A Polymorphic 3’UTR Element in ATP1B1 Regulates Alternative Polyadenylation and Is Associated with Blood Pressure

Megana K. Prasad1*, Kavita Bhalla2*, Zhen Hua Pan3, Jeffrey R. O’Connell1, Alan B. Weder4, Aravinda Chakravarti5, Bin Tian3, Yen-Pei C. Chang1

1 Division of Endocrinology, Diabetes and Nutrition, University of Maryland School of Medicine, Baltimore, Maryland, United States of America, 2 Marlene and Stewart Greenebaum Cancer Center, University of Maryland School of Medicine, Baltimore, Maryland, United States of America, 3 Department of Biochemistry & Molecular Biology, University of Medicine and Dentistry of New Jersey, New Jersey Medical School, Newark, New Jersey, United States of America, 4 Department of Internal Medicine, University of Michigan School of Medicine, Ann Arbor, Michigan, United States of America, 5 McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland, United States of America

Abstract

Although variants in many genes have previously been shown to be associated with blood pressure (BP) levels, the molecular mechanism underlying these associations are mostly unknown. We identified a multi-allelic T-rich sequence (TRS) in the 3’UTR of ATP1B1 that varies in length and sequence composition (T22,27 and T23 GT GTG). The 3’UTR of ATP1B1 contains 2 functional polyadenylation signals and the TRS is downstream of the proximal polyadenylation site (A2). Therefore, we hypothesized that alleles of this TRS might influence ATP1B1 expression by regulating alternative polyadenylation. In vitro, the T23 GT GTG allele increases polyadenylation at the A2 polyadenylation site as compared to the T23 allele. Consistent with our hypothesis, the relative abundance of the A2-polyadenylated ATP1B1 mRNA was higher in human kidneys with at least one copy of the T23 GT GTG allele than in those lacking this allele. The T23 GT GTG allele is also associated with higher systolic BP (beta = 3.3 mmHg, p = 0.014) and diastolic BP (beta = 2.4 mmHg, p = 0.003) in a European-American population. Therefore, we have identified a novel multi-allelic TRS in the 3’UTR of ATP1B1 that is associated with higher BP and may mediate its effect by regulating the polyadenylation of the ATP1B1 mRNA.

Introduction

Hypertension (HTN) is a major public health concern affecting over 1 billion people worldwide, with hypertensive patients having an elevated risk of developing kidney disease, coronary disease, and stroke [1]. Genetic factors are involved in the development of HTN and heritability estimates for systolic and diastolic blood pressure (BP) measures range from 32%–68% [2,3,4,5]. Genome-wide linkage and association studies have uncovered novel loci involved in BP regulation and HTN pathophysiology [6,7,8]. However, the mechanisms by which genetic variants at these loci contribute to HTN are often poorly understood.

Previously, we identified a BP-related linkage region in 1q23-24 and several single nucleotide polymorphisms (SNPs) associated with BP [9]. In humans, markers in 1q23–1q32 have been reproducibly linked with BP-related traits [10,11,12]. Homologous regions in mice and rats have also been shown to harbour BP-quantitative trait loci (QTLs) [13,14,15]. Of the biological candidate genes within this linkage region, we showed a 3’UTR SNP in ATP1B1 (rs12079745) to be strongly associated with BP in a Caucasian population [9]. ATP1B1 encodes the beta 1-subunit of Na-K ATPase, a ubiquitously expressed cotransporter essential for regulating cell volume, electrolyte and nutrient transport, and signal transduction [15]. In the kidney, Na-K ATPase provides the primary drive for Na and water reabsorption, which is necessary for the maintenance of fluid and electrolyte homeostasis [16]. Na-K ATPase consists of the catalytic alpha subunit, and the beta subunit, which is thought to be the "regulatory subunit" as it is
the rate-limiting factor in the dimerization of Na-K ATPase, and is necessary for enzyme activity [17]. Indeed, SNPs in ATP1B1 have also been found to be associated with BP in an African-American and in a Chinese population [18,19].

The human ATP1B1 gene has 6 exons and encodes multiple mRNAs that differ in their 5’ and 3’ UTR lengths due to the use of multiple transcription initiation and polyadenylation sites [20,21]. The structure of the 3’UTR of Alp1b1 has been best studied in rats, in which the 3’UTR has 5 canonical polyadenylation sites, A1-A5, of which A1, A2, and A5 are predominantly used in vivo [21]. The differentially polyadenylated mRNAs are functionally distinct. The shorter A2-polyadenylated mRNA is translationally more efficient than the longer A5-polyadenylated mRNA due to a translational repressor sequence in the region unique to the A5-polyadenylated mRNA [22,23]. By contrast, the human ATP1B1 gene has 4 canonical polyadenylation sites (rat A1 is absent) and only the A2 and A5 sites are used for the polyadenylation of the ATP1B1 transcript (Figure 1A) [24]. Because the BP-associated SNP, rs12079745, is located in the 3’UTR near the A2 polyadenylation site, we hypothesized that sequence variation in the 3’UTR may regulate alternative polyadenylation, with alleles that promote the use of the A2 polyadenylation site leading to a translationally more efficient mRNA and an overall increase in ATP1B1 expression.

Alternative polyadenylation plays an important role in regulating gene expression and thus contributes to normal development and disease [25]. Polyadenylation begins with the recognition of the canonical polyadenylation signal (either AAUAAA or AUUAAA in humans) and the less well-defined U/GU-rich core downstream element (DSE) by the cleavage and polyadenylation specificity factor (CPSF) and the cleavage stimulating factor (CstF), respectively [26]. It is believed that, upon forming a complex, CPSF and CstF recruit additional factors, including cleavage factors I and II and polyadenylate [poly(A)] polymerase, that together cleave the nascent premRNA at the cleavage/polyadenylation site, between the polyadenylation signal and the DSE and add a poly(A) tail. Over 50% of human genes have multiple polyadenylation sites. In these genes, which site is used in a given mRNA is influenced by the surrounding cis elements, including the upstream sequence element, and the auxiliary DSE [26]. In addition, the abundance of cleavage and polyadenylation factors can also play a role in polyadenylation site choice [26]. While some non-canonical polyadenylation signals maintain normal function, other variants in these sequences alter their function and are the molecular bases of diseases. In contrast, because the DSE is less well characterized, very few disease-causing variants have thus been identified in the DSE [27].

Here, we report the identification of a novel polymorphic T-rich sequence (TRS) in the 3’UTR of ATP1B1, downstream of the A2 polyadenylation signal. Based on sequence composition and distance to the A2 polyadenylation signal, this TRS (U-rich sequence in the corresponding mRNA) is the putative DSE of the A2 site. Although alleles of this site (T<sub>22</sub> and T<sub>23</sub>) are in linkage disequilibrium (LD) with rs12079745 (D’ = 0.73), this TRS is independently associated with BP in a European-American population. By in vitro assays, we demonstrate that this allele mediates its effect via increasing polyadenylation at the A2 site. Furthermore, we show higher relative abundance of the A2-polyadenylated mRNA in human kidneys heterozygous for the T<sub>22</sub>G<sub>7</sub>T<sub>3</sub> allele than in T<sub>22</sub> allele than in T<sub>22</sub> alleles.

Therefore, we have identified a common human polymorphism in the DSE of ATP1B1 that may contribute to BP regulation by regulating ATP1B1 mRNA processing.

Materials and Methods

Subjects

Samples used in the association study were ascertained through the GenNet network of the Family Blood Pressure Program [9]. Phenotype characteristics of the GenNet samples have been previously published [9]. Briefly, subjects’ ages ranged between 18–50 years. Their mean SBP and DBP were 123.2 ± 17.6 and 77.0 ± 10.0 mmHg, respectively. Overall, 32% were clinically hypertensive with SBP/DBP over 140/90 mmHg or were taking anti-hypertension medication. To measure ATP1B1 expression, we used RNA from Epstein Barr Virus (EBV)-transformed lymphocytes from healthy Amish subjects from the Amish Family Longevity Study (AFLS) [28], and human kidneys from the National Disease Research Interchange, Philadelphia, PA (http://ndiresource.org/). Approval for the collection and use of the AFLS samples was provided by the University of Maryland, Baltimore (UMB) Institutional Review Board (IRB). Written informed consent was obtained from AFLS subjects for the use of their samples. The kidney samples used were exempted from IRB approval by the UMB IRB.

Genotyping

The TRS was amplified from human genomic DNA using the following primers: F-5’-Gacaaaagaaaaagaaaaattgac-3’ and R-5’-aaataaagacctaacaccacaggaa-3’. Sequencing was performed using the BigDye Terminator reaction chemistry (Applied Biosystems, Foster City, CA) according to the manufacturer’s recommendations. Sequences were analyzed using Sequencher 4.5 (Gene Codes, Ann Arbor, MI). The length of the TRS was deduced by electrophoresis of fluorescently labelled PCR product along with size markers. Genomic DNA was amplified using the following primers: F-5’-CTGTGTCTACTTTATGTGAGCAAGG-3’ (FAM labelled) and R-5’-CTCATCGATGGCGCTTAC-3’ and the GoTaq Colourless Master Mix (Promega, Madison, WI). Size was determined by electrophoresis of the PCR products mixed with Hi-Di Formamide (Applied Biosystems) and the Gene Scan 600Liz size standard (Applied Biosystems). The results were analyzed using GeneMapper 4.0 software (Applied Biosystems) and examples of the electropherograms of T<sub>22</sub> and T<sub>23</sub> are shown in Figure S1.

Polyadenylation assay

A 372 base pair (bp) fragment of the ATP1B1 3’UTR spanning the TRS was amplified from genomic DNA using the following primers: F-5’-GACAAAAGAAAAAGAAAAATTGAGC-3’, and R-5’-
AAATAAGACCTAACACCACAGGAA-3'. The products were cloned into the pCR4-TOPO vector (Invitrogen) and then into the pRIG vector (Figure 2A) using EcoRI digestion and ligation. Successful cloning was verified by colony PCR and sequencing. Polyadenylation activity was determined as described in Pan et al. [30]. Briefly, Human Embryonic Kidney 293T (HEK293T) cells grown in 6-well plates were transfected with the various pRIG vectors (1 µg plasmid in each well) using the Fugene 6 transfection kit (Roche). Cells were harvested 48 h after transfection by treatment with Trypsin-EDTA. When cells are transfected with the empty pRIG vector, red and green fluorescence proteins (RFP and eGFP, respectively) are equally expressed. Insertion of a polyadenylation signal-containing sequence into pRIG at the multiple cloning site (MCS) leads to the cleavage of the transcript between the RFP and eGFP open reading frames (ORFs) causing cells to express more RFP than eGFP. The ratio of RFP to eGFP in cells reflects the strength of the polyadenylation signal tested. The relative abundance of RFP and eGFP were measured at 530 nm and 585 nm, respectively, using the FACScalibur system (BD Biosciences, San Jose, CA).

Figure 1. Structure and genotyping of the T-rich sequence in ATP1B1. A) Genomic architecture of the ATP1B1 gene. The red box highlights the 3'UTR, which is enlarged and shown below. TGA: translational stop codon, G/A: SNP rs12079745, A2: A2 polyadenylation signal, TRS: T-rich element, A5: A5 polyadenylation signal. B) Representative Sanger sequencing chromatograms of the T23 and T12GT6 alleles. Alleles are named based on the portion of the TRS that is unique to each allele (denoted by a horizontal bar). C) Sequence alignment of the TRS across mammals. The A2 polyadenylation signal and the TRS are highlighted in red and green respectively. D) Aligned are the 2a consensus sequence of the CstF binding site [38] and the sequences of the TRS with the T23 and T12GT6 alleles. Asterisks mark the nucleotides in the TRS that deviate from the consensus CstF binding sequence. Y = C/T, N = A/C/G/T. doi: 10.1371/journal.pone.0076290.g001
RNA isolation and real-time PCR

RNA was isolated from sections of frozen kidneys and EBV-transformed lymphocytes using the RNeasy Mini Kit from Qiagen (Valencia, CA) according to the manufacturer’s instructions. RNA was treated with DNAse I, Amplification grade (Life Technologies, Grand Island, NY) and reverse-transcribed using oligodT primers and the Superscript III First-Strand Synthesis Supermix (Life Technologies). Real-time PCR was performed using the LightCycler 480 SYBR Green I Master (Roche, Indianapolis, IN) on a Roche LightCycler 480. The following primers were used to amplify the A2- and A5-polyadenylated mRNAs and are indicated in Figure 3A: A2/A5-F-5’-CTGGCCCCTAAGTATTGCT-3’, A2-R-5’-TTTTTTTTTTTTTTACCATGCCAGTTTTATTC-3’, A5-R-5’-CACCACAGGAAAAGACTATGGA-3’. The reverse primer for the A2-polyadenylated mRNA had an additional 5’ T14 adaptor in order to selectively amplify A2-polyadenylated mRNA. Melt curve analysis was used to ensure a specific product. Expression was normalized to total ATP1B1 transcript levels (amplified using primers 5’-ATGTCCCTCTCCGCTAAGT-3’ and 5’-CAGAGGAAAAACCAGGGGAGT-3’) due to the known differences in Atp1b1 mRNA levels in different regions of the kidney [29]. All transcripts were quantified based on standard curves generated from pooled cDNA and using the LightCycler 480 Relative Quantification software (Roche). Each sample was tested in triplicate and in at least 2 independent experiments. Data from subjects without the T12GT3GT6 allele (0) and those with at least 1 copy of the T12GT3GT6 allele (1) are presented in all figures. Significance was calculated using Student’s t-test. p < 0.05 was considered significant.

Genetic analysis

Association analyses were performed using MMAP (http://edn.umaryland.edu/mmap/index.php) [31]. Briefly, this software uses a variance component model that assesses the effect of genotype, as an additive effect, on the phenotype of interest while simultaneously estimating the effects of age, age^2, sex, and a polygenic component to account for phenotypic correlation due to relatedness. An (n-1)-degree-of-freedom t test was used to assess the significance of the measured genotype. The genotype was coded as the dosage of the effect alleles shown in Table 1, with the remaining alleles being collapsed to the non-effect allele. For those GenNet subjects taking anti-hypertension medication (22%), analyses were performed either by removing these individuals or by adding the average treatment effect to raw BP readings (10 and 5 mmHg to SBP and DBP, respectively) [32]. LD calculations between TRS and rs12079745 were made using LD [33,34]. Sequence alignment was performed using ClustalW [35,36] and orthologous sequences were obtained.
Figure 3. The T\textsubscript{12}GT\textsubscript{3}GT\textsubscript{6} allele regulates the polyadenylation of ATP1B1 mRNA in human tissues. A) Schematic of the primer design used to distinguish between the A2- and A5- polyadenylated mRNAs. The A2-R primer had a 5' T\textsubscript{15} tail to selectively amplify A2-polyadenylated mRNA. Total ATP1B1 mRNA was measured using primers that amplified across the exon 5/exon 6 junction. B) Real-time PCR of mRNA from human kidneys. The A2-polyadenylated transcript was quantified relative to total ATP1B1 mRNA. 0 and 1 indicate subjects without the T\textsubscript{12}GT\textsubscript{3}GT\textsubscript{6} allele and with 1 copy of the T\textsubscript{12}GT\textsubscript{3}GT\textsubscript{6} allele, respectively. Error bars represent standard errors. *p < 0.0001. C) Real-time PCR of mRNA from human lymphocytes. Expression of the A2- and A5-polyadenylated mRNAs relative to total ATP1B1 mRNA in lymphocytes from subjects with 0 or at least 1 copy of the T\textsubscript{12}GT\textsubscript{3}GT\textsubscript{6} allele. *p = 0.0001 between the A5-polyadenylated transcript levels in the 2 groups. Data represent the average across samples from each genotype group with standard error. D) Expression of ATP1B1 protein in lymphocytes from individuals with 0 or at least 1 copy of the T\textsubscript{12}GT\textsubscript{3}GT\textsubscript{6} allele. Tubulin was used as a loading control. The graph quantifies the average level of ATP1B1 protein relative to tubulin with standard error. *p = 0.05.

doi: 10.1371/journal.pone.0076290.g003
from the UCSC genome browser (www.genome.ucsc.edu) from assemblies hg18, panTro2, rheMac2, mm9, mm4, bosTau4, canFam2, monDom5, oviArt1, and myoLuc2. The pteVam1 sequence was obtained from scaffold_1928 from Ensembl (www.ensembl.org).

**Protein extraction and Western blot analysis**

Protein was extracted by sonicating lymphocytes in lysis buffer (25 mM Tris-HCl, pH 7.5, 1% SDS, protease inhibitor tablet [Quality Biologicals, Gaithersburg, MD], and PIC). Protein was quantified using the BCA Protein Assay Kit (Thermo Scientific, Rockford, IL) according to the manufacturer’s instructions. Twenty micrograms of protein was deglycosylated using PNGaseF (New England Biolabs, Ipswich, MA). The deglycosylated lysates were separated on a 10% polyacrylamide gel (Bio-Rad, Hercules, CA) and transferred to a polyvinylidene difluoride membrane (PVDF, Bio-Rad). The membrane was blocked overnight, treated with a monoclonal ATP1B1 antibody (1:5000, MA3-930, Thermo Scientific) according to the manufacturer’s instructions. Twenty micrograms of protein was deglycosylated using PNGaseF (New England Biolabs, Ipswich, MA). The deglycosylated lysates were separated on a 10% polyacrylamide gel (Bio-Rad, Hercules, CA) and transferred to a polyvinylidene difluoride membrane (PVDF, Bio-Rad). The membrane was blocked overnight, treated with a monoclonal ATP1B1 antibody (1:5000, MA3-930, Thermo Scientific) for 1 h at room temperature, washed and incubated with an anti-mouse secondary antibody (1:5000, 115-035-003, Jackson ImmunoResearch) for 1 h at room temperature. The membrane was developed using the Chemiluminescent Western blotting kit (Thermo Scientific) and the protein was visualized on the FluorChem Q system (ProteinSimple, Santa Clara, CA). When needed, the membrane was stripped with Restore Plus Western Blot Stripper Buffer (Thermo Scientific).

**Results**

A multi-allelic T-rich sequence in the 3’UTR of ATP1B1 is a putative DSE

Examination of the 3’UTR of ATP1B1 revealed a highly polymorphic T-rich sequence (TRS) 13 nucleotides downstream of the A2 polyadenylation signal (Figure 1A). This TRS has 22 T nucleotides in the reference genome (hg18), but the presence of 5 annotated biallelic SNPs and insertions/deletions in dbSNP (rs34447553, rs1779795, rs1779796, rs56167655, and rs4656649) suggests that this TRS might be structurally complex and contain multiple alleles. To identify sequence variants in the TRS, we sequenced this region in 1,010 European American samples collected through the GenNet network. We identified 7 alleles in this TRS that vary the length of the T-track and the number of G nucleotides that interrupt the T-track: T$_{22}$, T$_{23}$, T$_{24}$, T$_{25}$, T$_{26}$, T$_{27}$, and T$_{12}$GT$_{3}$GT$_{6}$ (representative sequence traces of T$_{23}$ and T$_{12}$GT$_{3}$GT$_{6}$ are shown in Figure 1B). The most common alleles were T$_{23}$ and T$_{12}$GT$_{3}$GT$_{6}$ with allele frequencies of 0.82 and 0.12, respectively. The remaining alleles had allele frequencies ranging from 0.04 to 0.001.

Sequence inspection suggested a potential functional role for the TRS alleles. Species in which the A2 polyadenylation signal adheres to the AATAAA consensus polyadenylation sequence have a T-rich sequence downstream of the polyadenylation signal, whereas species such as mouse that have a variant polyadenylation signal have a relatively T-poor region downstream of the polyadenylation site (Figure 1C) [37]. This suggests that the TRS and the A2 polyadenylation signal may be functionally linked, and that the TRS may encode the uridine-rich element that serves as the DSE of the A2 polyadenylation site. In fact, this TRS contains a putative CstF binding site that is 72.2% similar to the 2a consensus CstF binding site (Figure 1D) [38]. Interestingly, the T$_{12}$GT$_{3}$GT$_{6}$ sequence is more similar to the 2a consensus CstF binding site than the alleles with pure T-track (83.3% vs 72.2%, Figure 1D). Therefore, we investigated whether the TRS alleles may be responsible for the association signal seen in ATP1B1 in our previous study, and whether any such association may be mediated through the effect of these alleles on the polyadenylation of the ATP1B1 mRNA.

**Alleles of the TRS are associated with BP independently of rs12079745**

In order to determine whether the TRS alleles have an effect on BP, we first performed an association analysis of the TRS alleles with systolic BP (SBP) and diastolic BP (DBP) in the GenNet samples in which the original association was identified. Subjects on HTN medication either had their SBP

### Table 1. Association analysis of the TRS alleles with blood pressure.

| Model       | Variant | Effect allele | SBP (n = 830) | DBP (n = 831) | SBP (n = 647) | DBP (n = 648) |
|-------------|---------|---------------|---------------|---------------|---------------|---------------|
| Single locus| rs12079745 | A            | 2.8           | 0.10          | 0.32          | 1.0           |
|            |         |               | 4.9           | 0.001         | 2.4           | 0.022         |
|            | TRS     | T$_{12}$GT$_{3}$GT$_{6}$ | 3.3           | 0.014         | 2.4           | 0.003         |
|            |         |               | 2.4           | 0.011         | 1.1           | 0.029         |
| Joint Analysis | rs12079745 & rs12079745 | A | 3.2 | 0.07 | 1.3 | 0.21 |
| TRS         |         | T$_{12}$GT$_{3}$GT$_{6}$ | 3.6 | 0.009 | 2.5 | 0.002 |
|            |         |               | 2.4           | 0.06          | 2.1           | 0.019         |

SBP: systolic blood pressure, DBP: diastolic blood pressure, *effect sizes are in mmHg

doi: 10.1371/journal.pone.0076290.t001
and DBP adjusted (see Methods) or were excluded from the analysis. Of primary interest were the T<sub>23</sub> and T<sub>12</sub>G<sub>T</sub>S<sub>GT</sub> alleles because they were the most common alleles in this population. Since too few subjects carried alleles T<sub>22</sub>, T<sub>24</sub>, T<sub>25</sub>, T<sub>26</sub>, and T<sub>27</sub> to provide sufficient power for association analysis, we analyzed our data by collapsing all non-T<sub>12</sub>G<sub>T</sub>S<sub>GT</sub> alleles as the reference allele (Table 1) and by removing subjects carrying alleles other than T<sub>23</sub> and T<sub>12</sub>G<sub>T</sub>S<sub>GT</sub> from the analysis (Table S1).

As shown in Table 1, when analyzed individually, rs12079745 and the TRS were both significantly associated with SBP and DBP (Single locus model). The rs12079745 A allele was associated with increased SBP (beta = 4.9, p = 0.001) and increased DBP (beta = 2.4, p = 0.022) in unmedicated subjects, but not in the full sample. On the other hand, the TRS T<sub>12</sub>G<sub>T</sub>S<sub>GT</sub> allele was associated with SBP (beta = 3.3, p = 0.014) and DBP (beta = 2.4, p value = 0.003) in the full sample, and only DBP (beta = 0.19, p = 0.029) in unmedicated subjects. Removing alleles with low allele frequencies gave a similar result (Table S1). Thus, both rs12079745 and the TRS were both significantly associated with DBP as compared to T<sub>23</sub> and T<sub>12</sub>G<sub>T</sub>S<sub>GT</sub> from the analysis (Table S1).

Next, we tested whether the T<sub>12</sub>G<sub>T</sub>S<sub>GT</sub> allele also regulates alternative polyadenylation of ATP1B1 in vivo. We determined the abundance of the A2- and A5-polyadenylated ATP1B1 mRNAs relative to total ATP1B1 levels in human kidneys (see Methods). Human kidneys express high levels of ATP1B1 [21], and are intimately involved in BP regulation [39]. Among the human kidneys available, 4 were heterozygous for the T<sub>12</sub>G<sub>T</sub>S<sub>GT</sub> allele while the rest were homozygous T<sub>23</sub>T<sub>23</sub>. All the kidneys used were homozygous for the G allele at rs12079745. We measured the relative abundance of the A2- and A5-polyadenylated mRNAs in these 4 heterozygotic kidneys and compared it to that of 5 T<sub>23</sub>T<sub>23</sub> kidneys by quantitative PCR using the primers depicted in Figure 2A. As shown in Figure 3B, the relative abundance of the A2-polyadenylated mRNA was significantly twice as high in kidneys with the T<sub>12</sub>G<sub>T</sub>S<sub>GT</sub> allele than in those lacking T<sub>12</sub>G<sub>T</sub>S<sub>GT</sub> alleles (p < 0.0001). The A5-polyadenylated transcript was undetectable in the kidneys tested.

We also quantified the relative abundance of the A2- and A5-polyadenylated mRNAs in lymphocytes, for which more samples overall and more carriers of the T<sub>12</sub>G<sub>T</sub>S<sub>GT</sub> allele were available. Lymphocytes also express Na-K ATPase pumps, albeit at a lower level than skeletal muscle and kidney tissue [15]. Furthermore, ATP1B1 mRNA levels have been shown to be detectable at higher levels in mature lymphocytes than in immature lymphocytes [40]. Since only 3 samples were homozygous for the T<sub>12</sub>G<sub>T</sub>S<sub>GT</sub> allele, we combined data from the heterozygote and homozygote T<sub>12</sub>G<sub>T</sub>S<sub>GT</sub> samples for the subsequent analysis. In contrast to human kidneys, in lymphocytes, the relative abundance of the A2-polyadenylated mRNA was independent of TRS genotypes, whereas the abundance of the A5-polyadenylated mRNA was significantly higher in the presence of the T<sub>12</sub>G<sub>T</sub>S<sub>GT</sub> allele (p = 0.0001) (Figure 3C). These results suggest that the TRS alleles modulate alternative polyadenylation of ATP1B1 transcripts in human tissues but their effects on mRNA isoform abundance differ across tissue types.

**Discussion**

We have identified a novel polymorphic and functional TRS in the 3’UTR of ATP1B1. Using human tissue samples and *in vitro* assays, we have shown that the alleles of this TRS regulate the alternative polyadenylation of ATP1B1. We also showed that the alleles are associated with BP in a European-
American population. Therefore, the TRS alleles likely regulate BP by regulating alternative polyadenylation of ATP1B1.

Polymorphisms in the TRS likely mediate their effect on polyadenylation by modulating the binding of cleavage stimulating factors. The T\textsubscript{G}GT\textsubscript{G}GT\textsubscript{A} allele of the TRS is more similar to the consensus binding site of the cleavage stimulating factor CstF64 than the T\textsubscript{G}GT\textsubscript{A} allele (Figure 1D) [38]. While Takagaki and Manley found that both GU- and U-rich sequences are recognized by CstF64 [41], the yeast homolog of CstF64, Rna15, was found to have a preference for GU-containing RNA [42], suggesting that the GUUs in T\textsubscript{G}GT\textsubscript{G}GT\textsubscript{A} may increase the affinity of the TRS for CstF64, thereby enhancing A2 polyadenylation site usage. Such an increased affinity for CstF64 may in turn lead to the increased abundance of the A2-polyadenylated ATP1B1 mRNA. In fact, regulation of CstF64 binding has been shown to regulate the switch between the synthesis of the secreted and membrane bound forms of immunoglobulin [43]. Interestingly, while multiple alleles that vary the length of the T-track were found, we did not find any alleles similar to T\textsubscript{G}GT\textsubscript{G}GT\textsubscript{A} that varied the length of the T-track, such as T\textsubscript{G}GT\textsubscript{G}GT\textsubscript{A}, again suggesting that the spacing of the 2 additional G nucleotides is functionally important.

Alternatively, sequence variants in the TRS may also cause changes to the secondary structure of ATP1B1 mRNA. The RNA folding program MFOLD predicts differences in the secondary structure of ATP1B1 mRNA containing the U\textsubscript{G}GU \textsubscript{GU} or U\textsubscript{G}GU \textsubscript{ GU} sequences (Figure S2 and Information S1) [44]. Effects on mRNA 3’ end processing or mRNA stability, mediated through changes in the secondary structure of either the A5-polyadenylated mRNA or the premature mRNA, may contribute to the different levels of A5-polyadenylated mRNA noted in lymphocytes. Further studies will be required to determine if the T\textsubscript{G}GT\textsubscript{G}GT\textsubscript{A} allele impacts ATP1B1 mRNA post-transcriptionally through mechanisms other than polyadenylation.

Our association analyses showed that while rs12079745 and the TRS are in LD, these 2 sites are independently associated with BP. The effect or higher BP-associated allele of the TRS (T\textsubscript{G}GT\textsubscript{G}GT\textsubscript{A}) and that of the SNP (A) are not found on the same haplotype in our study population. Furthermore, the association between rs12079745 and SBP is stronger than that between this SNP and DBP, whereas the TRS is more strongly associated with DBP than with SBP. In fact, having both loci in the same model provides stronger association to BP than having only one locus in the model. While rs12079745 is in strong LD with several other SNPs in and near the 3’UTR of ATP1B1, none of these SNPs are located in a sequence context in which a potential function can be computationally predicted and experimentally tested. However, these SNPs might contribute to the regulation of ATP1B1 expression and BP levels via other 3’UTR functions unrelated to alternative polyadenylation. Indeed, Wei et al. demonstrated the role for a common 3’UTR variant in ATP6V0A1 that alters hypertension risk by creating a microRNA binding site that affects ATP6V0A1 mRNA expression [45]. Additional studies are needed to determine how rs12079745 or another site in LD with this SNP may regulate ATP1B1 expression and explain the association between rs12079745 and BP.

Since the β1 subunit encoded by ATP1B1 is the rate-limiting factor in the assembly and activity of Na-K ATPase, an increase in ATP1B1 mRNA levels in the presence of the T\textsubscript{G}GT\textsubscript{G}GT\textsubscript{A} allele is expected to increase the abundance and function of Na-K ATPase activity [46]. Increased Na-K ATPase activities have been correlated with higher BP through studies in animal models. In the Milan Hypertensive Rat, an increase in ATP1B1 expression and Na-K ATPase activity in the kidney precedes the onset of hypertension [47,48]. Knock-in gene targeting of the alpha-2 isoform of Na-K ATPase in mice makes them insensitive to ouabain-induced hypertension [49]. A conditional cardiac knockout of Atp1b1 in a mouse model also demonstrated that Atp1b1 is crucial for cardiac growth, contractility, and ouabain-sensitivity [50]. Furthermore, a 2-kidney, 1-clip rat model of hypertension also showed increased Na-K ATPase levels and increased phosphorylation of Na-K ATPase in renovascular hypertension [51]. The importance of Na-K ATPase in human essential hypertension, due to its heterogeneous condition, is less well established. However, rostafuroxin, a compound that lowers blood pressure by restoring normal renal and vascular Na-K ATPase functions, is currently in clinical trial and might be effective in a genotype-specific manner [52]. Measuring the effect of the TRS alleles on Na-K ATPase activity was not feasible because the TRS appears to influence the polyadenylation of ATP1B1 transcripts in a tissue specific manner and human tissues most relevant for blood pressure homeostasis were not available for study. Therefore, further studies will be required to determine the effect of the identified variants on Na-K ATPase activity.

While the A5-polyadenylated mRNA is less abundant than the A2-polyadenylated mRNA in most human tissues that express high levels of ATP1B1 (Figure S3 and Information S1), we were unable to detect A5-polyadenylated mRNA in our collection of kidneys. This can be due to the low abundance of this mRNA in the kidney or suboptimal quality of RNA obtained from the kidneys used in this study. The frozen kidneys from which RNA was extracted in this study were obtained post-mortem, from donors who died from a variety of causes unrelated to acute kidney diseases. Although the overall quality of the RNAs was acceptable for standard molecular analysis, the intactness of transcripts with destabilizing motifs, such as those present in the A5-polyadenylated mRNA, is unknown. Furthermore, the kidney is a complex organ and many genes are differentially regulated throughout the nephron [29]. The kidney sections available to us varied in composition and tissue heterogeneity (versus studying only renal cortex, for example). This may have contributed to our inability to detect the A5-polyadenylated transcript. Therefore, we cannot exclude the possibility that technical limitations prevented us from determining how the change in the relative abundance of the A2-polyadenylated mRNA affects that of the A5-polyadenylated mRNA in human kidneys. Nevertheless, we were still able to detect an increase in A2-polyadenylated mRNA levels in kidneys containing the T\textsubscript{G}GT\textsubscript{G}GT\textsubscript{A} allele, suggesting that the T\textsubscript{G}GT\textsubscript{G}GT\textsubscript{A} allele does indeed have an effect on ATP1B1 polyadenylation.

In contrast, the A5-polyadenylated mRNA was detectable in freshly harvested lymphocytes, where only the relative
abundance of the A5-polyadenylated mRNA was higher in the presence of the T3GT6 allele. This may be due to the preferential use of one polyadenylation site versus another in different tissues. Indeed, over 70% of tandem 3'UTR polyadenylation sites in humans have been shown to be used in a tissue-specific manner [53]. Furthermore, the lower abundance of Na-K ATPase pumps in lymphocytes versus in kidneys may contribute to differences in the regulation of the ATP1B1 gene in lymphocytes versus kidneys [15].

Our study is one of the few examples, to our knowledge, of a polymorphic DSE involved in allele-specific gene expression via alternative polyadenylation that is associated with a clinically important phenotype. Uitte de Willige et al., found a C>T transversion (TACCT>TATCT, rs2066865) in the DSE of the fibrinogen γ (FGG) gene that has an impact on the putative CstF64 binding site of FGG similar to what we see with the T23 to T12GT6 change in the TRS [54]. This TATCT allele is also relatively common (0.22 and 0.50 in HapMap European-Americans and Asians, respectively), alters the fibrinogen γ'/γA ratio by regulating the alternative polyadenylation of FGG mRNA, and is associated with an increased risk of deep vein thrombosis [54,55]. Another example is provided by a rare gain-of-function C20221T mutation 11 bp downstream of the cleavage site of the prothrombin F2 gene. This mutation alters a TGCT sequence to a TGTT motif and shows greater mRNA expression than the wild-type sequence and was found in patients with abnormal thrombosis [56]. Given that an estimated 5–10% of all GWA hits are located in the 3'UTR [57], and 50% of all protein-coding genes possess multiple polyadenylation signals and are partially regulated through alternative polyadenylation [58], there are certainly more such variants to be uncovered that are clinically relevant and/or can yield important biological insights. In fact, many risk alleles from GWAS studies are correlated with alleles that can cause alternate polyadenylation [59]. Because genotyping this complex TRS site is more technically demanding than SNP genotyping and sequencing, the allele frequency of the TRS in other populations is still unknown. In fact, we identified another allele, T16GT6, in African Americans (data not shown) but its impact on polyadenylation and BP is still unknown. Furthermore, simple repeat sequences, such as the TRS in ATP1B1, are often prone to mutate. Given how polymorphic this TRS is (8 alleles identified so far), it is highly likely that other DSEs are also polymorphic and that the regulation of alternative polyadenylation may be a relatively common mechanism in the susceptibility of quantitative traits, such as BP, across human populations.

We have identified novel polymorphisms in the 3'UTR of ATP1B1 that are associated with BP. We have also shown that these polymorphisms alter what is most likely the DSE of the gene, and that these polymorphisms mediate changes in the relative abundance of differentially polyadenylated ATP1B1 transcripts by regulating alternative polyadenylation of the gene.

**Supporting Information**

**Figure S1.** Determination of the T track genotypes. Electropherograms from GeneMapper for the T23 and T24 alleles. (TIF)

**Figure S2.** Predicted mRNA secondary structures of the ATP1B1 mRNA transcripts with the U23 and U6 GU GU6 sequences. (A) MFOLD prediction of the secondary structure of the mRNA transcript containing the U23 sequence. (B) MFOLD prediction of the secondary structure of the mRNA transcript containing the U6 GU GU6 sequence. The locations of A2 polyadenylation signal sequence and the TRS are indicated. (TIF)

**Figure S3.** Expression of the A2- and A5-polyadenylated ATP1B1 transcripts in a human tissue panel. Real-time PCR analysis of the levels of the A2- and A5-polyadenylated ATP1B1 transcripts relative to GAPDH mRNA levels in a human tissue panel. (TIF)

**Information S1. Supporting Methods.** (DOC)

**Table S1.** Association analysis of only the T23 and T12GT6 alleles with blood pressure. Other alleles with frequencies <0.05 were removed from the model. (DOC)

**Acknowledgements**

The authors would like to thank Xiaochun Wang for technical assistance.

**Author Contributions**

Conceived and designed the experiments: MKP KB ZP BT YCC. Performed the experiments: MKP KB. Analyzed the data: MKP JRO YCC. Contributed reagents/materials/analysis tools: JRO ABW AC BT. Wrote the manuscript: MKP YCC.
1. Keamy PM, Whelton M, Reynolds K, Muntner P, Whelton PK et al. (2005) Global burden of hypertension: Analysis of worldwide data. Lancet 365: 217-223. doi:10.1016/S0140-6736(05)17744-1. PubMed: 15652604.

2. Tobin MD, Raleigh SM, Newhouse S, Braund P, Bodycote C et al. (2005) Association of WNK1 gene polymorphisms and haplotypes with arterial blood pressure in the general population. Circulation 112: 3423-3429. doi:10.1161/CIRCULATIONAHA.105.555474. PubMed: 16301342.

3. Pilga G, Chen WM, Scuteri A, Orru M, Alabi G et al. (2006) Heterotrimorphism of cardiovascular and personality traits in 6,148 sardinians. PLOS Genet 2: e132. doi:10.1371/journal.pgen.0020132. PubMed: 16934002.

4. Levy D, DeStefano AL, Larson MG, O’Donnell CJ, Lifton RP et al. (2000) 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. doi:10.1161/01.HYP.36.4.477. PubMed: 11040222.

5. van Rijn MJ, Schut AF, Aulchenko YS, Deimun J, Sayed-Tabatabae FA et al. (2007) Heterotrimorphism of blood pressure traits and the genetic contribution to blood pressure variance explained by four blood-pressure-related genes. J Hypertens 25: 565-570. doi:10.1097/HJH.0b013e32801449fb. PubMed: 17267962.

6. Binder A (2007) A review of the genetics of essential hypertension. Curr Opin Cardiol 22: 176-184. doi:10.1097/HCO.0b013e3280135f79. PubMed: 17413273.

7. Ehtet GB, Munroe PB, Rice KM, Bochud M, Johnson AD et al. (2001) Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. Nature 478: 103-109. doi:10.1038/nature04105. PubMed: 21090115.

8. Gentino J, Rac G, Freedman BI (2012) Novel findings and future directions on the genetics of hypertension. Curr Opin Nephrol Hypertens, 21: 500–7. PubMed: 22614628.

9. Chang YP, Liu X, Kim JD, Ikeda MA, Layton MR et al. (2000) Multiple genes for the human Na,K-ATPase beta subunit. Genomics 5: 25-36. doi:10.1006/geno.2000.6401. PubMed: 11161799.

10. Stoll M, Kwitek-Black AE, Cowley AW, Harris EL, Harrap SB et al. (2000) Target regions for human hypertension via comparative genome. Genome Res 10: 473-482. doi:10.1101/gr.4.4.73. PubMed: 10779487.

11. Kochunov P, Glahn D, Lancaster J, Winkler A Jr, KJW et al. (2010) Whole brain and regional hyperintense white matter volume and blood pressure: Overlap of genetic loci produced by bivariate, whole-genome linkage analyses. Stroke 41: 2137-2142.

12. Perola M, Kainulainen K, Pajukanta P, Tervilliger JD, Hiekkanlina T et al. (2000) Genome-wide scan of predisposing loci for increased diastolic blood pressure in finnish siblings. J Hypertens 18: 1579-1585. doi:10.1097/00004872-200018110-00008. PubMed: 11081770.

13. Stoll M, Kweit-Black AE, Cowley AW, Harris EL, Harrap SB, et al. (2000) Target regions for human hypertension via comparative genomics. Genome Res 10: 473-482. doi:10.1101/gr.4.4.73. PubMed: 10779487.

14. Sugiyama F, Churchill GA, Higgins DC, Johns C, Makaritski BP et al. (2001) Concordance of murine quantitative trait loci for salt-induced hypertension with rat and human loci. Genomics 71: 70-77. doi:10.1006/geno.2000.6401. PubMed: 11161799.

15. Clausen T (1998) Clinical and therapeutic significance of the nα,K+ pump*. Clin Sci 95: 3-17.

16. Jørgensen PL (1986) Structure, function and regulation of nα,K-ATPase in the kidney. Kidney Int 29: 10-20. doi:10.1038/kj.1986.37. PubMed: 2421041.

17. McDonough AA, Geering K, Farley RA (1990) The sodium pump needs its beta subunit. FASEB J 4: 1596-1598. PubMed: 2156741.

18. Faruque MU, Chen G, Dounatay A, Huang H, Zhou J et al. (2011) Association of ATP1B1, RGS5 and SELE polymorphisms with hypertension and blood pressure in african-americans. J Hypertens 29: 176-184. doi: 10.1097/HCO.0b013e3280d357f9. PubMed: 21925375.

19. Cost AD, Chovit LK, Shull GE, Lingel JB (1987) Characterization of two genes for the human nα,K-ATPase beta subunit. Genomics 5: 445-453. doi:10.1016/0888-7543(89)90008-6. PubMed: 2559024.

20. Young RM, Shull GE, Lingel JB (1987) Multiple mRNAs from rat kidney and brain encode a single nα,K-ATPase beta subunit protein. J Biol Chem 262: 4905-4910..
specificity in transcriptional 3'-end processing factors. Nucleic Acids Res 38: 3119-3132. doi:10.1093/nar/gkq002. PubMed: 20097654.

43. Takagaki Y, Seipelt RL, Peterson ML, Manley JL (1996) The polyadenylation factor CstF-64 regulates alternative processing of IgM heavy chain pre-mRNA during B cell differentiation. Cell 87: 941-952. doi:10.1016/S0092-8674(00)82000-0. PubMed: 8945520.

44. Zuger M (2003) Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res 31: 3406-3415. doi:10.1093/nar/gkg595. PubMed: 12824337.

45. Wei Z, Biswas N, Wang L, Couré M, Zhang K et al. (2011) A common genetic variant in the 3'-UTR of vacuolar H+-ATPase ATP6V0A1 creates a micro-RNA motif to alter chromogranin A processing and hypertension risk. Circ Cardiovasc Genet 4: 381-389. doi:10.1161/ CIRCGENETICS.111.955767. PubMed: 21558123.

46. McDonough AA, Brown TA, Horowitz B, Chiu R, Schlotterbeck J et al. (1988) Thyroid hormone coordinately regulates Na+-K+-ATPase alpha- and beta-subunit mRNA levels in kidney. Am J Physiol 254: 323-329. PubMed: 2831727.

47. Ferrandi M, Tripodi G, Salardi S, Florio M, Modica R et al. (1996) Renal Na,K-ATPase in genetic hypertension. Hypertension 28: 1018-1025. doi:10.116101.HYP.28.6.1018. PubMed: 8952591.

48. Metz ML, Syrén ML, Asaell BM, Sereni F, Aperia A (1991) Increased renal tubular Na+-K+-ATPase activity in Milan hypertensive rats in the prehypertensive period. Pediatr Nephrol 5: 700-703. doi:10.1007/ BF00857877. PubMed: 1662981.

49. Dostanic I, Paul RJ, Lorenz JN, Theriault S, Huysse JWV et al. (2005) The alpha2-isoform of Na-K-ATPase mediates ouabain-induced hypertension in mice and increased vascular contractility in vitro. Am J Physiol Heart Circ Physiol 288: H477-H485. PubMed: 15458945.

50. Barwe SP, Jordan MC, Skay A, Inge L, Rajasekaran SA et al. (2009) Dysfunction of ouabain-induced cardiac contractility in mice with heartspecific ablation of Na-K-ATPase beta1-subunit. J Mol Cell Cardiol 47: 552-560. doi:10.1016/j.yjmcc.2009.07.018. PubMed: 19683723.

51. Kava L, Rossi NF, Mattingly R, Yingst DR (2012) Increased expression of Na,K-ATPase and a selective increase in phosphorylation at ser-11 in the cortex of the 2-kidney, 1-clip hypertensive rat. Am J Hypertens 25: 487-491. doi:10.1038/ajh.2011.247. PubMed: 22237155.

52. Lanzani C, Citterio L, Glorioso N, Manunta P, Tripodi G et al. (2010) Adducin- and ouabain-related gene variants predict the antihypertensive activity of rostafuroxin, part 2: Clinical studies. Sci Transl Med 2: 59ra87.

53. Wang ET, Sandberg R, Luo S, Krehbukova I, Zhang L et al. (2008) Alternative isoform regulation in human tissue transcriptomes. Nature 456: 470-476. doi:10.1038/nature07509. PubMed: 18978772.

54. Williams SuD, Visser MCD, Vissing-Duistermaat JJ, Rosendaal FR, Vos HL et al. (2005) Genetic variation in the fibrinogen gamma gene increases the risk for deep venous thrombosis by reducing plasma fibrinogen gamma' levels. Blood 106: 4176-4183. doi:10.1182/ blood-2005-05-2180. PubMed: 16144795.

55. Willigie SuD, Rietveld IM, Visser MCD, Vos HL, Bertina RM (2007) Polymorphism 10034C>T is located in a region regulating polyadenylation of FGG transcripts and influences the fibrinogen gamma'/gammaA mRNA ratio. J Thromb Haemost 5: 1243-1249. doi:10.1111/j.1538-7836.2007.02566.x. PubMed: 17403086.

56. Danckwardt S, Gehring N, Neu-Yilik G, Hundsdorfer P, Pforrrich M et al. (2004) The prothrombin 3'end formation signal reveals a unique architecture that is sensitive to thrombophilic gain-of-function mutations. Blood 104: 428-435. doi:10.1182/blood-2003-08-2894. PubMed: 15059842.

57. Manolio TA (2010) Genomewide association studies and assessment of the risk of disease. N Engl J Med 363: 166-176. doi:10.1056/ NEJMra0905980. PubMed: 20847212.

58. Tian B, Hu J, Zhang H, Lutz CS (2005) A large-scale analysis of mRNA polyadenylation of human and mouse genes. Nucleic Acids Res 33: 201-212. doi:10.1093/nar/gki158. PubMed: 15647503.

59. Thomas LF, Saetrom P (2012) Single nucleotide polymorphisms can create alternative polyadenylation signals and affect gene expression through loss of microRNA-regulation. PLOS Comput Biol 8: e1002621. PubMed: 22915998.