate in Whites. In 2006, another study<sup>2</sup> replicated the association of a *SLC4A5* polymorphism with blood pressure phenotypes in Utah pedigrees at baseline and after a 10-year follow-up period.

These findings coincide with a meta-analysis<sup>3</sup> from the NHLBI GeneLink network, which combined observations from three independent projects (HyperGEN, HERITAGE Family Study and Genetics of Hypertension in Blacks)<sup>4–6</sup> to obtain LOD scores for SBP and DBP in African-American and Nigerian families. The *SLC4A5* gene is located within the chromosomal region 2p14–p13.1 where linkage evidence was identified. The region flanking the *SLC4A5* locus (2p12–q22) was also suggested to be important for hypertension and blood pressure variation in Whites in another meta-analysis using data from nine different studies, including the HERITAGE Family Study.<sup>7</sup> Finally, a recent study identified a significant linkage (LOD = 4.84) with antihypertensive response to renin-angiotensin system inhibiting drugs in the British Genetics of Hypertension Study and suggested *SLC4A5* as a promising candidate.<sup>8</sup>

Although the *SLC4A5* gene was already cloned<sup>9</sup> in 2000,<sup>9</sup> not much is known about its physiological role.<sup>10</sup> *SLC4A5* is a transmembrane protein which functions as an electrogenic cotransporter of bicarbonate ( $\text{HCO}_3^-$ ) and sodium ( $\text{Na}^+$ ) in the same direction.<sup>11,12</sup> So far, *SLC4A5* expression has been shown in liver, testis, spleen, heart, kidney, brain, lung, pancreas and skeletal muscle, a profile that suggests implications for common metabolic processes.<sup>9,12–14</sup>

The primary aim of the current study was to replicate in the HERITAGE Family Study the associations with resting blood pressure phenotypes and *SLC4A5* SNPs initially reported in FBPP.<sup>1</sup> Furthermore, we extended the analysis of a full panel of tag-SNPs to cardiovascular and metabolic traits measured during submaximal exercise before and after a standardized exercise program. As a second step, the promoter and all known coding SNPs were investigated as possible mechanisms for the associations. The promoter of *SLC4A5* was identified, functionally tested, and re-sequenced together with all exons containing known coding SNPs in an informative subset of the HERITAGE Family Study.

## Materials and methods

### Subjects

The design and methods of the HERITAGE Family Study have been described earlier.<sup>15</sup> Briefly, all subjects were required to be sedentary and healthy and to fulfill a series of inclusion criteria.<sup>15</sup> A total of 276 Blacks from 116 family units and 503 Whites from 99 families were available for the study. The institutional review board of each of the HERITAGE Family Study research centers approved the study protocol. Written informed consent was obtained from each participant.

### Phenotype measurements

Systolic blood pressure (SBP), diastolic blood pressure (DBP), and heart rate (HR) were measured twice on separate days under resting conditions and during submaximal exercise at a fixed workload of 50W, both at baseline and after a 20-week endurance training program, as described earlier.<sup>15,16</sup> Likewise, oxygen consumption ( $\text{VO}_{250}$ ), expired carbon dioxide ( $\text{VCO}_{250}$ ) and minute ventilation ( $\text{VE}_{50}$ ) during submaximal exercise at 50W were measured on 2 separate days, both pre- and post-training.<sup>15</sup> Pulse pressure (PP) was calculated as the difference between SBP and DBP and rate-pressure product (RPP) as SBP multiplied by HR.

### Genetic analyses

Genomic DNA was extracted from lymphoblastoid cell lines using standard procedures. SNP selection was based on CEU Hapmap data, [www.hapmap.org](http://www.hapmap.org) ( $r^2 = 0.8$ ,  $\text{MAF} > 0.1$ ) and included the significant SNPs from the FBPP study.<sup>1</sup> Genotyping of the *SLC4A5* polymorphisms was carried out by the primer extension method with fluorescence polarization detection. Details for PCR conditions and primer sequences are available upon request. All SNPs were in Hardy–Weinberg equilibrium in both Blacks and Whites.

### Statistical analyses

Haplotypes were generated with MERLIN. The associations between either single SNPs or *SLC4A5* haplotypes and hemodynamic and metabolic phenotypes were analyzed using a variance components and likelihood ratio test based procedure in the QTD software package <http://www.sph.umich.edu/csg/abecasis/><sup>17</sup> with details elsewhere.<sup>18</sup> Baseline hemodynamic phenotypes were adjusted for age, sex and BMI, whereas  $\text{VO}_{250}$ ,  $\text{VCO}_{250}$  and  $\text{VE}_{50}$  were adjusted for age, sex and body weight. All training responses (calculated as post- minus pre-measurements) were additionally adjusted for the respective baseline values. The Multtest procedure using FDR in SAS (Cary, NC, USA) was used to account for multiple testing. All analyses were carried out separately in Blacks and Whites. Renilla-normalized promoter activity data from the functional studies were analyzed using a general linear model. First, the constructs, the experiments, and the replicates

within each experiment were modeled as main effects. Second, when construct main effect was statistically significant, post-hoc analyses were carried out to test the pair-wise differences between the constructs.

### Cloning of human SLC4A5 promoter constructs

Promoter constructs (size and primers available upon request) were amplified from human genomic DNA with the Expand High Fidelity polymerase (Roche, Indianapolis, IN, USA). Furthermore, the remaining chromosomal DNA between the Long SLC4A5 promoter construct and the Dynactin gene (15.9 kb) was cloned as five overlapping constructs as follows: first (3131 bp), second (3185 bp), third (3205 bp), fourth (3543 bp), and fifth (3221 bp). In addition to the Long, the Alternative and the third construct, reverse complement constructs were created. All PCR fragments were cloned into the promoterless vector pGL3basic (Promega, Madison, WI, USA), and their correctness was verified by DNA sequencing. All DNAs for transfections were generated with the Endofree-Maxiprep kit (Qiagen, Valencia, CA, USA).

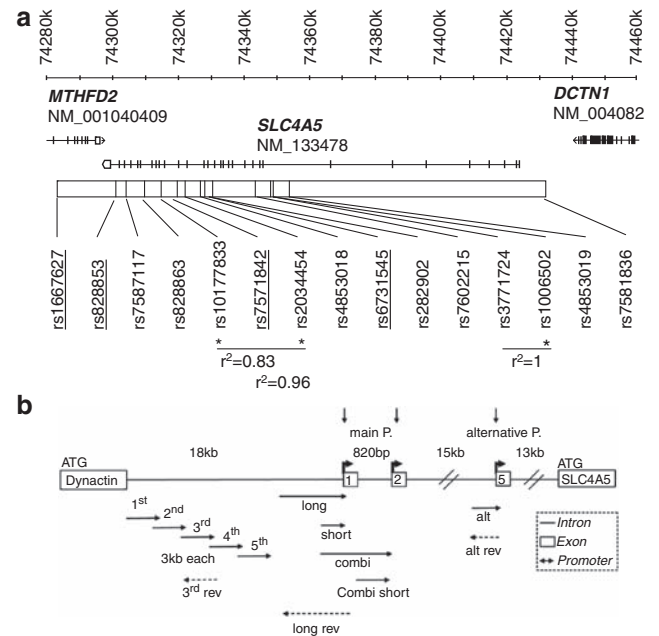
### Cell culture and reagents

The human epithelial lung cell line A549, the murine myoblastoma cell line C2C12, and the human embryonic kidney cell line HEK293 (ATCC, Manassas, VA, USA) were cultured in F12K, DMEM or MEM medium (all Invitrogen, Carlsbad, CA, USA), respectively, each supplemented with 10% FBS (Invitrogen) and 1% Penicillin/streptomycin. Cells were grown in a 5% CO<sub>2</sub> humidified incubator and seeded into 48-well plates the day before use and transfected at 80% confluence.

### Transient transfection

Cells were transfected with Lipofectamine 2000 (Invitrogen) according to the manufacturer's protocol using 300–400 ng reporter DNA and 30–80 ng Renilla vector pRL-TK (Promega). After incubation for 28 h, cells were

lysed with 1 × Passive Lysis buffer (Promega). Firefly and Renilla luciferase activity was measured using the Dual Luciferase Kit (Promega) and a FB12 luminometer (Bertold Detection systems, Oak Ridge, TN, USA). All firefly luciferase data were normalized to equal expression of Renilla luciferase of the same tube. Shown are the

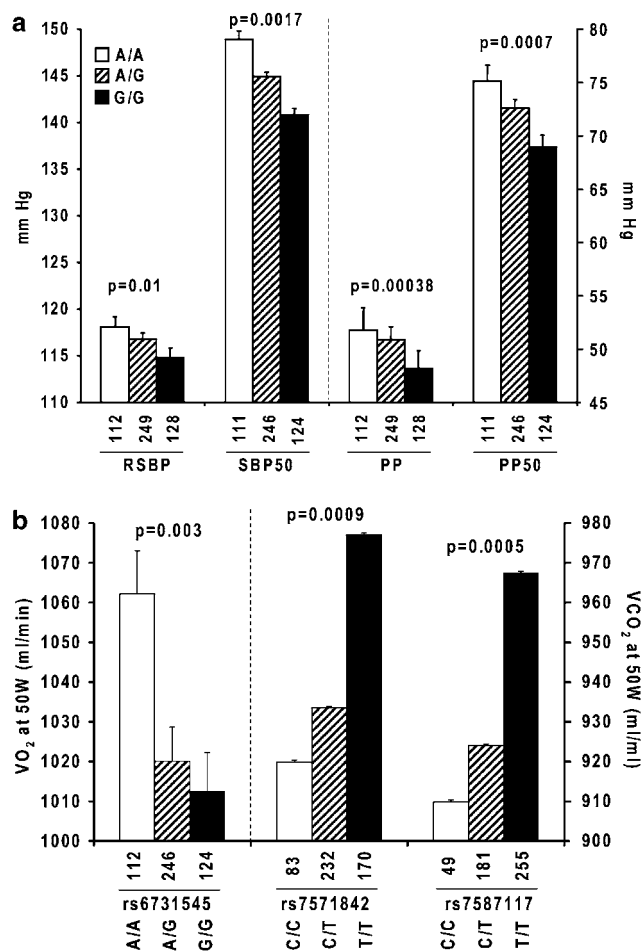


**Figure 1** SLC4A5 gene structure, location of promoter constructs and the position of SLC4A5 SNPs. (a) SNPs marked with an asterisk were significant in the earlier association studies and were not genotyped in this study. However, pair-wise linkage disequilibrium (LD) with tag-SNPs used in our study has been marked below the graph. SNPs that were significant in other studies and were also genotyped in this study have their names underlined. (b) SLC4A5 exons are boxed and numbered according to the SLC4A5 c isoform. Solid arrows mark the position and orientation of promoter reporter constructs, the names are given below. Dotted arrows depict identical constructs in reverse complement orientation.

**Table 1** Summary of associations between SLC4A5 SNPs and resting and submaximal phenotypes at baseline in Whites

Whites	rs1667627	rs828853	rs7587117	rs828863	rs7571842	rs4853018	rs6731545	rs828902	rs7602215	rs3771724	rs4853019	rs7581836
	1	2	3	4	5	6	7	8	9	10	11	12
RSBP	0.20	0.82	0.26	<b>0.02</b>	<b>0.05</b>	0.16	<b>0.01</b>	0.51	0.60	0.48	0.89	0.66
RDBP	0.73	0.14	0.35	0.58	0.63	0.58	0.71	0.20	0.65	0.45	0.33	0.24
RPP	0.75	0.68	0.58	0.23	0.53	0.73	0.58	0.44	0.09	0.37	0.89	0.92
RHR	0.60	0.54	0.89	0.92	0.63	0.26	0.36	0.13	0.15	0.14	0.92	0.74
PP	0.05	0.07	<b>0.03</b>	<b>0.01</b>	<i>0.003</i>	<b>0.03</b>	<b>0.0004</b>	0.06	0.74	0.12	0.38	0.06
SBP50	0.48	0.63	0.31	<b>0.01</b>	0.06	<b>0.05</b>	<b>0.002</b>	0.12	0.14	0.23	0.34	0.29
DBP50	0.15	<b>0.01</b>	0.13	0.16	0.27	0.79	0.66	0.31	0.28	0.82	0.63	0.39
RPP50	0.65	0.89	0.10	<b>0.04</b>	0.052	<b>0.05</b>	<b>0.002</b>	0.12	0.06	0.24	0.26	0.09
HRSO	0.74	0.69	<b>0.04</b>	0.27	0.12	0.17	0.07	0.41	0.08	0.41	0.35	0.07
PP50	0.09	0.33	<b>0.02</b>	<b>0.01</b>	<i>0.002</i>	<b>0.02</b>	<b>0.0007</b>	<b>0.02</b>	0.24	0.15	0.33	0.47
VO250	0.09	<b>0.02</b>	<b>0.02</b>	<b>0.03</b>	<b>0.006</b>	<b>0.03</b>	<b>0.003</b>	0.20	<b>0.04</b>	<b>0.05</b>	0.50	1
VCO250	0.07	<b>0.004</b>	<b>0.0005</b>	<b>0.02</b>	<b>0.0009</b>	<b>0.03</b>	<b>0.008</b>	0.22	<b>0.01</b>	<b>0.04</b>	0.35	1
VE50	0.09	<b>0.002</b>	<b>0.006</b>	0.26	<b>0.02</b>	0.31	0.15	0.11	0.07	0.07	0.82	0.92

RSBP, resting systolic blood pressure; RPP, rate pressure product; PP, pulse pressure; SBP50, systolic blood pressure at 50W; VE50, minute ventilation at 50W. P-values are from the QTDT Total association model analysis. Bold P-values are <0.05 and the italicised P-values remain significant after multiple testing adjustment (FDR < 0.05).



**Figure 2** Genotypes at *SLC4A5* SNPs exhibit significantly different cardiovascular and metabolic phenotypic values in White HERITAGE subjects. (a) Systolic blood pressure and pulse pressure measured at rest (SBP, PP) and during submaximal exercise at 50W (SBP50, PP50) according to the rs6731545 genotypes. Phenotype values are means and standard error and are depicted on the left y-axis for SBPs and the right y-axis for PPs, divided by a dotted line. The numbers on the x-axis represent the number of subjects for each genotype. *P*-values are from the total association model of the QTD software package. (b) Submaximal exercise VO<sub>2</sub> and VCO<sub>2</sub> phenotype data according to the alleles of three SNPs.

means + SEM of three independent experiments with each construct in triplicates ( $n=9$ ) for each cell line.

### Resequencing of the *SLC4A5* promoters

Sixteen individuals who were CC homozygotes for rs7571842 and GG homozygotes for rs6731545 and 16 individuals who were TT and AA homozygotes at the respective loci, (matched for sex, age and BMI), were resequenced in the Alternative and the Long *SLC4A5* promoters. Furthermore, all exons containing known *SLC4A5* coding SNPs (coding exons 6, 8, 18 and 25) were also sequenced in the same two groups.

## Results

Basic characteristics of the subjects are presented in Supplementary Table I (available online). Minor allele frequencies and pairwise linkage disequilibria (LD) of the 12 tag-SNPs, which include the significant SNPs of the FBPP study,<sup>1</sup> are given in Supplementary Tables II and III. Their physical location in regard to the *SLC4A5* gene and the way they overlap with significant SNPs used in the two earlier studies are given in Figure 1a.<sup>1,2</sup>

### Single marker associations

A summary of the associations between the *SLC4A5* SNPs and cardiovascular and metabolic phenotypes is presented in Table 1 and Supplementary Table IV. In Whites, rs6731545 was significantly associated with resting PP ( $P=0.0004$ ) and submaximal exercise (50W) PP (PP50), submaximal exercise SBP (SBP50), and rate pressure product (RPP50) ( $P=0.002$  for both), explaining 1.5 to 3.2% of the trait variance. These associations retained a FDR<0.05 after accounting for multiple testing for 12 genotypes and 13 phenotypes (Table 1). The rs6731545 G/G homozygotes showed 3.3 and 8.1 mmHg lower SBP and 3.6 and 6.8 mmHg lower PP values at rest and during exercise, respectively, than the A/A homozygotes (Figure 2a). Similar results for resting and submaximal exercise PPs were also observed with rs7571842 (Table 1).

Consistent associations were found in Whites with VO<sub>2</sub>50 and VCO<sub>2</sub>50 across SNPs between rs828853 and rs6731545 ( $0.0005 < P < 0.03$ ) (Table 1, Figure 2b). A similar trend was observed for VE50, with the strongest signal from rs828853 ( $P=0.002$ ). Associations between rs6731545 and VO<sub>2</sub>50 ( $P=0.003$ ), rs7587117 and rs7571842 with VCO<sub>2</sub>50 ( $0.0005 < P < 0.0009$ ) and rs828853 with VE50 ( $P=0.002$ ) had a FDR<0.05 (Table 1).

### Haplotype associations

The associations were confirmed in analyses with haplotypes constructed from the six SNPs between rs828853 and rs6731545 (Table 2). Associations were found with three main haplotypes, numbers 4, 6 and 11 (Supplementary Table V). The haplotypes 4 and 6, defined by a G at rs828853 and CG at rs7571842 and rs6731545, were consistently associated with lower trait values, whereas haplotype 11 (defined by a G at rs828853 and TA at rs7571842 and rs6731545) was associated with higher phenotype levels. Haplotypes 4 and 6 differ in SNPs rs7587117 and rs4853018. Haplotype 6 was strongly associated with metabolic but not cardiovascular phenotypes whereas haplotype 4 was associated with both classes of phenotypes. In summary, opposite alleles in rs7571842 and rs6731545 with the fixed G allele at rs828853 are required for the associations, confirming the importance of these SNPs found in the single marker analysis. Neither haplotype nor single SNP analyses showed any significant

**Table 2** Summary of associations between *SLC4A5* SNP2-7 haplotypes with phenotypes in Whites using the QTDT model

Haplotype	Freq (%)	Resting					Submaximal exercise at 50 W							
		SBP	DBP	RPP	RHR	PP	SBP	DBP	RPP	HR	PP	VO2	VCO2	VE
Global	NA	<b>0.008</b>	0.61	0.21	0.88	<b>0.001</b>	<b>0.002</b>	0.08	<b>0.003</b>	0.08	<b>0.002</b>	<b>0.001</b>	<b>0.001</b>	<b>0.014</b>
1	2.40	0.58	0.13	0.89	0.89	0.61	0.16	0.01	0.1	0.54	0.89	0.66	1	0.48
2	0.10	0.42	0.26	0.44	0.72	1	0.57	0.65	0.76	0.28	0.59	1	0.29	0.11
3	1.20	0.07	0.49	0.11	0.55	0.08	0.42	1	1	0.56	0.22	0.35	1	0.18
4	23.90	<b>0.01</b>	0.86	0.11	0.55	<b>0.003</b>	0.08	0.44	<b>0.02</b>	0.03	<b>0.002</b>	<b>0.05</b>	<b>0.007</b>	<b>0.007</b>
5	0.10	0.76	0.76	0.82	1	0.58	0.86	0.58	0.6	0.45	0.81	0.22	0.11	0.13
6	4.30	0.12	0.84	0.35	1	<b>0.05</b>	0.39	1	0.45	0.73	0.23	<b>0.0004</b>	<b>0.007</b>	<b>0.03</b>
7	0.30	0.6	0.19	0.19	0.23	0.51	0.29	0.02	0.01	0.004	0.92	0.56	0.53	0.13
8	2.10	1	0.86	0.65	0.62	0.86	0.73	0.82	0.49	0.59	0.38	0.17	0.66	0.48
11	8.70	<b>0.001</b>	0.32	0.03	0.45	<b>0.001</b>	<b>0.001</b>	0.03	<b>0.0009</b>	0.02	<b>0.005</b>	<b>0.01</b>	<b>0.009</b>	0.19
12	0.20	1	0.92	0.51	0.49	1	0.82	0.81	0.39	0.28	0.92	0.25	0.24	0.24
13	0.20	0.11	0.5	0.59	0.68	0.17	0.04	0.34	0.37	0.73	0.02	0.26	0.45	0.43
14	7.40	0.58	0.46	1	0.61	1	0.03	0.11	0.16	0.92	0.13	0.53	0.35	0.07
15	0.40	0.24	0.73	0.76	0.18	0.19	0.84	0.84	0.06	0.003	0.72	0.89	0.24	0.12
16	32.90	0.51	0.26	0.92	0.51	0.07	0.21	0.43	0.15	0.24	0.08	0.03	0.13	0.74
17	2.10	0.58	0.92	0.57	0.79	0.52	0.67	0.89	0.45	0.6	0.89	0.53	0.75	0.82
18	12.00	0.55	0.71	0.38	0.48	0.67	0.58	0.17	0.53	0.69	0.75	1	0.25	0.06
19	1.00	0.84	0.86	0.11	0.03	0.64	0.89	0.39	0.15	0.01	0.31	0.4	0.66	0.67
20	0.30	0.13	0.45	0.09	0.31	0.23	0.82	0.78	0.82	0.82	0.89	1	0.72	0.75

Bold *P*-values are <0.05.

associations with baseline phenotypes in Blacks (Supplementary Tables IV and VI) or with training response phenotypes in both Blacks and Whites (data not shown).

### Promoter identification

To assess whether any of the significant SNPs identified in this and other studies are markers for causal variation located in the nearby putative *SLC4A5* promoter sequence, the promoter of *SLC4A5* was further defined and functionally tested. The NCBI reference gene model for *SLC4A5* suggested a complex structure with four non-coding exons spanning 27.5 kb of genomic sequence (Figure 1). However, manual EST assembly suggested the presence of another 5' non-coding exon located 820 bp further upstream, based on three EST sequences (DB455137, DB062381, BX492763). In addition, existence of an alternative promoter directly upstream of the non-coding exon 5 was suggested by the EST DA112469 (Figure 1b). To clarify which of these regions contained the promoter, several promoter reporter constructs were generated (Figure 1b). Additionally, the 15.5 kb region between the predicted main *SLC4A5* promoter and the 5' neighboring gene, Dynactin (*DCTN1*) was cloned as five overlapping constructs.

### Functional testing of the promoter

Promoter activity of all 13 constructs was tested in three different cell lines: A549 (human lung epithelia), C2C12 (mouse myoblastoma) and HEK293 (human embryonic kidney) (Figure 3). In all three cell lines, constructs targeting the NCBI promoter in front of exon 2 (combi and combi short) did not show any promoter activity compared with the empty vector pGL3basic. In contrast,

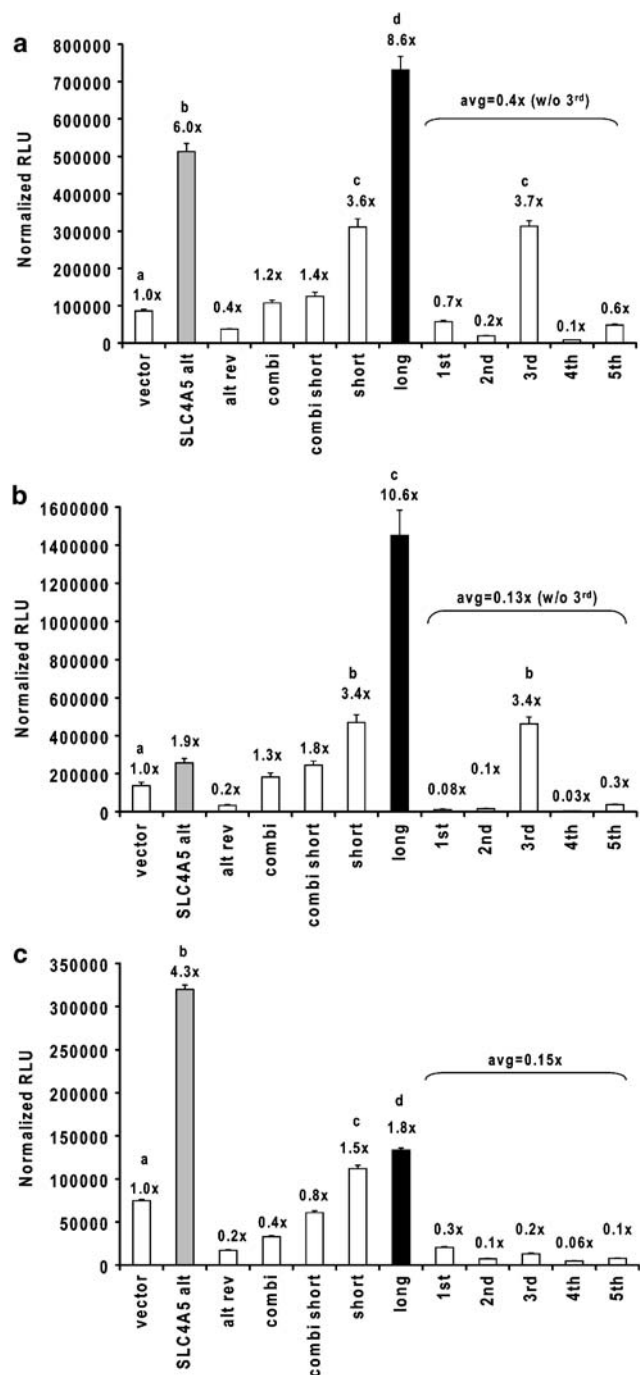
the Long construct showed significantly higher promoter activity compared with the empty vector in all cell lines and was also the strongest active construct in A549 (8.6 fold) and C2C12 (10.6 fold). In the HEK293 cell line (Figure 3c), the 1.8 × activity over pGL3basic should be considered meaningful, as compared with the average of five constructs without activity (mean = 16% of vector); this translated into an 11.7-fold higher activity level. Constructs with DNA preceding new exon 1 identified by EST assembly showed the most convincing promoter activity, suggesting that the main promoter resides within this region.

The alternative promoter and the third construct also showed significant promoter activity in two of the three cell lines. To distinguish between promoter and enhancer function, reverse forms of those constructs were generated and tested. All three reverse constructs completely lost activity in all cell lines (Figure 3 and data not shown), showing the presence of a unidirectional promoter in those regions.

### Re-sequencing of promoter and exons containing known coding SNPs

The active promoters and all exons containing known coding SNPs were re-sequenced in 32 White HERITAGE subjects. Half of the subjects (*n* = 16) were C/C homozygotes at rs7571842 and G/G homozygotes at rs6731545, whereas the other half of the subjects were T/T and A/A homozygotes in the respective loci (Supplementary Table VII).

No SNPs were found in the alternative promoter, whereas three SNPs, rs3771742, rs6714954 and rs6732913 were identified in the main promoter in 5 individuals (Table 3).



**Figure 3** Transient transfection results of promoter reporter constructs in three cell lines. Promoter constructs were transfected into the (a) human epithelial lung cell line A549, (b) murine myoblastoma cell line C2C12 and (c) human embryonic kidney cell line HEK293. Data is depicted as normalized relative light units (RLU) after normalization to equal activity of cotransfected Renilla. Shown are the means  $\pm$  SEM of three independent experiments with each construct in triplicates ( $n = 9$ ). Fold promoter activities over the empty vector are written over each bar.

Their distribution of 2:3 between the groups was not statistically different, and therefore those SNPs were not further functionally evaluated. The NCBI dbSNP database (<http://www.ncbi.nlm.nih.gov/projects/SNP/>) contained putative coding SNPs in exons 6, 8, 18 and 25, which were sequenced (Table 3). Neither the two non-synonymous SNPs (rs17009792 and rs36081793) nor the frame-shift SNP (rs35866044) were found in any of the 32 individuals. Only one of the two synonymous SNPs, rs3796109, was found in three individuals in group 1. One intronic SNP near coding exon 25 was found in five individuals of group 1 but not in group 2. The synonymous SNP rs4853018 in exon 16 was already genotyped in the full cohort and showed only marginal associations (Table 1).

## Discussion

The first main goal of the present study was to test whether the earlier reported associations between *SLC4A5* polymorphisms and blood pressure phenotypes could be replicated in the HERITAGE Family Study cohort. Our results using comprehensive tag-SNP approach support the earlier reports from the FBPP study<sup>1</sup> and the Utah Pedigree study.<sup>2</sup> However, our data also extends the earlier findings to hemodynamic phenotypes measured during submaximal exercise and thereby strengthens the observations. Furthermore, our data showed new associations between the *SLC4A5* gene locus and submaximal exercise metabolic phenotypes. The second main goal was to investigate potential mechanisms behind the associations. The promoter(s) of the *SLC4A5* gene was successfully identified, functionally tested and re-sequenced. The results indicate that promoter sequence variation does not seem to be responsible for the associations. Finally, most known *SLC4A5* coding SNPs were not found in an informative subset of the HERITAGE Family Study thus excluding them as a likely mechanism.

The original FBPP study focused on a hypertensive cohort and used PP and HR as primary continuous hemodynamic phenotypes.<sup>1</sup> In comparison, our study used healthy, sedentary individuals, allowing analysis of SBP, DBP, RPP and PP without the confounding of medication. Our primary findings in Whites were mainly related to SBP, as both resting and submaximal SBP, resting and submaximal PP (SBP-DBP), and submaximal rate-pressure product ( $HR \times SBP$ ) showed significant associations in single SNP and/or haplotype analyses. This is in agreement with the significant findings with baseline SBP and 10-year follow-up SBP in the Utah pedigree study<sup>2</sup> observed with a SNP in strong LD ( $r^2 = 0.83$ ) with one of the main SNPs (rs7571842) of our study. In addition, the phenotypic effect goes in the same direction, as in both cases the minor allele carriers have lower SBP. In summary, of the four SNPs within the *SLC4A5* gene that were found

**Table 3** Summary of re-sequencing 32 individuals in key areas of the *SLC4A5* gene

Position	Coding exon 6	Coding exon 8	Coding exon 18	Coding exon 25	Alternative promoter	Main promoter
Class	Non-synonymous	Synonymous + frameshift	Synonymous	Non-synonymous	Regulatory	Regulatory
SNP(s)	rs17009792	rs12621832/ rs35866044	rs3796109	rs36081793	None	rs3771742 /rs6714954/ rs6732913
MAF CEU	0.042	0/unknown	0.034	unknown	—	0.093
Group 1	None	None/none	3 AB	None	None	3AB/3AB/3AB
Group 2	None	None/none	0	None	None	2AB/2AB/2AB
Intronic	None	rs7602215 / rs26121729	rs7571842	rs6745054	—	—
MAF CEU	—	0.108	0.367	0.083	—	—
Group 1	—	5 AB/3AB	16 BB	4 AB+1 BB	—	—
Group 2	—	3 AB/0	0	0	—	—

A, common allele; B, rare allele; AB, heterozygote. Only the rare alleles per group are counted.

to be significantly associated with hemodynamic and metabolic traits in this study, three were either identical to or in strong LD with earlier reported associated SNPs.

Our study is the first to investigate putative mechanisms behind these associations, such as the promoter or coding SNPs as either quantitative or qualitative changes affecting gene function. Furthermore, characterization of the *SLC4A5* promoter was warranted because some of the associated SNPs in this and other studies are located not too far from the putative promoter sequence. First, it is important to functionally show the location of the promoter of a gene. Many studies focus on a region either directly upstream of the start codon or of exon 1 of the reference sequence, often missing the true promoter by tens of kilobases. In the present study, we functionally identified three regions with promoter activity. The main *SLC4A5* promoter lies 29 kb upstream of the start codon-containing exon. It is located in front of the newly-identified first 5' non-coding exon and would initiate most *SLC4A5* related EST transcripts in the NCBI database. An alternative promoter is located in front of the fifth 5' non-coding exon (13 kb upstream of the ATG exon) resulting in a shorter 5'-UTR-containing transcript, but the same protein coding information as the main promoter. However, the alternative promoter could potentially be differentially regulated. Both existing ESTs and 5'-RACE products initiating from the main promoter miss this very last 5'-non-coding exon, supporting the idea that this area contains important regulatory elements affecting both transcription and splicing. Another transcriptionally active region was identified between the main *SLC4A5* promoter and the neighboring gene *DCTN1* (third construct out of 5). It is unclear whether this promoter initiates *SLC4A5* transcripts with even longer 5'-UTR or if it belongs to another gene or a non-coding RNA. All three identified areas act as unidirectional promoters and not merely as enhancers, as the reverse construct did not have any activity.

No DNA sequence variation was identified in the alternative promoter, and SNPs identified in the main promoter were found equally in both re-sequencing groups. Also, SNP rs7581836 tags a LD cluster that covers the whole promoter area (including all non-coding exons and the complete neighboring *Dynactin* gene) but did not show any significant associations with the phenotypes tested. This study therefore concludes that variation in the *SLC4A5* promoter is not responsible for the observed associations.

It is also unlikely that any of the known *SLC4A5* coding SNPs are responsible for the observed associations. The genotyped synonymous SNP rs4853018 (Gly729Gly) was not associated with any of the tested hemodynamic phenotypes. Furthermore, only two non-synonymous *SLC4A5* SNPs exist in the current NCBI SNP database. One of them has not been verified and would change a valine at an unconserved position into an isoleucine-amino acids with almost identical properties. The other one is located in coding exon 6. It is important that, re-sequencing of 32 individuals did not detect either of the non-synonymous SNPs. Although it is possible that they exist, they would be rare and could not explain the associations identified in the present study. Therefore, our results strongly suggest that DNA sequence variant(s) responsible for our association signals are located at the 3' end rather than 5' end of the *SLC4A5* gene, which now needs to be fully re-sequenced to identify the causal variation.

The consistent associations of *SLC4A5* with cardiovascular phenotypes found herein and in two other studies<sup>1,2</sup> may be explained by the sodium component of the sodium/bicarbonate co-transporter. It is known that increased Na<sup>+</sup> excretion by the kidneys will reduce both blood and extra-cellular fluid volumes and therefore blood pressure.<sup>19</sup> Furthermore, the reduction of salt intake has proven to be beneficial for elevated blood pressure in several clinical trials.<sup>20</sup> Also, it was recently speculated that

*SLC4A5* might be a promising candidate for the strong linkage signal found in response to renin-angiotensin system inhibiting drugs highlighting its potential importance for a salt-sensitive form of hypertension.<sup>8</sup>

On the other hand, bicarbonate transport functions of *SLC4A5* may be related to our observations on metabolic phenotypes involving respiratory gas exchanges. The majority of CO<sub>2</sub> is carried in blood as bicarbonate, which can reversibly dissociate back to CO<sub>2</sub> at the lungs, where it is exhaled. A key enzyme class for this process is carbonic anhydrases,<sup>21</sup> which are known to bind to at least three of the SLC4 family members including *SLC4A4*, the sodium-bicarbonate cotransporter most similar to *SLC4A5*.<sup>10</sup> Interestingly, *SLC4A5* expression was recently detected in sarcolemmal vesicles of human skeletal muscle.<sup>22</sup> It is therefore possible that one of the steps in the balance of bicarbonate generation, transport, and dissociation is affected by alterations in *SLC4A5* function as a sodium/bicarbonate cotransporter. Results in this study therefore warrant a more extensive analysis in regard to *SLC4A5* expression and functional characterization in the lung and in blood as thus far attention has focused on kidney, liver, and muscle.

Although no significant associations were found between the SNPs and hemodynamic phenotypes in Blacks, we cannot completely exclude them. The tag-SNPs were selected from HapMap Caucasian database and most SNPs were more polymorphic in Whites than in Blacks. The combination of lower minor allele frequencies and smaller sample size in Blacks resulted in lower statistical power. Furthermore, it is possible that the contribution of the *SLC4A5* gene to hemodynamic traits in Blacks is expressed in hypertensives rather than in healthy normotensives. These possibilities need to be addressed in future studies.

Our study confirmed and extended the associations observed between DNA sequence variation in the *SLC4A5* gene locus and blood pressure-related phenotypes in two earlier studies. It is important that, the strong and consistent associations with metabolic phenotypes related to gas exchange during submaximal exercise suggest additional physiological functions of this gene. We started to address possible mechanisms by identifying and studying the promoter but it does not seem to be responsible for these associations. Also, none of the known coding SNPs for *SLC4A5* seems to be involved. Studies are needed to clarify the importance of both sodium and bicarbonate transport functions of this cotransporter. Finally, identifying the target tissues responsible for the associations with cardiovascular and metabolic phenotypes represents an important next step. Evidence for an involvement of *SLC4A5* in the linkage signal for hypertension is growing with our results. Nevertheless, further studies will be needed to unambiguously clarify its contribution and continue looking for other responsible genes. They will allow a more definite assessment of the importance of

*SLC4A5* as a candidate gene for cardiovascular and respiratory diseases.

#### Acknowledgements

We thank Jessica Watkins and Christina Riley for their expert technical assistance in genotyping and sequencing. The HERITAGE Family Study is currently supported by NIH/NHLBI grant R01 HL-45670. Part of the study was supported by an R&D grant from the Health and Performance Enhancement Division of the Pennington Biomedical Research Center. C Bouchard is partially supported by the George A Bray Chair in Nutrition. AMS, TRa, TRi and MT conceived and designed the experiments. AMS and TRa performed the experiments. AMS and TRa analyzed the data. All authors contributed to writing the paper. CB was the principal investigator and oversaw experimental design and analysis.

#### Conflict of interest

The authors declare no conflict of interest.

#### References

- 1 Barkley RA, Chakravarti A, Cooper RS *et al*: Positional identification of hypertension susceptibility genes on chromosome 2. *Hypertension* 2004; **43**: 477–482.
- 2 Hunt SC, Xin Y, Wu LL *et al*: Sodium bicarbonate cotransporter polymorphisms are associated with baseline and 10-year follow-up blood pressures. *Hypertension* 2006; **47**: 532–536.
- 3 Rice T, Cooper RS, Wu X *et al*: Meta-analysis of genome-wide scans for blood pressure in African American and Nigerian samples. The National Heart, Lung, and Blood Institute GeneLink Project. *Am J Hypertens* 2006; **19**: 270–274.
- 4 Rao DC, Province MA, Leppert MF *et al*: A genome-wide affected sibpair linkage analysis of hypertension: the HyperGEN network. *Am J Hypertens* 2003; **16**: 148–150.
- 5 Cooper RS, Luke A, Zhu X *et al*: Genome scan among Nigerians linking blood pressure to chromosomes 2, 3, and 19. *Hypertension* 2002; **40**: 629–633.
- 6 Rice T, Rankinen T, Chagnon YC *et al*: Genomewide linkage scan of resting blood pressure: HERITAGE Family Study. Health, Risk Factors, Exercise Training, and Genetics. *Hypertension* 2002; **39**: 1037–1043.
- 7 Koivukoski L, Fisher SA, Kanninen T *et al*: Meta-analysis of genome-wide scans for hypertension and blood pressure in Caucasians shows evidence of susceptibility regions on chromosomes 2 and 3. *Hum Mol Genet* 2004; **13**: 2325–2332.
- 8 Padmanabhan S, Wallace C, Munroe PB *et al*: Chromosome 2p shows significant linkage to antihypertensive response in the British Genetics of Hypertension Study. *Hypertension* 2006; **47**: 603–608.
- 9 Pushkin A, Abuladze N, Newman D, Lee I, Xu G, Kurtz I: Cloning, characterization and chromosomal assignment of NBC4, a new member of the sodium bicarbonate cotransporter family. *Biochim Biophys Acta* 2000; **1493**: 215–218.
- 10 Pushkin A, Kurtz I: SLC4 base (HCO<sub>3</sub><sup>-</sup>, CO<sub>3</sub><sup>2-</sup>) transporters: classification, function, structure, genetic diseases, and knockout models. *Am J Physiol Renal Physiol* 2006; **290**: F580–F599.
- 11 Virkki LV, Wilson DA, Vaughan-Jones RD, Boron WF: Functional characterization of human NBC4 as an electrogenic Na<sup>+</sup>-HCO<sub>3</sub><sup>-</sup> cotransporter (NBCe2). *Am J Physiol Cell Physiol* 2002; **282**: C1278–C1289.
- 12 Sassani P, Pushkin A, Gross E *et al*: Functional characterization of NBC4: a new electrogenic sodium-bicarbonate cotransporter. *Am J Physiol Cell Physiol* 2002; **282**: C408–C416.
- 13 Abuladze N, Pushkin A, Tatishchev S, Newman D, Sassani P, Kurtz I: Expression and localization of rat NBC4c in liver



- and renal uroepithelium. *Am J Physiol Cell Physiol* 2004; **287**: C781–C789.
- 14 Damkier HH, Nielsen S, Praetorius J: Molecular expression of SLC4-derived Na<sup>+</sup>-dependent anion transporters in selected human tissues. *Am J Physiol Regul Integr Comp Physiol* 2007; **293**: R2136–R2146.
  - 15 Bouchard C, Leon AS, Rao DC, Skinner JS, Wilmore JH, Gagnon J: The HERITAGE family study. Aims, design, and measurement protocol. *Med Sci Sports Exerc* 1995; **27**: 721–729.
  - 16 Wilmore JH, Stanforth PR, Gagnon J *et al*: Endurance exercise training has a minimal effect on resting heart rate: the HERITAGE Study. *Med Sci Sports Exerc* 1996; **28**: 829–835.
  - 17 Abecasis GR, Cardon LR, Cookson WO: A general test of association for quantitative traits in nuclear families. *Am J Hum Genet* 2000; **66**: 279–292.
  - 18 Spielmann N, Leon AS, Rao DC *et al*: CETP genotypes and HDL-cholesterol phenotypes in the HERITAGE Family Study. *Physiol Genomics* 2007; **31**: 25–31.
  - 19 Meneton P, Jeunemaitre X, de Wardener HE, MacGregor GA: Links between dietary salt intake, renal salt handling, blood pressure, and cardiovascular diseases. *Physiol Rev* 2005; **85**: 679–715.
  - 20 He FJ, MacGregor GA: Effect of longer-term modest salt reduction on blood pressure. *Cochrane Database Syst Rev* 2004; (1): CD004937.
  - 21 Geers C, Gros G: Carbon dioxide transport and carbonic anhydrase in blood and muscle. *Physiol Rev* 2000; **80**: 681–715.
  - 22 Kristensen JM, Kristensen M, Juel C: Expression of Na<sup>+</sup>/HCO<sub>3</sub><sup>-</sup>-co-transporter proteins (NBCs) in rat and human skeletal muscle. *Acta Physiol Scand* 2004; **182**: 69–76.

Supplementary Information accompanies the paper on European Journal of Human Genetics website (<http://www.nature.com/ejhg>)