

Analysis of putative cis-regulatory elements regulating blood pressure variation

Priyanka Nandakumar¹, Dongwon Lee^{1,2}, Thomas J. Hoffmann^{3,4}, Georg B. Ehret^{1,2,5}, Dan Arking¹, Dilrini Ranatunga⁶, Man Li⁷, Megan L. Grove⁸, Eric Boerwinkle⁸, Catherine Schaefer⁶, Pui-Yan Kwok⁴, Carlos Iribarren⁶, Neil Risch^{3,4,6}, Aravinda Chakravarti^{1,2}

¹ McKusick-Nathans Institute of Genetic Medicine, Baltimore, MD, USA.

² Center for Human Genetics and Genomics, NYU School of Medicine, New York, NY, USA.

³ Department of Epidemiology and Biostatistics, University of California San Francisco, San Francisco, CA, USA.

⁴ Institute for Human Genetics, University of California San Francisco, San Francisco, CA, USA.

⁵ Cardiology, Department of Specialties of Internal Medicine, University of Geneva, Geneva, Switzerland.

⁶ Kaiser Permanente Northern California Division of Research, Oakland, California, USA.

⁷ Division of Nephrology and Department of Human Genetics, University of Utah, Salt Lake City, Utah, USA.

⁸ Human Genetics Center, University of Texas Health Science Center, Houston, Texas, USA.

* Corresponding author

E-mail: aravinda.chakravarti@nyulangone.org (AC)

Abstract

Hundreds of loci have been associated with blood pressure traits from many genome-wide association studies. We identified an enrichment of these loci in aorta and tibial artery expression quantitative trait loci in our previous work in ~100,000 Genetic Epidemiology Research on Aging (GERA) study participants. In the present study, we subsequently focused on determining putative regulatory regions for these and other tissues of relevance to blood pressure, to both fine-map these loci by pinpointing genes and variants of functional interest within them, and to identify any novel genes.

We constructed maps of putative cis-regulatory elements using publicly available open chromatin data for the heart, aorta and tibial arteries, and multiple kidney cell types. Sequence variants within these regions may be evaluated quantitatively for their tissue- or cell-type-specific regulatory impact using deltaSVM functional scores, as described in our previous work. In order to identify genes of interest, we aggregate these variants in these putative cis-regulatory elements within 50Kb of the start or end of genes considered as “expressed” in these tissues or cell types using publicly available gene expression data, and use the deltaSVM scores as weights in the well-known group-wise sequence kernel association test (SKAT). We test for association with both blood pressure traits as well as expression within these tissues or cell types of interest, and identify several genes, including *MTHFR*, *C10orf32*, *CSK*, *NOV*, *ULK4*, *SDCCAG8*, *SCAMP5*, *RPP25*, *HDGFRP3*, *VPS37B*, and *PPCDC*. Although our study centers on blood pressure traits, we additionally examined two known genes, *SCN5A* and *NOS1AP* involved in the cardiac trait QT interval, in the Atherosclerosis Risk in Communities Study (ARIC), as a positive control, and observed an expected heart-specific effect. Thus, our method may be used to

identify variants and genes for further functional testing using tissue- or cell-type-specific putative regulatory information.

Author Summary

Sequence change in genes (“variants”) are linked to the presence and severity of different traits or diseases. However, as genes may be expressed in different tissues and at different times and degrees, using this information is expected to more accurately identify genes of interest. Variants within the genes are essential, but also in the sequences (“regulatory elements”) that control the genes’ expression in different tissues or cell types. In this study, we aim to use this information about expression and variants potentially involved in gene expression regulation to better pinpoint genes and variants in regulatory elements of interest for blood pressure regulation. We do so by taking advantage of such data that are publicly available, and use methods to combine information about variants in aggregate within a gene’s putative regulatory elements in tissues thought to be relevant for blood pressure, and identify several genes, meant to enable experimental follow-up.

Introduction

Genetic studies of complex disorders have identified hundreds to thousands of sequence variants in the human non-coding genome. However, despite significant mapping progress, we do not yet know the identity of most of the underlying genes and variants, nor have a mechanistic understanding of how these genes, individually and together, contribute to a phenotype. Thus, we need to consider how such genomic studies can improve our knowledge of trait physiology. One approach would be to focus study genetic analyses by organs and tissues of interest.

Pritchard and colleagues have hypothesized that the majority of genome-wide association study (GWAS) signals may be functionally spurious and arise from genes peripheral to the core functions affected in a trait or disease [1]. These false positives dominate because most genes in a cell-type are connected by gene expression to one another through very shallow functional networks, a working hypothesis that fails to explain the stability of network perturbations (robustness) or their specificity (phenotypic effects) [2–4]. To resolve this question, connecting genotypes to phenotypes through gene expression variation is of primary importance since eQTLs (expression quantitative trait loci) are identifiable causal factors [5,6]. However, utilizing gene expression in trait-related tissues is necessary [7], as genes exert their activities in the context of a core genetic network with intrinsic (cell autonomous) and extrinsic (non-autonomous) feedback [8].

Transcription within mammalian genomes is locally regulated within ~400 kilobase (kb) chromatin segments called topological associating domains (TADs), largely invariant across cell types [9]. TADs contain numerous dispersed spatiotemporal expression cis-regulatory elements (CREs or enhancers) that are brought together by DNA looping to allow binding of various transcription factors (TF) to enable gene expression control [10]. Many enhancers are recognized by their DNaseI hyper-sensitivity (DHS), ATAC-seq (Assay for Transposase-Accessible Chromatin using sequencing) assays [11], or adjacent histone (H3K4me1, H3K4me3, H3K27ac) modifications [12,13]. Their phenotypic importance is evident from the fact that only 2.6% of the genome comprises DHS and histone marks [14] but explains ~30% of the heritability of traits [15]. Thus, trait variation is from sequence changes within TFs, their binding sites (TFBS) and CREs, all detectable through epigenomic marks in cell lines and tissues. In this study, we

propose an approach wherein these types of epigenomic data are used to identify genes within a GWAS locus in tissues of interest.

The analyses we propose are enabled by numerous public genomic resources. The Encyclopedia of DNA Elements (ENCODE) Project (<https://www.encodeproject.org/>) has generated open chromatin, RNA and DNA sequencing, genotyping, and histone modification data, among other data types. The Genotype-Tissue Expression (GTEx) Project (<https://www.gtexportal.org/>) includes genotype and expression data across 53 tissues and is useful as a reference transcriptome and eQTL dataset. These public resources also enable the development of an annotation score, deltaSVM [16], in which the quantitative impact of a non-coding variant on tissue or cell type specific gene regulation is predicted, based on a reference training set of regulatory regions. In this study, we exemplify this reverse genetic approach by focusing on blood pressure (BP) and QT interval variation.

Although the roles of the kidney and adrenal gland are well established in blood pressure regulation and syndromes [17–19], our previous work in the Kaiser Permanente Research Program on Genes, Environment and Health (RPGEH) Genetic Epidemiology Research on Adult Health and Aging (GERA) [20,21] study demonstrated that associated variants at BP GWAS loci were enriched in eQTLs specific to the aorta and tibial arteries. Expanding on this work in this study, we aimed to connect groups of proximal putative regulatory variants within and around each gene to both the gene's expression and also to BP traits, inferring that the gene's expression in a potentially relevant tissue affected the regulation of BP. To accomplish this, within each artery dataset, we identified putative CREs, and by extension, putative CRE variants, for every gene, and tested these variants in aggregate for association with BP in the GERA study, as well as with expression in the GTEx study. We used the sequence kernel association test (SKAT) [22]

for these association analyses, with each variant weighted by their deltaSVM score, to up-weight variants with greater predicted effects on gene regulatory activity. We supplemented our expression analyses with the software MetaXcan [23] to test whether the predicted expression of genes in each individual could be associated with BP. Prior to the novel BP gene discovery analyses of tissue involvement, and as a positive control, we first examined genes for the cardiac trait QT interval for which there is strong functional evidence of primarily heart involvement, using data from the Atherosclerosis Risk in Communities (ARIC) [24,25] study. Finally, we examined the effects of putative regulatory variation for monogenic BP syndrome genes, all known to be renal or adrenal disorders, in four available kidney cell types to test for a group effect on BP.

Our results demonstrate the feasibility of identifying BP genes by tissue, which we expect will facilitate more comprehensive functional analyses of BP genes and BP control mechanisms.

Results

We conducted several tissue-specific analyses to identify tissues and genes of interest for BP regulation using the GERA study. We initially focused on identifying tissues relevant to BP GWAS loci, and subsequently expanded on this by using tissue-specific information to analyze putative CRE variation of genes in these tissues. The aim was to identify specific genes and variants of interest at these GWAS loci. We also studied putative regulatory variation at 20 monogenic syndromic hypertension and hypotension genes in several kidney cell types. To begin, our study also includes an analysis of QT interval as a positive control to demonstrate the identification of well-characterized genes for that trait.

Partitioned heritability of BP

We examined heritability for SBP and DBP using 80,792 GERA EUR subjects with stratified LD score regression (LDSC) across several functional categories [15], to identify functional categories in which BP heritability was enriched. We found that the top-ranked enriched categories were enhancer-associated histone marks H3K27ac, H3K4me1, and the Hnisz “super-enhancer” category (Tables S1-S2). This is in accordance with a previous study in which BP heritability was determined to be mostly from within DNaseI hypersensitivity sites [26], and, taken together with the results of the eQTL enrichment analyses, supports the study of regulatory elements in specific tissues of interest for BP.

Constructing CRE maps

With knowledge of tissues highly relevant to characterizing BP GWAS loci, our next aim was to test each gene’s putative cis-regulatory variation for association with both gene expression and BP, in a tissue-specific context. This is expected to assist in identifying novel genes of interest, as well as provide tissue- or cell-type-specific information about known genes.

We first constructed CRE maps for the aorta and tibial arteries, as well as four kidney cell types (renal cortical epithelial cell, glomerular endothelial cell, epithelial cell of proximal tubule, and glomerular visceral epithelial cell), because of the known involvement of the kidney in blood pressure regulation [17,18], using ENCODE data (Table 1) (though many monogenic forms of blood pressure disorders occur due to an effect of the adrenal gland on renal function [19]). These CRE maps were completed as an extension of the construction of our recent cardiac CRE map [27]. We specifically focused on identifying putative enhancers for the aorta and tibial arteries (see Methods). We subsequently used these maps for training with the software gkm-

SVM [28,29] in order to generate deltaSVM functional scores for all non-coding variants from the 1000 Genomes European ancestry sample, to be tested for association on a gene-level basis. The performance for each model is available in Table S3 (AUC range: 0.84-0.96), with the best performance in the renal cell types. A possible reason for the improved performance of the renal cell types is that the data were from individual cell types as opposed to a mixture of cell types comprising the arteries. The magnitude of the deltaSVM score for a variant reflects its predicted impact on regulatory functional activity, while its sign reflects the prediction with respect to the reference allele. Therefore, to represent the predicted impact of each variant irrespective of allele, we show the distributions of the absolute values of the deltaSVM scores for the arteries and kidney cell types in Fig 1.

Table 1. Summary of datasets for analysis in this study.

Study/Project	Dataset	Description	Analysis
ENCODE Project			
Artery	Open Chromatin		
	ENCSR000EIH	aortic smooth muscle cell	Aorta SKAT analyses groupings
	ENCSR000EMC	aortic adventitia cell	Aorta SKAT analyses groupings
	ENCSR000EOG	pulmonary artery endothelial cell	Aorta SKAT analyses groupings
	ENCSR000EOH	pulmonary artery fibroblast	Aorta SKAT analyses groupings
	ENCSR630REB	tibial artery	Aorta, Tibial Artery SKAT analyses groupings
	Histone Modification		
	ENCSR519CFV	aorta H3K27Ac	Aorta SKAT analyses groupings
	ENCSR015GFK	aorta H3K27Ac	Aorta SKAT analyses groupings
	ENCSR318HUC	aorta H3K27Ac	Aorta SKAT analyses groupings

	ENCSR069UMW	aorta H3K27Ac	Aorta SKAT analyses groupings
	ENCSR322TJD	aorta H3K27Ac	Aorta SKAT analyses groupings
	ENCSR233LCT	tibial artery H3K4me1	Tibial Artery SKAT analyses groupings
Kidney	Open Chromatin		
	ENCSR000EOK	renal cortical epithelial cell	Kidney SKAT analyses groupings
	ENCSR000EOM	glomerular endothelial cell	Kidney SKAT analyses groupings
	ENCSR000EPW	proximal tubule epithelial cell	Kidney SKAT analyses groupings
	ENCSR785BDQ	glomerular visceral epithelial cell (3yo)	Kidney SKAT analyses groupings
GERA	Genotypes, BP phenotype	71404 European-Ancestry individuals	SKAT, BP traits
	Summary statistics	80,792 European-Ancestry individuals	Partitioned heritability analyses, MetaXcan
ARIC	Genotypes, QT interval	9,083 European-Ancestry individuals	SKAT, QT interval
	Summary statistics	9,083 European-Ancestry individuals	Partitioned heritability analyses, MetaXcan
GTE _x	Genotypes, Expression		
	Aorta	197 samples	SKAT, expression (for BP)
	Tibial Artery	285 samples	SKAT, expression (for BP)
	Heart Atrial Appendage	159 samples	SKAT, expression (for QT)
	Heart Left Ventricle	190 samples	SKAT, expression (for QT)

Fig 1. Distributions of deltaSVM scores for each tissue or cell type.

Tissue-specific gene identification

As our emphasis in this section is to connect a gene's putative CRE variants to both a phenotype of interest and to its expression in relevant tissues (Fig 2), we first describe the overall analysis scheme as applied to a general phenotype of interest. We then describe how we applied

these analyses, first to the QT interval in the ARIC study, as proof of principle to demonstrate the utility of these analyses, and then to our BP traits of interest in the GERA study.

Fig 2. Overview of CRE SKAT analysis.

We defined a gene's "cis"-regulatory variants in this analysis as those variants falling in putative CREs within 50Kb of the gene's start and end. We tested their aggregate effect for each gene using SKAT [22], for association tests with the phenotype(s) of interest in the relevant population, SBP and DBP in the GERA study, and QT interval in the ARIC study. SKAT is a test that has generally been used to study groups of variants together and is useful when variants can have bidirectional effects; rare variants are more highly weighted than common variants by default. In addition to the default weights, we ran the analysis using equal weights for all variants. We finally used the tissue- or cell-type-specific deltaSVM scores for the analyzed variants as weights for a customized SKAT test; the score scaling with the effect of the variant on functional regulatory activity.

We then tested these groupings with expression data from GTEx v6p in the tissues of interest to link variants in the genes of interest to their gene expression. The groupings tested in the GTEx data with expression were not always identical to the groupings tested in the GERA or ARIC studies because of differences in imputation quality score filtering, missingness of genotypes from genotype probabilities to hard call conversion, and variants present in the reference populations studied. However, this analysis still connects a given gene to its expression and to the phenotype via a highly overlapping set of CRE variants, and was completed this way to test the most complete set of variants available meeting our criteria. In addition to testing

putative regulatory variants with gene expression in GTEx, we used the recently developed MetaXcan [23] software to augment SKAT to identify any new associations by this method. This software predicts the association of gene expression with a phenotype, given genotypes for the population of interest based on training from reference genotypes and expression data.

Analysis of CREs in QT interval

As mentioned earlier, we considered the cardiac trait QT interval first to demonstrate proof of concept for tissue-specific gene identification. The QT interval is the time in ms between the onset of the Q wave and the end of the T wave in the surface 12-lead electrocardiogram [30], which has ~30% heritability [31–34]. In our recent work, we have demonstrated that a significant proportion of the heritability is explained by predicted cardiac regulatory variants [27]. We analyzed the genes at previously published QT interval GWAS loci to determine whether or not a heart-specific effect could be observed. Two of the genes with major effects in a GWAS and functionally validated in QT interval heritability are *NOS1AP* [34–36] and *SCN5A* [36,37]. The full results are presented in the Text S1 results, Table S4, and Figs S1 and S2; to summarize here briefly, we aimed to discover if a heart-specific effect could be revealed for each of these two genes. We observed a heart-specificity for *SCN5A*; *NOS1AP* showed signal across all the cell types in the equal-weighted analyses, though considerably attenuated in some of the deltaSVM-weighted non-heart tissues. Considering both sets of effects, certainly variants with detectable signals present in open chromatin regions specific to the relevant tissue/cell types will allow the detection of a tissue-specific signal, as for *SCN5A*. It also appears, however, that gene-level signals may be captured by analyses in which all variants are weighted equally, and when local open chromatin boundaries across tissues/cell types overlap

considerably, especially when variants with strong signals are present within these shared regions. In this situation, we will not necessarily be able to differentiate between different tissue/cell types. Weighting with the tissue-specific deltaSVM scores introduces an additional tier of tissue specificity and is based on global open chromatin differences, and is also not expected to be impacted by linkage disequilibrium (LD) in the ways that the other two weighting schemes are, as the generation of the scores are only dependent on sequence context. Finally, using the default weights shows least concordance with the other two sets of results, indicating that for this analysis, rare variants are not driving the signal as compared to common variants. This is as expected, as we prioritized non-coding variation for these analyses, and the rare variants with larger effects expected to make a detectable contribution are more likely to be in the exome.

Analysis of CREs at GWAS loci for BP regulation

We then applied these analyses to the tissues of interest for BP regulation, namely aorta, tibial artery, and four kidney cell types, in a subset of 71,404 unrelated GERA EUR individuals. We tested 14,548 genes expressed at $\text{RPKM} \geq 0.3$ in 197 aorta GTEx samples and 13,963 genes expressed at $\text{RPKM} \geq 0.3$ in 285 tibial artery GTEx samples for the SKAT analyses. We used summary statistics available from 80,792 individuals[38] to maximize the sample size for which the MetaXcan analyses were run, for the aorta and tibial arteries. Results for each of the arteries are presented in Tables 2-5. In some cases, shared variants drive the positive signal for multiple genes at the same locus; expression in the relevant tissue or cell type may pinpoint a specific gene. However, it may be noted that the genes *CERS5*, *COX14*, and *RP4-605O3.4* are all present at the same locus in the arteries (Tables 2-5), but evidence of expression association is present

256 for many of these genes; this may be indicative of proximal variants affecting different genes, or
257 pleiotropy of single variants affecting expression of multiple genes.
258

259 **Table 2. Aorta SBP SKAT and MetaXcan results.**

Gene	N.sbp p	p.sbp.dsvm	p.sbp.eq	p.sbp.def	N.GTex	p.GTex. dsvm	p.GTex. eq	p.GTex. def	N.MetX	p.MetX	previous
<i>NR3C1</i>	8	5.44×10^{-6}	7.19×10^{-6}	0.064	7	0.002	0.001	0.529	-	-	FALSE
<i>WBP1L</i>	39	6.66×10^{-6}	5.91×10^{-7}	0.085	37	2.72×10^{-4}	4.76×10^{-4}	0.137	21	1.35×10^{-5}	FALSE
<i>SBF2</i>	34	9.27×10^{-6}	7.60×10^{-6}	2.58×10^{-5}	34	0.003	4.71×10^{-4}	0.022	17	2.01×10^{-5}	TRUE
<i>CLCN6</i>	28	1.07×10^{-5}	1.73×10^{-8}	2.32×10^{-6}	34	0.014	0.003	2.63×10^{-4}	19	0.028	TRUE
<i>MTHFR</i>	29	1.08×10^{-5}	3.35×10^{-8}	4.22×10^{-6}	35	1.85×10^{-5}	3.15×10^{-5}	0.015	2	0.001	TRUE
<i>C10orf32</i>	17	1.68×10^{-5}	5.41×10^{-7}	0.026	15	4.63×10^{-15}	7.10×10^{-13}	0.059	-	-	TRUE
<i>RP4-605O3.4</i>	7	4.51×10^{-5}	3.47×10^{-4}	0.069	5	0.016	5.11×10^{-4}	0.425	-	-	FALSE
<i>COX14</i>	6	4.51×10^{-5}	1.47×10^{-4}	0.069	4	7.64×10^{-4}	5.97×10^{-5}	0.338	5	0.034	TRUE
<i>CSK</i>	3	6.34×10^{-5}	5.65×10^{-5}	0.110	3	3.18×10^{-5}	2.49×10^{-5}	0.579	17	1.24×10^{-4}	TRUE
<i>ULK3</i>	2	6.34×10^{-5}	6.42×10^{-5}	0.952	3	0.001	7.02×10^{-4}	0.164	9	9.14×10^{-4}	TRUE

260 N.sbp, number of variants analyzed in SKAT analysis of SBP in GERA; p.sbp.*, p-values from SKAT analysis of BP in GERA with
 261 deltasvm (dsvm), eq (equal), or default (def) weights; N.GTex, number of variants analyzed in SKAT analysis of expression in GTex
 262 p.GTex.*, p-values from SKAT analysis of expression in GTex with deltasvm (dsvm), eq (equal), or default (def) weights; N.MetX,
 263 number of variants used in MetaXcan prediction results; p.MetX, MetaXcan p-value; previous, TRUE if found at a previously
 264 identified GWAS locus from the UKB list, FALSE otherwise

265

266 **Table 3. Aorta DBP SKAT and MetaXcan results.**

Gene	N.db p	p.dbp.ds vm	p.dbp.eq	p.dbp.def	N.GTE x	p.GTEEx.d svm	p.GTEEx.eq	p.GTEEx.d ef	N.Me tX	p.Met X	previous
<i>NOV</i>	10	1.40×10^{-8}	1.98×10^{-8}	0.664	7	4.26×10^{-5}	4.04×10^{-5}	1	99	0.102	TRUE
<i>ULK4</i>	5	1.75×10^{-8}	5.04×10^{-8}	0.001	5	2.49×10^{-22}	2.95×10^{-22}	0.094	42	2.95×10^{-10}	TRUE
<i>COX14</i>	6	1.35×10^{-7}	7.49×10^{-7}	0.138	4	7.64×10^{-4}	5.97×10^{-5}	0.338	5	0.003	TRUE
<i>IGFBP3</i>	4	3.41×10^{-7}	7.67×10^{-8}	0.006	3	0.011	0.014	0.842	-	-	FALSE
<i>SDCCAG8</i>	10	6.92×10^{-7}	1.74×10^{-6}	0.359	10	5.57×10^{-6}	2.26×10^{-8}	0.481	9	2.95×10^{-9}	TRUE
<i>CEP170</i>	6	7.06×10^{-7}	1.49×10^{-6}	8.44×10^{-7}	7	2.52×10^{-4}	1.72×10^{-4}	3.59×10^{-4}	19	0.002	TRUE
<i>CSK</i>	3	9.93×10^{-7}	8.91×10^{-7}	0.068	3	3.18×10^{-5}	2.49×10^{-5}	0.579	17	7.93×10^{-5}	TRUE
<i>ULK3</i>	2	9.93×10^{-7}	9.76×10^{-7}	0.826	3	0.001	7.02×10^{-4}	0.164	9	2.09×10^{-4}	TRUE
<i>SCAMP5</i>	15	3.98×10^{-6}	1.23×10^{-5}	0.009	13	6.70×10^{-8}	2.97×10^{-11}	4.15×10^{-10}	14	1.93×10^{-5}	FALSE
<i>RPP25</i>	3	4.67×10^{-6}	4.76×10^{-6}	0.928	2	3.58×10^{-17}	3.93×10^{-17}	4.79×10^{-9}	-	-	FALSE ^a
<i>HDGFRP3</i>	3	1.52×10^{-5}	1.21×10^{-5}	0.578	3	3.90×10^{-7}	1.35×10^{-6}	0.042	48	0.001	TRUE
<i>COX4I2</i>	31	2.12×10^{-5}	1.36×10^{-5}	0.003	7	0.002	8.86×10^{-4}	0.729	-	-	FALSE
<i>SBF2</i>	34	2.27×10^{-5}	1.37×10^{-5}	0.004	34	0.003	4.71×10^{-4}	0.022	17	7.41×10^{-4}	TRUE
<i>RNF40</i>	6	2.39 x	3.72 x	0.242	5	$2.74 \times 10^{-}$	2.50×10^{-4}	0.338	-	-	TRUE

		10^{-5}	10^{-5}			10^{-4}					
<i>RP11-382A20.2</i>	2	4.82×10^{-5}	4.94×10^{-5}	0.627	2	0.005	0.004	0.066	-	-	FALSE
<i>RNASEH2C</i>	1	7.50×10^{-5}	7.50×10^{-5}	7.50×10^{-5}	1	0.002	0.002	0.002	11	0.690	TRUE
<i>SLC25A37</i>	31	1.16×10^{-4}	1.80×10^{-4}	0.121	28	4.01×10^{-4}	2.07×10^{-4}	0.537	13	0.169	FALSE
<i>SENP2</i>	8	1.36×10^{-4}	1.09×10^{-4}	0.711	7	0.004	0.002	0.202	43	0.050	TRUE
<i>VPS37B</i>	11	1.46×10^{-4}	2.82×10^{-5}	7.73×10^{-5}	10	8.50×10^{-6}	8.90×10^{-7}	2.03×10^{-4}	16	3.43×10^{-5}	FALSE
<i>ZNF652</i>	1	1.69×10^{-4}	1.69×10^{-4}	1.69×10^{-4}	1	8.73×10^{-4}	8.73×10^{-4}	8.73×10^{-4}	7	1.53×10^{-5}	TRUE
<i>NR3C1</i>	8	2.21×10^{-4}	2.14×10^{-4}	0.018	7	0.002	0.001	0.529	-	-	FALSE
<i>PPCDC</i>	14	2.25×10^{-4}	6.05×10^{-5}	0.009	12	6.68×10^{-8}	4.20×10^{-8}	4.82×10^{-8}	38	0.002	FALSE

N.dbp, number of variants analyzed in SKAT analysis of DBP in GERA; p.dbp.*, p-values from SKAT analysis of DBP in GERA

with deltasvm (dsvm), eq (equal), or default (def) weights; N.GTEx, number of variants analyzed in SKAT analysis of expression in

GTEx; p.GTEx.*, p-values from SKAT analysis of expression in GTEx with deltasvm (dsvm), eq (equal), or default (def) weights;

N.MetX, number of variants used in MetaXcan prediction results; p.MetX, MetaXcan p-value; previous, TRUE if found at a

previously identified GWAS locus from the UKB list, FALSE otherwise

^a *RPP25* has been previously identified in GWAS, but our method to identify genes at GWAS loci was conservative and missed this

gene; see text

275 **Table 4. Tibial Artery SBP SKAT and MetaXcan results.**

Gene	N.sbp	p.sbp.dsv m	p.sbp.eq	p.sbp.def	N.GTE x	p.GTEx.ds vm	p.GTEx.e q	p.GTEx.d ef	N.Me tX	p.Met X	previous
<i>CLCN6</i>	24	9.65×10^{-9}	6.54×10^{-8}	2.62×10^{-4}	28	5.07×10^{-8}	6.14×10^{-8}	3.58×10^{-6}	10	2.67×10^{-9}	TRUE
<i>MTHFR</i>	24	9.65×10^{-9}	6.54×10^{-8}	2.62×10^{-4}	28	8.79×10^{-8}	1.13×10^{-5}	0.161	35	0.076	TRUE
<i>C10orf32</i>	8	6.98×10^{-8}	2.08×10^{-8}	0.004	7	3.07×10^{-14}	2.05×10^{-18}	7.30×10^{-8}	-	-	TRUE
<i>HOXC- AS1</i>	11	4.00×10^{-5}	1.42×10^{-4}	0.638	14	3.24×10^{-6}	6.04×10^{-9}	0.194	-	-	FALSE
<i>CCDC6</i>	37	4.49×10^{-5}	7.66×10^{-5}	0.151	37	3.32×10^{-4}	3.81×10^{-4}	0.035	30	1.08×10^{-4}	FALSE
<i>ATE1</i>	15	5.67×10^{-5}	1.06×10^{-4}	0.583	17	8.74×10^{-4}	2.29×10^{-4}	0.086	71	0.816	FALSE
<i>SOX7</i>	20	6.76×10^{-5}	6.42×10^{-5}	0.036	18	0.007	0.016	0.121	-	-	FALSE
<i>AGT</i>	14	8.29×10^{-5}	4.69×10^{-5}	0.090	14	0.001	5.87×10^{-7}	6.97×10^{-6}	24	0.937	TRUE
<i>NT5C2</i>	25	8.75×10^{-5}	8.58×10^{-4}	1.44×10^{-4}	22	0.004	0.008	0.163	7	0.006	TRUE
<i>DHX33</i>	28	9.67×10^{-5}	1.34×10^{-4}	0.372	29	3.67×10^{-7}	5.69×10^{-8}	0.033	7	0.277	FALSE
<i>SFMBT1</i>	8	1.15×10^{-4}	7.26×10^{-4}	0.521	8	0.003	0.023	0.363	13	0.006	FALSE
<i>NPPA</i>	15	1.16×10^{-4}	2.86×10^{-7}	0.003	16	0.007	0.012	0.025	-	-	TRUE
<i>ERII</i>	29	1.33×10^{-4}	2.37×10^{-4}	0.280	27	0.001	0.002	0.185	-	-	FALSE
<i>BCL2L2</i>	6	1.59×10^{-4}	2.67×10^{-4}	0.009	5	5.20×10^{-4}	0.006	0.356	-	-	FALSE
<i>BCL2L2-</i>	6	1.59×10^{-4}	2.67×10^{-4}	0.009	6	0.013	0.104	0.717	-	-	FALSE

<i>PABPN1</i>			10 ⁻								
<i>NPPA-ASI</i>	17	1.66 x 10 ⁻⁴	2.92 x 10 ⁻⁷	0.003	19	1.89 x 10 ⁻²³	5.62 x 10 ⁻²⁴	4.67 x 10 ⁻¹⁰	-	-	FALSE
<i>Clorf132</i>	26	1.74 x 10 ⁻⁴	5.09 x 10 ⁻⁶	0.055	26	0.009	1.30 x 10 ⁻⁴	0.400	-	-	FALSE
<i>RPAIN</i>	12	1.89 x 10 ⁻⁴	2.89 x 10 ⁻⁴	0.627	12	0.003	7.26 x 10 ⁻⁴	0.479	5	0.546	FALSE
<i>CTC-524C5.2</i>	12	1.89 x 10 ⁻⁴	2.89 x 10 ⁻⁴	0.627	12	0.0001628	2.97 x 10 ⁻⁵	0.301	-	-	FALSE

N.sbp, number of variants analyzed in SKAT analysis of SBP in GERA; p.sbp.*, p-values from SKAT analysis of SBP in GERA with deltasvm (dsvm), eq (equal), or default (def) weights; N.GTEx, number of variants analyzed in SKAT analysis of expression in GTEx; p.GTEx.*, p-values from SKAT analysis of expression in GTEx with deltasvm (dsvm), eq (equal), or default (def) weights; N.MetX, number of variants used in MetaXcan prediction results; p.MetX, MetaXcan p-value; previous, TRUE if found at a previously identified GWAS locus from the UKB list, FALSE otherwise

Table 5. Tibial Artery DBP SKAT and MetaXcan results.

Gene	N.d bp	p.dbp.dsvm	p.dbp.eq	p.dbp.def	N.GT Ex	p.GTEx.dsvm	p.GTEx.eq	p.GTEx.def	N.Met X	p.Met X	previous
<i>NOV</i>	17	1.70 x 10 ⁻⁸	1.15 x 10 ⁻⁸	0.731	16	7.00 x 10 ⁻⁶	7.95 x 10 ⁻⁶	0.943	34	1.44 x 10 ⁻⁶	TRUE
<i>CERS5</i>	6	1.66 x 10 ⁻⁶	3.75 x 10 ⁻⁶	0.004	2	3.41 x 10 ⁻⁴	3.90 x 10 ⁻⁴	3.55 x 10 ⁻⁴	33	1.17 x 10 ⁻⁴	TRUE
<i>COX14</i>	6	1.66 x 10 ⁻⁶	3.75 x 10 ⁻⁶	0.004	2	4.26 x 10 ⁻⁵	1.62 x 10 ⁻⁵	6.54 x 10 ⁻⁴	33	0.004	TRUE
<i>RP4-605O3.4</i>	6	1.66 x 10 ⁻⁶	3.75 x 10 ⁻⁶	0.004	2	1.57 x 10 ⁻⁷	1.76 x 10 ⁻⁷	2.36 x 10 ⁻⁷	-	-	FALSE

<i>JAG1</i>	37	4.59×10^{-6}	6.04×10^{-6}	0.032	41	0.018	0.003	0.304	-	-	TRUE
<i>ULK4</i>	12	5.69×10^{-6}	2.60×10^{-10}	0.008	11	2.78×10^{-11}	3.61×10^{-19}	0.158	50	5.40×10^{-12}	TRUE
<i>IPO9</i>	18	6.08×10^{-6}	1.65×10^{-4}	0.416	17	0.009	5.04×10^{-4}	8.59×10^{-4}	33	0.952	FALSE
<i>LIMA1</i>	9	1.86×10^{-5}	3.60×10^{-5}	0.638	6	0.008	0.003	0.926	6	0.120	TRUE
<i>NAVI</i>	44	3.83×10^{-5}	7.79×10^{-4}	0.554	36	2.64×10^{-4}	0.009	0.010	-	-	FALSE
<i>COX4I2</i>	13	3.92×10^{-5}	1.88×10^{-5}	0.360	4	1.94×10^{-4}	1.82×10^{-4}	0.656	-	-	FALSE
<i>UBN1</i>	11	4.23×10^{-5}	1.10×10^{-5}	0.417	10	3.19×10^{-7}	1.52×10^{-7}	0.299	9	1.51×10^{-5}	TRUE
<i>SCAMP5</i>	7	5.29×10^{-5}	7.75×10^{-5}	0.331	5	4.47×10^{-11}	1.36×10^{-7}	5.09×10^{-5}	35	8.04×10^{-8}	FALSE
<i>RNASEH2C</i>	3	5.88×10^{-5}	5.65×10^{-5}	0.056	1	0.006	0.006	0.006	10	0.831	TRUE
<i>CEP120</i>	8	6.36×10^{-5}	6.87×10^{-7}	0.003	8	4.43×10^{-5}	3.42×10^{-5}	0.123	12	1.63×10^{-5}	TRUE
<i>CLCN6</i>	24	7.30×10^{-5}	3.87×10^{-5}	0.001	28	5.07×10^{-8}	6.14×10^{-8}	3.58×10^{-6}	10	1.16×10^{-5}	TRUE
<i>MTHFR</i>	24	7.30×10^{-5}	3.87×10^{-5}	0.001	28	8.79×10^{-8}	1.13×10^{-5}	0.161	35	0.086	TRUE
<i>SDCCAG8</i>	6	9.04×10^{-5}	5.15×10^{-5}	0.130	7	4.55×10^{-5}	1.73×10^{-5}	0.273	17	8.08×10^{-8}	TRUE
<i>ACSF3</i>	9	1.05×10^{-4}	2.68×10^{-4}	5.00×10^{-4}	11	0.003	8.21×10^{-5}	0.548	56	0.328	FALSE
<i>RPP25</i>	4	1.11×10^{-4}	9.49×10^{-5}	0.559	2	4.36×10^{-12}	2.56×10^{-12}	0.001	-	-	FALSE ^a
<i>COX5A</i>	4	1.11×10^{-4}	9.49×10^{-5}	0.559	2	0.004	0.004	0.053	-	-	TRUE
<i>MKL2</i>	31	$1.12 \times$	$1.34 \times$	0.031	25	0.017	0.010	0.020	-	-	FALSE

		10^{-4}	10^{-4}								
<i>VPS37B</i>	11	1.40×10^{-4}	0.002	5.38×10^{-4}	9	2.08×10^{-15}	1.13×10^{-14}	1.12×10^{-14}	37	0.007	FALSE
<i>FAM20B</i>	13	1.91×10^{-4}	0.004	0.690	10	3.35×10^{-5}	1.17×10^{-13}	0.337	17	0.030	FALSE
<i>PLA2G4B</i>	10	2.10×10^{-4}	2.18×10^{-4}	1	10	3.41×10^{-7}	4.20×10^{-7}	0.320	16	0.005	FALSE
<i>ALDH2</i>	3	2.15×10^{-4}	9.05×10^{-5}	0.658	2	5.44×10^{-8}	3.91×10^{-8}	1.26×10^{-6}	8	0.002	FALSE
<i>DUSP15</i>	15	2.39×10^{-4}	1.57×10^{-4}	0.008	4	0.018	0.038	0.923	5	0.028	FALSE
<i>RP11-65J21.3</i>	8	2.40×10^{-4}	1.88×10^{-4}	0.660	5	0.015	0.020	0.382	-	-	FALSE
<i>MAPKBPI</i>	8	2.52×10^{-4}	2.34×10^{-4}	0.077	6	7.04×10^{-5}	6.34×10^{-5}	0.056	23	0.019	FALSE
<i>JMJD7</i>	7	2.54×10^{-4}	2.38×10^{-4}	0.077	6	1.47×10^{-9}	4.30×10^{-8}	0.203	20	3.34×10^{-4}	FALSE
<i>CENPW</i>	3	2.79×10^{-4}	2.43×10^{-4}	0.075	2	0.004	0.004	0.066	21	0.002	FALSE
<i>ATF1</i>	6	2.81×10^{-4}	3.36×10^{-4}	0.639	2	4.80×10^{-19}	5.97×10^{-19}	9.06×10^{-5}	14	2.13×10^{-5}	FALSE

283 N.dbp, number of variants analyzed in SKAT analysis of DBP in GERA; p.dbp.*, p-values from SKAT analysis of DBP in GERA
 284 with deltasvm (dsvm), eq (equal), or default (def) weights; N.GTEx, number of variants analyzed in SKAT analysis of expression in
 285 GTEx; p.GTEx.*, p-values from SKAT analysis of expression in GTEx with deltasvm (dsvm), eq (equal), or default (def) weights;
 286 N.MetX, number of variants used in MetaXcan prediction results; p.MetX, MetaXcan p-value; previous, TRUE if found at a
 287 previously identified GWAS locus from the UKB list, FALSE otherwise

288 ^a *RPP25* has been previously identified in GWAS, but our method to identify genes at GWAS loci was conservative and missed this
289 gene; see text

290

291

On the whole, the 25 genes reported here across aorta and tibial artery genes have been identified at previous BP loci [39]. While there are several genes in each analysis with interesting associations with BP traits, here we only highlight the genes that have statistical significance of $p < 1 \times 10^{-4}$ for both expression and BP in the aorta analyses (Tables 2-3). In our previous work, the aorta was demonstrated as the greatest outlier in an analysis of eQTL enrichment among GTEx tissues for BP traits [38]. These genes include: *MTHFR*[40,41] (SBP), *C10orf32*[40] (SBP), *CSK* (SBP), *NOV*[42] (DBP), *ULK4*[43] (DBP), *SDCCAG8* (DBP), *SCAMP5* (DBP), *RPP25* (DBP), *HDGFRP3* (DBP), *VPS37B* (DBP), and *PPCDC* (DBP). Most of these genes are present at or near previously replicated BP GWAS loci; *SDCCAG8* was identified as part of Hoffmann et al.[38] It is noteworthy that both *SCAMP5* and *PPCDC* are neighboring genes, but have independent expression support in the same tissue.

Analysis of CREs at GWAS loci for monogenic BP genes

We also studied the genes involved in monogenic forms of hypotension or hypertension in four kidney cell types available from the ENCODE project (see earlier). As the expression data available for kidney are insufficient, we studied each cell type individually and carried out only SKAT analyses for these genes; the results are in Tables 6-7. The most notable result is that of *CYP17A1*, which shows an effect ($p \sim 10^{-5} - 10^{-7}$) across all four cell types in the unweighted variants analyses for SBP only, and more specifically, only in the glomerular endothelial cell (ENCSR000EOM) ($p_{\text{SKAT.dsvm.ENCSR000EOM}} = 3.64 \times 10^{-8}$) in the deltaSVM-weighted results. However, as *C10orf32* is a gene of interest at the same locus, based on the artery results above, we examined and noted that the results are somewhat similar for this gene, although not

as striking, due to variant set sharing in the SKAT analyses for these genes (deltaSVM p-values: ENCSR000EOK, 8 variants, $p=1.87 \times 10^{-3}$; ENCSR000EOM, 12 variants, $p=2.87 \times 10^{-5}$; ENCSR000EPW, 10 variants, $p=6.33 \times 10^{-4}$; ENCSR786BDQ, 9 variants, $p=0.031$). The breakdown of individual variants analyzed for these two genes is in Table S5. The variant rs3824754, with an SBP association $p=1.40 \times 10^{-11}$, appears in the groupings of both genes for all four cell types, but has the highest deltaSVM magnitude in the endothelial cell. Additionally, there is a set of four variants with SBP association ($p < 1 \times 10^{-4}$; rs284853, rs284854, rs284855, rs284856) which only appear in the ENCSR000EOM groupings. We observed that while *CYP17A1* was similarly associated with, or demonstrated evidence of association with, SBP in the deltaSVM and unweighted variants analysis (aorta deltaSVM $p=2.06 \times 10^{-5}$, 34 variants; tibial artery deltaSVM $p=1.40 \times 10^{-8}$, 15 variants), the analysis of variants in GTEx (33 variants for aorta and 14 variants for tibial artery) did not reflect any significant association ($p > 0.01$). In contrast, *C10orf32* demonstrated association with SBP (aorta deltaSVM $p=1.68 \times 10^{-5}$, 17 variants; tibial artery deltaSVM $p=6.98 \times 10^{-8}$, 8 variants, Tables 2 and 4) and with expression in GTEx (aorta deltaSVM $p=4.63 \times 10^{-15}$, 15 variants; tibial artery deltaSVM $p=3.07 \times 10^{-14}$, 7 variants, Tables 2 and 4). The same four variants unique to the ENCSR000EOM groupings above with strong associations with SBP are also present in the artery groupings. Three of these variants (rs284854, rs284855, rs284856) are eQTLs for *C10orf32* in the aorta and tibial arteries; these variants, however, do not show association with *CYP17A1* expression in these tissues (all $p > 0.03$ for aorta, all $p > 0.21$ for tibial artery, from eQTL data available from the GTEx portal (<https://www.gtportal.org/>), accessed 09/08/17). Additionally, as the *CYP17A1* gene primarily demonstrates an adrenal effect in the monogenic disorder [44], we also examined the associations of these three variants in the GTEx portal with adrenal gland expression data for

both genes; all have $p > 0.26$ for *CYP17A1* and $p > 0.04$ for *C10orf32*. This may reflect an endothelial-cell-specific effect for *C10orf32* rather than a tissue-type effect, especially as this locus has been identified in several previous BP GWAS studies, [40,43,45–47]; it may also not be very informative for the kidney, though suitable expression data for kidney would be required to assess this.

Table 6. Kidney SBP SKAT results.

Gene	Experiment ^a	N.sbp	p.sbp.dsvm	p.sbp.eq	p.sbp.def
<i>BSND</i>	ENCSR000EOK	13	0.836	0.796	0.611
<i>CASR</i>	ENCSR000EOK	21	0.628	0.794	0.605
<i>CLCNKA</i>	ENCSR000EOK	8	0.444	0.502	0.426
<i>CLCNKB</i>	ENCSR000EOK	9	0.475	0.382	0.454
<i>CUL3</i>	ENCSR000EOK	18	0.318	0.388	0.049
<i>CYP11B1</i>	ENCSR000EOK	12	0.020	0.016	0.001
<i>CYP11B2</i>	ENCSR000EOK	10	0.014	0.014	0.003
<i>CYP17A1</i>	ENCSR000EOK	7	0.069	1.11×10^{-6}	0.018
<i>HSD11B2</i>	ENCSR000EOK	4	0.150	0.104	0.248
<i>KCNJ1</i>	ENCSR000EOK	13	0.992	0.958	0.789
<i>KCNJ5</i>	ENCSR000EOK	11	0.550	0.742	0.982
<i>KLHL3</i>	ENCSR000EOK	12	0.573	0.431	0.345
<i>NR3C2</i>	ENCSR000EOK	26	0.375	0.270	0.148
<i>SCNN1A</i>	ENCSR000EOK	7	0.502	0.452	0.767
<i>SCNN1B</i>	ENCSR000EOK	7	0.053	0.310	0.479
<i>SCNN1G</i>	ENCSR000EOK	8	0.170	0.276	1.000
<i>SLC12A1</i>	ENCSR000EOK	6	1.000	0.909	0.713
<i>SLC12A3</i>	ENCSR000EOK	16	0.233	0.168	0.491
<i>WNK1</i>	ENCSR000EOK	11	0.365	0.557	0.637
<i>WNK4</i>	ENCSR000EOK	1	0.010	0.010	0.010
<i>BSND</i>	ENCSR000EOM	12	0.861	0.737	0.562
<i>CASR</i>	ENCSR000EOM	2	0.860	0.790	0.672
<i>CLCNKA</i>	ENCSR000EOM	10	0.688	0.657	0.540
<i>CLCNKB</i>	ENCSR000EOM	10	0.688	0.657	0.540
<i>CUL3</i>	ENCSR000EOM	8	0.081	0.250	0.044
<i>CYP11B1</i>	ENCSR000EOM	4	0.018	0.018	0.281
<i>CYP11B2</i>	ENCSR000EOM	3	0.015	0.015	0.281
<i>CYP17A1</i>	ENCSR000EOM	11	3.64×10^{-8}	2.88×10^{-7}	0.013

<i>HSD11B2</i>	ENCSR000EOM	6	0.024	0.044	0.595
<i>KCNJ1</i>	ENCSR000EOM	9	0.086	0.554	0.060
<i>KCNJ5</i>	ENCSR000EOM	6	0.084	0.150	0.163
<i>KLHL3</i>	ENCSR000EOM	12	0.065	0.012	0.208
<i>NR3C2</i>	ENCSR000EOM	10	0.278	0.126	0.149
<i>SCNN1A</i>	ENCSR000EOM	5	0.517	0.572	0.727
<i>SCNN1B</i>	ENCSR000EOM	5	0.815	0.930	0.873
<i>SCNN1G</i>	ENCSR000EOM	5	0.353	0.298	0.540
<i>SLC12A1</i>	ENCSR000EOM	4	0.742	0.896	0.704
<i>SLC12A3</i>	ENCSR000EOM	12	0.062	0.077	0.245
<i>WNK1</i>	ENCSR000EOM	7	0.151	0.158	0.864
<i>WNK4</i>	ENCSR000EOM	1	0.010	0.010	0.010
<i>BSND</i>	ENCSR000EPW	10	0.649	0.702	0.591
<i>CASR</i>	ENCSR000EPW	11	0.665	0.798	0.335
<i>CLCNKA</i>	ENCSR000EPW	3	0.190	0.203	0.433
<i>CLCNKB</i>	ENCSR000EPW	4	0.268	0.176	0.464
<i>CUL3</i>	ENCSR000EPW	8	0.030	0.195	0.039
<i>CYP11B1</i>	ENCSR000EPW	3	0.015	0.015	0.281
<i>CYP11B2</i>	ENCSR000EPW	5	0.015	0.017	0.488
<i>CYP17A1</i>	ENCSR000EPW	9	0.006	9.92×10^{-7}	0.045
<i>HSD11B2</i>	ENCSR000EPW	1	0.051	0.051	0.051
<i>KCNJ1</i>	ENCSR000EPW	7	0.960	0.954	0.411
<i>KCNJ5</i>	ENCSR000EPW	10	0.519	0.699	0.417
<i>KLHL3</i>	ENCSR000EPW	8	0.380	0.525	0.428
<i>NR3C2</i>	ENCSR000EPW	24	0.264	0.189	0.027
<i>SCNN1A</i>	ENCSR000EPW	6	0.492	0.360	0.767
<i>SCNN1B</i>	ENCSR000EPW	7	0.141	0.181	0.482
<i>SCNN1G</i>	ENCSR000EPW	6	0.778	0.565	1.000
<i>SLC12A1</i>	ENCSR000EPW	4	1.000	0.896	0.704
<i>SLC12A3</i>	ENCSR000EPW	13	0.135	0.116	0.358
<i>WNK1</i>	ENCSR000EPW	11	0.526	0.515	0.054
<i>WNK4</i>	ENCSR000EPW	1	0.010	0.010	0.010
<i>BSND</i>	ENCSR785BDQ	10	0.888	0.854	0.668
<i>CASR</i>	ENCSR785BDQ	1	0.465	0.465	0.465
<i>CLCNKA</i>	ENCSR785BDQ	10	0.466	0.606	0.682
<i>CLCNKB</i>	ENCSR785BDQ	11	0.457	0.476	0.695
<i>CUL3</i>	ENCSR785BDQ	2	0.641	0.314	0.982
<i>CYP11B1</i>	ENCSR785BDQ	4	0.387	0.353	0.026
<i>CYP11B2</i>	ENCSR785BDQ	2	0.025	0.026	0.996
<i>CYP17A1</i>	ENCSR785BDQ	12	0.008	3.93×10^{-5}	0.020
<i>HSD11B2</i>	ENCSR785BDQ	6	0.205	0.067	0.061

<i>KCNJ1</i>	ENCSR785BDQ	7	0.961	0.879	0.400
<i>KCNJ5</i>	ENCSR785BDQ	2	1.000	1.000	1.000
<i>KLHL3</i>	ENCSR785BDQ	10	0.498	0.126	0.888
<i>NR3C2</i>	ENCSR785BDQ	9	0.314	0.290	0.037
<i>SCNN1A</i>	ENCSR785BDQ	9	0.521	0.700	0.980
<i>SCNN1B</i>	ENCSR785BDQ	4	1.000	1.000	0.983
<i>SCNN1G</i>	ENCSR785BDQ	6	0.727	0.565	1.000
<i>SLC12A1</i>	ENCSR785BDQ	2	0.848	0.848	0.848
<i>SLC12A3</i>	ENCSR785BDQ	28	0.250	0.440	0.666
<i>WNK1</i>	ENCSR785BDQ	7	0.289	0.414	0.128
<i>WNK4</i>	ENCSR785BDQ	-	-	-	-

N.sbp, number of variants analyzed in SKAT analysis of SBP in GERA; p.sbp.*, p-values from

SKAT analysis of BP in GERA with deltasvm (dsvm), eq (equal), or default (def) weights

^a Experiments: ENCSR000EOK, renal cortical epithelial cell; ENCSR000EOM, glomerular

endothelial cell; ENCSR000EPW, epithelial cell of proximal tubule; ENCSR785BDQ,

glomerular visceral epithelial cell

Table 7. Kidney DBP SKAT results.

Gene	Experiment ^a	N.dbp	p.dbp.dsvm	p.dbp.eq	p.dbp.def
<i>BSND</i>	ENCSR000EOK	13	0.938	0.831	0.061
<i>CASR</i>	ENCSR000EOK	21	0.990	0.986	0.856
<i>CLCNKA</i>	ENCSR000EOK	8	0.126	0.283	0.108
<i>CLCNKB</i>	ENCSR000EOK	9	0.131	0.410	0.137
<i>CUL3</i>	ENCSR000EOK	18	0.270	0.183	0.855
<i>CYP11B1</i>	ENCSR000EOK	12	0.102	0.095	0.013
<i>CYP11B2</i>	ENCSR000EOK	10	0.091	0.094	0.023
<i>CYP17A1</i>	ENCSR000EOK	7	0.505	0.189	0.337
<i>HSD11B2</i>	ENCSR000EOK	4	0.045	0.023	0.114
<i>KCNJ1</i>	ENCSR000EOK	13	0.969	0.879	0.398
<i>KCNJ5</i>	ENCSR000EOK	11	0.722	0.598	0.487
<i>KLHL3</i>	ENCSR000EOK	12	0.835	0.922	0.622
<i>NR3C2</i>	ENCSR000EOK	26	0.730	0.766	0.104
<i>SCNN1A</i>	ENCSR000EOK	7	0.751	0.621	0.599
<i>SCNN1B</i>	ENCSR000EOK	7	3.33 x 10 ⁻⁴	0.012	0.053
<i>SCNN1G</i>	ENCSR000EOK	8	0.074	0.027	0.461
<i>SLC12A1</i>	ENCSR000EOK	6	0.361	0.259	0.860

<i>SLC12A3</i>	ENCSR000EOK	16	0.069	0.072	0.175
<i>WNK1</i>	ENCSR000EOK	11	0.375	0.300	0.442
<i>WNK4</i>	ENCSR000EOK	1	0.828	0.828	0.828
<i>BSND</i>	ENCSR000EOM	12	0.613	0.776	0.052
<i>CASR</i>	ENCSR000EOM	2	0.877	0.825	0.759
<i>CLCNKA</i>	ENCSR000EOM	10	0.349	0.546	0.133
<i>CLCNKB</i>	ENCSR000EOM	10	0.349	0.546	0.133
<i>CUL3</i>	ENCSR000EOM	8	0.290	0.266	0.798
<i>CYP11B1</i>	ENCSR000EOM	4	0.117	0.129	0.689
<i>CYP11B2</i>	ENCSR000EOM	3	0.096	0.105	0.689
<i>CYP17A1</i>	ENCSR000EOM	11	0.053	0.076	0.187
<i>HSD11B2</i>	ENCSR000EOM	6	0.012	0.012	0.387
<i>KCNJ1</i>	ENCSR000EOM	9	0.098	0.317	0.101
<i>KCNJ5</i>	ENCSR000EOM	6	0.106	0.499	0.598
<i>KLHL3</i>	ENCSR000EOM	12	0.897	0.924	0.515
<i>NR3C2</i>	ENCSR000EOM	10	0.671	0.302	0.072
<i>SCNN1A</i>	ENCSR000EOM	5	0.782	0.709	0.547
<i>SCNN1B</i>	ENCSR000EOM	5	0.763	0.907	0.913
<i>SCNN1G</i>	ENCSR000EOM	5	0.167	0.085	0.920
<i>SLC12A1</i>	ENCSR000EOM	4	0.118	0.244	0.865
<i>SLC12A3</i>	ENCSR000EOM	12	0.008	0.044	0.234
<i>WNK1</i>	ENCSR000EOM	7	0.499	0.589	1.000
<i>WNK4</i>	ENCSR000EOM	1	0.828	0.828	0.828
<i>BSND</i>	ENCSR000EPW	10	0.866	0.769	0.062
<i>CASR</i>	ENCSR000EPW	11	0.921	0.845	0.872
<i>CLCNKA</i>	ENCSR000EPW	3	0.431	0.438	0.367
<i>CLCNKB</i>	ENCSR000EPW	4	0.712	0.639	0.413
<i>CUL3</i>	ENCSR000EPW	8	0.239	0.085	0.783
<i>CYP11B1</i>	ENCSR000EPW	3	0.104	0.105	0.689
<i>CYP11B2</i>	ENCSR000EPW	5	0.104	0.102	0.599
<i>CYP17A1</i>	ENCSR000EPW	9	0.360	0.185	0.262
<i>HSD11B2</i>	ENCSR000EPW	1	0.006	0.006	0.006
<i>KCNJ1</i>	ENCSR000EPW	7	0.709	0.873	0.366
<i>KCNJ5</i>	ENCSR000EPW	10	0.680	0.546	0.190
<i>KLHL3</i>	ENCSR000EPW	8	0.817	0.867	0.448
<i>NR3C2</i>	ENCSR000EPW	24	0.489	0.374	0.032
<i>SCNN1A</i>	ENCSR000EPW	6	0.733	0.446	0.599
<i>SCNN1B</i>	ENCSR000EPW	7	0.079	0.017	0.053
<i>SCNN1G</i>	ENCSR000EPW	6	0.647	0.061	0.297
<i>SLC12A1</i>	ENCSR000EPW	4	0.415	0.244	0.865
<i>SLC12A3</i>	ENCSR000EPW	13	0.017	0.050	0.217

<i>WNK1</i>	ENCSR000EPW	11	0.300	0.173	0.031
<i>WNK4</i>	ENCSR000EPW	1	0.828	0.828	0.828
<i>BSND</i>	ENCSR785BDQ	10	0.645	0.492	0.062
<i>CASR</i>	ENCSR785BDQ	1	0.967	0.967	0.967
<i>CLCNKA</i>	ENCSR785BDQ	10	0.291	0.457	0.268
<i>CLCNKB</i>	ENCSR785BDQ	11	0.296	0.568	0.303
<i>CUL3</i>	ENCSR785BDQ	2	0.145	0.040	0.248
<i>CYP11B1</i>	ENCSR785BDQ	4	0.154	0.134	0.003
<i>CYP11B2</i>	ENCSR785BDQ	2	0.105	0.107	0.508
<i>CYP17A1</i>	ENCSR785BDQ	12	0.551	0.218	0.355
<i>HSD11B2</i>	ENCSR785BDQ	6	0.079	0.008	0.003
<i>KCNJ1</i>	ENCSR785BDQ	7	0.848	0.795	0.128
<i>KCNJ5</i>	ENCSR785BDQ	2	1.000	1.000	0.875
<i>KLHL3</i>	ENCSR785BDQ	10	0.802	0.778	0.424
<i>NR3C2</i>	ENCSR785BDQ	9	0.877	0.784	0.019
<i>SCNN1A</i>	ENCSR785BDQ	9	0.789	0.949	0.925
<i>SCNN1B</i>	ENCSR785BDQ	4	0.345	0.445	0.555
<i>SCNN1G</i>	ENCSR785BDQ	6	0.083	0.061	0.297
<i>SLC12A1</i>	ENCSR785BDQ	2	0.282	0.282	0.282
<i>SLC12A3</i>	ENCSR785BDQ	28	0.044	0.415	0.393
<i>WNK1</i>	ENCSR785BDQ	7	0.741	0.416	0.079
<i>WNK4</i>	ENCSR785BDQ	-	-	-	-

N.dbp, number of variants analyzed in SKAT analysis of DBP in GERA; p.dbp.*, p-values from

SKAT analysis of DBP in GERA with deltasvm (dsvm), eq (equal), or default (def) weights

^a Experiments: ENCSR000EOK, renal cortical epithelial cell; ENCSR000EOM, glomerular

endothelial cell; ENCSR000EPW, epithelial cell of proximal tubule; ENCSR785BDQ,

glomerular visceral epithelial cell

Quantile-quantile (QQ) plots are shown in Figs S3-S14 for each of the six tissues and two

BP phenotypes, and each of three delta SVM weighting schemes. Although the dsvm weighting

scheme demonstrates a greater enrichment of genes than the default weighting scheme, the equal

weighting scheme marginally presents the greatest enrichment. In many cases, deltaSVM

discriminates between different tissue/cell types while equal-weighted results do not; this is especially clear with the QT interval results.

We analyzed the union of genes that met significance in the association analysis with BP, regardless of association of expression, to maximize our gene list for annotation, using DAVID 6.8 [48,49]. The results are shown in Table S6.

Discussion

Our previous genetic analyses identified the aorta and tibial arteries as relevant to blood pressure regulation [38]. In this study we have now identified several genes with regulatory variants linking significantly to both BP traits and to expression data in these tissues, most at previously replicated BP loci. Although the involvement of the kidney is well established in BP regulation through physiological evidence, we sought to identify genes at any of the hundreds of BP GWAS loci in a broader set of tissues. We examined groupings of multiple proximal and putatively causal variants defined around genes within a single tissue in order to identify specific genes of interest. We also examined QT interval genes at previous GWAS loci to highlight the identification of functionally characterized genes for this trait.

We identified several genes of potential interest to the aorta/arteries for BP, mostly at previously identified GWAS loci: *MTHFR*, *C10orf32*, *CSK*, *NOV*, *ULK4*, *SDCCAG8*, *SCAMP5*, *RPP25*, *HDGFRP3*, *VPS37B*, and *PPCDC*. We note here that our method of identifying genes at previous loci was conservative: *RPP25* was not present in this list, but is present just outside the TAD boundary used. In addition to its role in the progression of various cancers, the *NOV* gene has been identified as a player in angiogenesis [50,51] and vascular homeostasis [52]. The *ULK4* gene has been previously associated with DBP [43], and variation in this gene has also

been associated with aortic disease and acute aortic dissections [53]. The association of a homozygous variant (C677T) in its neighboring gene, *MTHFR*, has long been associated with BP and vascular disease [54–57]; more generally, this locus has been identified in large BP GWAS [40,41]. The locus including *C10orf32* has been identified previously [40] and neighbors the well-studied *CYP17A1* gene. Though we initially examined only the latter among kidney cell types, because of its known role in monogenic hypertension, we note that both genes show BP association in endothelial contexts as well, but it is *C10orf32* that has strong expression support in the artery datasets in our study, while *CYP17A1* does not [58]. The gene *SDCCAG8* is a centrosomal protein linked with nephronophthisis-related ciliopathies (OMIM: Senior-Loken Syndrome 7, 613615; Bardet-Biedl syndrome-16, 615993, and Airik et al. [59]), and is expressed in the kidney and lung epithelia [59]. The *CSK* gene, encoding a tyrosine kinase, is at a previous BP GWAS locus [60] and has been found to be associated with SBP in young children [61]; there is also prior evidence through experiments in mouse aortas that this gene regulates blood pressure through Src [62]. Finally, the *SCAMP5* and *PPCDC* genes (within the same locus) [40], and *RPP25* [43], are previously identified BP genes.

As mentioned above, one major limitation in our study is the statistical power of the SKAT eQTL analysis, with small sample sizes available for each of the GTEx tissues. The power of implicating effects for a given tissue also depends on its total contribution, and the numbers of eQTLs identified. The requirement in our study that a gene meet significance for both BP and expression therefore produced a more conservative list. However, the QT interval results, especially for the *SCN5A* gene, still illustrate the utility of this method. The availability of additional samples in the future will contribute to the success of this method in identifying genes of interest with greater statistical power. The gene annotation analyses revealed no clear

BP-specific pathways or annotation, so these will also benefit from producing more specific and possibly larger gene sets. Additionally, we used hard genotype calls for analysis, necessitating some missing genotype data; the power of our methods could be improved by using imputed probabilities of genotypes.

Our attempts to expand findings beyond the known pathogenic coding variation with respect to the 20 genes involved in monogenic forms of hypertension or hypotension were inconclusive. We attribute this to the dearth of publicly available data for the kidney at this time, and expect that the availability of more extensive data will resolve some of the issues in further studies. Additionally, though it is beyond the scope of this study, as the effects of many of these monogenic disorders are likely through the adrenal gland, a full analysis of adrenal gland data will be necessary to assess them.

The MetaXcan software has supported most of the genes highlighted here and identified novel associations, although there were some limitations with the availability of the models for all genes. Additionally, our results indicated that deltaSVM weighting might be validly discriminatory between cell types; this is most evident with several QT interval genes, such as *NOS1AP* and *SCN5A*. It is also suggestive of cell-type specificity with the results for *CYP17A1* in the kidney cell types. It may be informative moving forward to characterize these BP genes at the individual cell-type level in the arteries as well.

The question of identification of core genes networks may be facilitated by our approach in this study, which includes using eQTL information from tissues or cell types of interest and genotypes to identify potentially relevant genes for a trait. As the expansion of publicly available resources continues, more information may be used for these purposes. Our analysis implicates

specific variants that can be functionally tested for their effect on both gene expression and the phenotype.

Materials and Methods

Study participants and summary of genotypes, phenotypes, and association results used in this study

The full descriptions of the prior underlying studies, phenotypes, and association results for the GERA cohort are in Hoffmann et al. [38] and are briefly recapitulated here. The Genetic Epidemiology Research on Adult Health and Aging (GERA) cohort, part of the Kaiser Permanente Research Program on Genes, Environment, and Health (RPGEH), consists of individuals from five ethnic backgrounds; the majority is non-Hispanic white (EUR), with the remainder including Latino, East Asians, African Americans, and South Asians. A total of 99,785 individuals were analyzed, of which 80,792 were EUR individuals. The populations were each genotyped on custom population-specific Affymetrix Axiom SNP genotyping arrays [63,64] and imputed to the 1000 Genomes Phase I Integrated Release Version 3 haplotype panel. Analyses of GERA alone, with the results of the International Consortium for Blood Pressure (ICBP, n=69,396) study [65], and with the ICBP and the UK Biobank (UKB, n=152,081) study [66], identified 316 novel BP loci. Combined with the set of replicated BP GWAS loci available at that time, there were a total of 390 BP loci we considered to be of interest. Of these, 367 had minor allele frequency (MAF) > 0.001 in the GERA EUR study, which was used as the reference population for the eQTL analyses described below.

For the purpose of several of the analyses described in this paper, we used these association results, as well as summary statistics available from 80,792 GERA EUR individuals

from the Hoffmann et al. [38] study, and genotypes from a subset of 71,404 GERA EUR ‘unrelated’ individuals (third degree or beyond, pruned by the KING software for relationship inference) [67]. We converted genotypes prepared in the Hoffmann et al. [38] study after imputation from IMPUTE2 genotype probabilities format to PLINK ‘hard’ calls (the most likely genotype), setting genotypes with uncertainty greater than 0.25 to missing, and retaining variants with < 10% missing data, a Hardy Weinberg equilibrium test $p < 1 \times 10^{-6}$, and imputation quality score ≥ 0.3 . In order to report univariate summary statistics within the 71,404 individuals, we used the --assoc option for analysis of a quantitative trait (Wald test) with PLINK v1.9 [68]. We analyzed covariate-adjusted longitudinal systolic (SBP) and diastolic (DBP) blood pressure in this study, as also described in Hoffmann et al. [38].

ARIC genotypes, phenotypes, and association

The Atherosclerosis Risk in Communities (ARIC) study cohort is a longitudinal population-based study of 15,792 individuals, including 11,478 European-Americans (EUR) and 4,266 African-Americans (AA) from four study centers: Washington County, MD; Forsyth County, NC; Jackson, MS; and, Minneapolis, MN [24,25]. The initial examination occurred from 1987-1989, with participants aged between 45 and 64 years. Subsequent examinations occurred in 1990-1992, 1993-1995, 1996-1998, and 2011-2013, with the most recent visits (6+) beginning in 2016. We analyzed 9,083 individuals of European ancestry with genotypes and QT interval at baseline. The genotyping of these samples on the Affymetrix genome-wide Human SNP Array 6.0, quality control, and imputation to the 1000 Genomes Phase I Integrated Release Version 3 haplotype panel are described elsewhere [69,70]. We converted IMPUTE2 genotype probabilities to PLINK ‘hard’ calls, setting genotypes with uncertainty greater than 0.25 to

missing, and retaining variants with < 10% missing data, a Hardy Weinberg equilibrium test $p < 1 \times 10^{-6}$, and imputation quality score ≥ 0.3 , (as for the GERA study). The phenotypes were analyzed as previously described [71] with QT residuals generated by adjusting raw QT intervals for age, sex and resting heart rate. Summary statistics were generated for single variants using the --assoc option for analysis of a quantitative trait (Wald test) with PLINK v1.9 [68].

GTEx genotypes and expression

We analyzed genotypes and expression data from the Genotype-Tissue Expression (GTEx; phs000424.v6.p1) Project [72] v6p for the SKAT analysis (see below) from the aorta, tibial artery, heart left ventricle, and heart atrial appendage tissues. Normalized expression was analyzed for these tissues, with the top three principal components, available PEER factors (15-35, depending on sample size), genotyping array platform, and sex used as covariates, all available from the GTEx portal. We used SNP-gene associations from the associated *.v6p.all_snpgene_pairs.txt.gz files from the authors' eQTL analyses.

Partitioned heritability analyses

We used the stratified LD score regression method and software [15] for estimating the heritability of the trait partitioned by genomic element using summary statistics for SBP and DBP from 80,792 GERA EUR individuals [38]. The 'mungestats.py' script was used to format the summary statistics as appropriate, and we analyzed them using the baseline model with 53 categories which included coding, UTR, and intronic regions, in addition to various open chromatin and histone modification annotations as described by the authors, as well as the 1000 Genomes Phase 3 reference files with the weights from their weights_hm3_no_hla.tgz file,

which were provided and described by the authors on their website

(<https://github.com/bulik/ldsc/wiki/Partitioned-Heritability>).

Generation of putative regulatory element maps and deltaSVM scores are described in the supplementary methods.

Gene-based testing with SKAT

We used the sequence-kernel association test (SKAT) [22,73] to test genes with median reads per kilobase of transcript, per million mapped reads (RPKM) ≥ 0.3 in GTEx samples for the aorta (n=197) and tibial (n=285) arteries with their respective variant sets. For each gene, we tested all variants within 50Kb of the gene start or end, inclusive of the entire gene body, per GENCODE v19 annotations (<https://www.encodegenes.org/releases/19.html>). The weights used were taken as the absolute value of the deltaSVM score for each variant to reflect its predicted impact; for comparison, we also ran SKAT using default weights with beta density parameters (weights.beta=c(1,25), which up-weights rare variants as compared to common variants), as well as equal weights to all variants (weights.beta=c(1,1)). We tested association of each gene with adjusted SBP and DBP phenotype residuals (see above), as well as the GTEx normalized expression data with covariates (release v6p, <https://www.gtexportal.org/>), from the aorta and tibial arteries. We restricted our primary analyses in each of the kidney cell types to the 20 monogenic hypertension and hypotension genes. We additionally tested tissue- or cell-type-specific groupings in the ARIC dataset with the adjusted QT interval phenotype using the sets for the heart and heart tissues from GTEx, arteries and kidney cell types, as described above.

Prediction of gene expression association with blood pressure

We used the MetaXcan [23] software with prebuilt HapMap training models for the GTEx (<https://www.gtexportal.org/>) tissues aorta and tibial arteries, provided by the authors at <http://predictdb.hakyimlab.org/>, with summary statistics from 80,792 GERA EUR individuals for SBP and DBP. We also used the software with the provided models for heart left ventricle and atrial appendage, for the QT interval analysis using summary statistics from 9,083 ARIC EUR individuals. MetaXcan is an extension of the PrediXcan [74] method, which predicts gene expression from genotypes and tests association of predicted expression with phenotypes using summary association results.

Statistical significance

Statistical significance was determined using the Benjamini-Hochberg [75] (BH) method for multiple test correction to adjust for the number of genes within each analysis. We made no additional adjustments for the number of tissues, in part due to the correlation of specific subsets (the arteries, and individual kidney cell types), and as we examined genes across multiple analyses, for phenotype and for expression.

Annotation of artery-significant genes

We used DAVID 6.8 [48,49] to annotate the set of genes that met significance in the association analysis with either SBP or DBP. We retained terms that met a BH threshold of $p < 0.05$.

Acknowledgements and Funding

The Atherosclerosis Risk in Communities Study is carried out as a collaborative study supported by National Heart, Lung, and Blood Institute contracts (HHSN268201100005C,

HHSN268201100006C, HHSN268201100007C, HHSN268201100008C,
HHSN268201100009C, HHSN268201100010C, HHSN268201100011C, and
HHSN268201100012C), R01HL087641, R01HL59367 and R01HL086694; National Human
Genome Research Institute contract U01HG004402; and National Institutes of Health contract
HHSN268200625226C. The authors thank the staff and participants of the ARIC study for their
important contributions. Infrastructure was partly supported by Grant Number UL1RR025005, a
component of the National Institutes of Health and NIH Roadmap for Medical Research. The
Genotype-Tissue Expression (GTEx) Project was supported by the Common Fund of the Office
of the Director of the National Institutes of Health, and by NCI, NHGRI, NHLBI, NIDA, NIMH,
and NINDS. The data used for the analyses described in this manuscript were obtained from: the
GTEx Portal and dbGaP accession number phs000424.v6.p1. This research was funded by NIH
grants HL128782 and HL0-86694 to A.C.

We are grateful to the Kaiser Permanente Northern California members who have
generously agreed to participate in the Kaiser Permanente Research Program on Genes,
Environment, and Health. Support for participant enrollment, survey completion, and
biospecimen collection for the RPGEH was provided by the Robert Wood Johnson Foundation,
the Wayne and Gladys Valley Foundation, the Ellison Medical Foundation, and Kaiser
Permanente Community Benefit Programs. Genotyping of the GERA cohort was funded by a
grant from the National Institute on Aging, National Institute of Mental Health, and the National
Institute of Health Common Fund (RC2 AG036607 to CAS and NJR). GE receives support from
Geneva University Hospitals and The Foundation of Medical Researchers, Geneva.

References

- 571 1. Boyle EA, Li YI, Pritchard JK. An Expanded View of Complex Traits: From Polygenic to
572 Omnigenic. *Cell*. 2017 Jun 15;169(7):1177–86.
- 573 2. Kirschner MW. The meaning of systems biology. *Cell*. 2005 May 20;121(4):503–4.
- 574 3. Alon U. *An Introduction to Systems Biology: Design Principles of Biological Circuits*. 1st
575 ed. Chapman and Hall/CRC; 2006.
- 576 4. Segal E, Raveh-Sadka T, Schroeder M, Unnerstall U, Gaul U. Predicting expression
577 patterns from regulatory sequence in *Drosophila* segmentation. *Nature*. 2008 Jan
578 31;451(7178):535–40.
- 579 5. Emilsson V, Thorleifsson G, Zhang B, Leonardson AS, Zink F, Zhu J, et al. Genetics of
580 gene expression and its effect on disease. *Nature*. 2008 Mar 27;452(7186):423–8.
- 581 6. Zhong H, Yang X, Kaplan LM, Molony C, Schadt EE. Integrating pathway analysis and
582 genetics of gene expression for genome-wide association studies. *Am J Hum Genet*. 2010
583 Apr 9;86(4):581–91.
- 584 7. Pickrell JK, Marioni JC, Pai AA, Degner JF, Engelhardt BE, Nkadori E, et al.
585 Understanding mechanisms underlying human gene expression variation with RNA
586 sequencing. *Nature*. 2010 Apr 1;464(7289):768–72.
- 587 8. Davidson EH. Emerging properties of animal gene regulatory networks. *Nature*. 2010 Dec
588 16;468(7326):911–20.
- 589 9. Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, et al. Topological domains in
590 mammalian genomes identified by analysis of chromatin interactions. *Nature*. 2012 Apr
591 11;485(7398):376–80.
- 592 10. Phillips-Cremins JE. Unraveling architecture of the pluripotent genome. *Curr Opin Cell*
593 *Biol*. 2014 Jun;28:96–104.
- 594 11. Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. Transposition of native
595 chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding
596 proteins and nucleosome position. *Nat Methods*. 2013 Dec;10(12):1213–8.
- 597 12. ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human
598 genome. *Nature*. 2012 Sep 6;489(7414):57–74.
- 599 13. Roadmap Epigenomics Consortium, Kundaje A, Meuleman W, Ernst J, Bilenky M, Yen A,
600 et al. Integrative analysis of 111 reference human epigenomes. *Nature*. 2015 Feb
601 19;518(7539):317–30.
- 602 14. Maurano MT, Haugen E, Sandstrom R, Vierstra J, Shafer A, Kaul R, et al. Large-scale
603 identification of sequence variants influencing human transcription factor occupancy in
604 vivo. *Nat Genet*. 2015 Dec;47(12):1393–401.

15. Finucane HK, Bulik-Sullivan B, Gusev A, Trynka G, Reshef Y, Loh P-R, et al. Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nat Genet.* 2015 Nov;47(11):1228–35.
16. Lee D, Gorkin DU, Baker M, Strober BJ, Asoni AL, McCallion AS, et al. A method to predict the impact of regulatory variants from DNA sequence. *Nat Genet.* 2015 Aug;47(8):955–61.
17. Lifton RP. Molecular genetics of human blood pressure variation. *Science.* 1996 May 3;272(5262):676–80.
18. Lifton RP, Gharavi AG, Geller DS. Molecular mechanisms of human hypertension. *Cell.* 2001 Feb 23;104(4):545–56.
19. Choi M, Scholl UI, Yue P, Björklund P, Zhao B, Nelson-Williams C, et al. K⁺ channel mutations in adrenal aldosterone-producing adenomas and hereditary hypertension. *Science.* 2011 Feb 11;331(6018):768–72.
20. Banda Y, Kvale MN, Hoffmann TJ, Hesselson SE, Ranatunga D, Tang H, et al. Characterizing Race/Ethnicity and Genetic Ancestry for 100,000 Subjects in the Genetic Epidemiology Research on Adult Health and Aging (GERA) Cohort. *Genetics.* 2015 Aug;200(4):1285–95.
21. Kvale MN, Hesselson S, Hoffmann TJ, Cao Y, Chan D, Connell S, et al. Genotyping Informatics and Quality Control for 100,000 Subjects in the Genetic Epidemiology Research on Adult Health and Aging (GERA) Cohort. *Genetics.* 2015 Aug;200(4):1051–60.
22. Wu MC, Lee S, Cai T, Li Y, Boehnke M, Lin X. Rare-variant association testing for sequencing data with the sequence kernel association test. *Am J Hum Genet.* 2011 Jul 15;89(1):82–93.
23. Barbeira A, Shah KP, Torres JM, Wheeler HE, Torstenson ES, Edwards T, et al. MetaXcan: Sum- 537 mary Statistics Based Gene-Level Association Method Infers Accurate PrediXcan Results. *bioRxiv.* 538 2016 mar; Available from: <http://biorxiv.org/content/early/2016/03/23/045260.abstract>.
24. Atherosclerosis Risk in Communities. [Internet]. [Accessed: 23 October 2015]. Available at: <https://www2.csc.unc.edu/aric>.
25. The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. The ARIC investigators. *Am J Epidemiol.* 1989 Apr;129(4):687–702.
26. Salfati E, Morrison AC, Boerwinkle E, Chakravarti A. Direct Estimates of the Genomic Contributions to Blood Pressure Heritability within a Population-Based Cohort (ARIC). *PloS One.* 2015;10(7):e0133031.

- 639 27. Lee D, Kapoor A, Safi A, Song L, Halushka MK, Crawford GE, et al. Human cardiac cis-
640 regulatory elements, their cognate transcription factors, and regulatory DNA sequence
641 variants. *Genome Res.* 2018;28(10):1577–88.
- 642 28. Ghandi M, Lee D, Mohammad-Noori M, Beer MA. Enhanced regulatory sequence
643 prediction using gapped k-mer features. *PLoS Comput Biol.* 2014 Jul;10(7):e1003711.
- 644 29. Lee D. LS-GKM: a new gkm-SVM for large-scale datasets. *Bioinforma Oxf Engl.* 2016 Jul
645 15;32(14):2196–8.
- 646 30. Postema PG, Wilde AAM. The measurement of the QT interval. *Curr Cardiol Rev.* 2014
647 Aug;10(3):287–94.
- 648 31. Dekker JM, Crow RS, Hannan PJ, Schouten EG, Folsom AR, ARIC Study. Heart rate-
649 corrected QT interval prolongation predicts risk of coronary heart disease in black and
650 white middle-aged men and women: the ARIC study. *J Am Coll Cardiol.* 2004 Feb
651 18;43(4):565–71.
- 652 32. Newton-Cheh C, Larson MG, Corey DC, Benjamin EJ, Herbert AG, Levy D, et al. QT
653 interval is a heritable quantitative trait with evidence of linkage to chromosome 3 in a
654 genome-wide linkage analysis: The Framingham Heart Study. *Heart Rhythm.* 2005
655 Mar;2(3):277–84.
- 656 33. Busjahn A, Knoblauch H, Faulhaber HD, Boeckel T, Rosenthal M, Uhlmann R, et al. QT
657 interval is linked to 2 long-QT syndrome loci in normal subjects. *Circulation.* 1999 Jun
658 22;99(24):3161–4.
- 659 34. Arking DE, Pfeufer A, Post W, Kao WHL, Newton-Cheh C, Ikeda M, et al. A common
660 genetic variant in the NOS1 regulator NOS1AP modulates cardiac repolarization. *Nat*
661 *Genet.* 2006 Jun;38(6):644–51.
- 662 35. Tomás M, Napolitano C, De Giuli L, Bloise R, Subirana I, Malovini A, et al.
663 Polymorphisms in the NOS1AP gene modulate QT interval duration and risk of
664 arrhythmias in the long QT syndrome. *J Am Coll Cardiol.* 2010 Jun 15;55(24):2745–52.
- 665 36. Avery CL, Wassel CL, Richard MA, Highland HM, Bien S, Zubair N, et al. Fine mapping
666 of QT interval regions in global populations refines previously identified QT interval loci
667 and identifies signals unique to African and Hispanic descent populations. *Heart Rhythm.*
668 2017 Apr;14(4):572–80.
- 669 37. Arking DE, Pulit SL, Crotti L, van der Harst P, Munroe PB, Koopmann TT, et al. Genetic
670 association study of QT interval highlights role for calcium signaling pathways in
671 myocardial repolarization. *Nat Genet.* 2014 Aug;46(8):826–36.
- 672 38. Hoffmann TJ, Ehret GB, Nandakumar P, Ranatunga D, Schaefer C, Kwok P-Y, et al.
673 Genome-wide association analyses using electronic health records identify new loci
674 influencing blood pressure variation. *Nat Genet.* 2017 Jan;49(1):54–64.

39. Evangelou E, Warren HR, Mosen-Ansorena D, Mifsud B, Pazoki R, Gao H, et al. Genetic analysis of over one million people identifies 535 novel loci for blood pressure. *bioRxiv* [Internet]. 2017 Jan 1; Available from: <http://biorxiv.org/content/early/2017/10/11/198234.abstract>
40. Newton-Cheh C, Johnson T, Gateva V, Tobin MD, Bochud M, Coin L, et al. Genome-wide association study identifies eight loci associated with blood pressure. *Nat Genet.* 2009 Jun;41(6):666–76.
41. Johnson T, Gaunt TR, Newhouse SJ, Padmanabhan S, Tomaszewski M, Kumari M, et al. Blood pressure loci identified with a gene-centric array. *Am J Hum Genet.* 2011 Dec 9;89(6):688–700.
42. Wain LV, Verwoert GC, O'Reilly PF, Shi G, Johnson T, Johnson AD, et al. Genome-wide association study identifies six new loci influencing pulse pressure and mean arterial pressure. *Nat Genet.* 2011 Sep 11;43(10):1005–11.
43. Levy D, Ehret GB, Rice K, Verwoert GC, Launer LJ, Dehghan A, et al. Genome-wide association study of blood pressure and hypertension. *Nat Genet.* 2009 Jun;41(6):677–87.
44. Goldsmith O, Solomon DH, Horton R. Hypogonadism and mineralocorticoid excess. The 17-hydroxylase deficiency syndrome. *N Engl J Med.* 1967 Sep 28;277(13):673–7.
45. Li Q, Gao T, Yuan Y, Wu Y, Huang Q, Xie F, et al. Association of CYP17A1 Genetic Polymorphisms and Susceptibility to Essential Hypertension in the Southwest Han Chinese Population. *Med Sci Monit Int Med J Exp Clin Res.* 2017 May 24;23:2488–99.
46. Li C, Kim YK, Dorajoo R, Li H, Lee I-T, Cheng C-Y, et al. Genome-Wide Association Study Meta-Analysis of Long-Term Average Blood Pressure in East Asians. *Circ Cardiovasc Genet.* 2017 Apr;10(2):e001527.
47. Nguyen K-DH, Pihur V, Ganesh SK, Rakha A, Cooper RS, Hunt SC, et al. Effects of rare and common blood pressure gene variants on essential hypertension: results from the Family Blood Pressure Program, CLUE, and Atherosclerosis Risk in Communities studies. *Circ Res.* 2013 Jan 18;112(2):318–26.
48. Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc.* 2009;4(1):44–57.
49. Huang DW, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* 2009 Jan;37(1):1–13.
50. Lin CG, Chen C-C, Leu S-J, Grzeszkiewicz TM, Lau LF. Integrin-dependent functions of the angiogenic inducer NOV (CCN3): implication in wound healing. *J Biol Chem.* 2005 Mar 4;280(9):8229–37.

51. Lin CG, Leu S-J, Chen N, Tebeau CM, Lin S-X, Yeung C-Y, et al. CCN3 (NOV) is a novel angiogenic regulator of the CCN protein family. *J Biol Chem*. 2003 Jun 27;278(26):24200–8.
52. Shimoyama T, Hiraoka S, Takemoto M, Koshizaka M, Tokuyama H, Tokuyama T, et al. CCN3 inhibits neointimal hyperplasia through modulation of smooth muscle cell growth and migration. *Arterioscler Thromb Vasc Biol*. 2010 Apr;30(4):675–82.
53. Guo D-C, Grove ML, Prakash SK, Eriksson P, Hostetler EM, LeMaire SA, et al. Genetic Variants in LRP1 and ULK4 Are Associated with Acute Aortic Dissections. *Am J Hum Genet*. 2016 Sep 1;99(3):762–9.
54. Niu W-Q, You Y-G, Qi Y. Strong association of methylenetetrahydrofolate reductase gene C677T polymorphism with hypertension and hypertension-in-pregnancy in Chinese: a meta-analysis. *J Hum Hypertens*. 2012 Apr;26(4):259–67.
55. Frosst P, Blom HJ, Milos R, Goyette P, Sheppard CA, Matthews RG, et al. A candidate genetic risk factor for vascular disease: a common mutation in methylenetetrahydrofolate reductase. *Nat Genet*. 1995 May;10(1):111–3.
56. Nakata Y, Katsuya T, Takami S, Sato N, Fu Y, Ishikawa K, et al. Methylenetetrahydrofolate reductase gene polymorphism: relation to blood pressure and cerebrovascular disease. *Am J Hypertens*. 1998 Aug;11(8 Pt 1):1019–23.
57. Kosmas IP, Tatsioni A, Ioannidis JPA. Association of C677T polymorphism in the methylenetetrahydrofolate reductase gene with hypertension in pregnancy and pre-eclampsia: a meta-analysis. *J Hypertens*. 2004 Sep;22(9):1655–62.
58. Van Woudenberg M, Shin J, Bernard M, Syme C, Abrahamowicz M, Leonard G, et al. CYP17A1 and Blood Pressure Reactivity to Stress in Adolescence. *Int J Hypertens*. 2015;2015:734586.
59. Airik R, Slaats GG, Guo Z, Weiss A-C, Khan N, Ghosh A, et al. Renal-retinal ciliopathy gene Sdccag8 regulates DNA damage response signaling. *J Am Soc Nephrol JASN*. 2014 Nov;25(11):2573–83.
60. Ganesh SK, Tragante V, Guo W, Guo Y, Lanktree MB, Smith EN, et al. Loci influencing blood pressure identified using a cardiovascular gene-centric array. *Hum Mol Genet*. 2013 Apr 15;22(8):1663–78.
61. Ahn S-Y, Gupta C. Genetic Programming of Hypertension. *Front Pediatr*. 2017;5:285.
62. Lee H-J, Kang J-O, Kim S-M, Ji S-M, Park S-Y, Kim ME, et al. Gene Silencing and Haploinsufficiency of Csk Increase Blood Pressure. *PloS One*. 2016;11(1):e0146841.
63. Hoffmann TJ, Kvale MN, Hesselson SE, Zhan Y, Aquino C, Cao Y, et al. Next generation genome-wide association tool: design and coverage of a high-throughput European-optimized SNP array. *Genomics*. 2011 Aug;98(2):79–89.

64. Hoffmann TJ, Zhan Y, Kvale MN, Hesselson SE, Gollub J, Iribarren C, et al. Design and coverage of high throughput genotyping arrays optimized for individuals of East Asian, African American, and Latino race/ethnicity using imputation and a novel hybrid SNP selection algorithm. *Genomics*. 2011 Dec;98(6):422–30.
65. International Consortium for Blood Pressure Genome-Wide Association Studies, Ehret GB, Munroe PB, Rice KM, Bochud M, Johnson AD, et al. Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. *Nature*. 2011 Sep 11;478(7367):103–9.
66. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med*. 2015 Mar;12(3):e1001779.
67. Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M, Chen W-M. Robust relationship inference in genome-wide association studies. *Bioinforma Oxf Engl*. 2010 Nov 15;26(22):2867–73.
68. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007 Sep;81(3):559–75.
69. Maruthur NM, Li M, Halushka MK, Astor BC, Pankow JS, Boerwinkle E, et al. Genetics of Plasma Soluble Receptor for Advanced Glycation End-Products and Cardiovascular Outcomes in a Community-based Population: Results from the Atherosclerosis Risk in Communities Study. *PloS One*. 2015;10(6):e0128452.
70. Li M, Maruthur NM, Loomis SJ, Pietzner M, North KE, Mei H, et al. Genome-wide association study of 1,5-anhydroglucitol identifies novel genetic loci linked to glucose metabolism. *Sci Rep*. 2017 Jun 6;7(1):2812.
71. Kapoor A, Bakshy K, Xu L, Nandakumar P, Lee D, Boerwinkle E, et al. Rare coding TTN variants are associated with electrocardiographic QT interval in the general population. *Sci Rep*. 2016 Jun 20;6:28356.
72. GTEx Consortium, Laboratory, Data Analysis & Coordinating Center (LDACC)—Analysis Working Group, Statistical Methods groups—Analysis Working Group, Enhancing GTEx (eGTEx) groups, NIH Common Fund, NIH/NCI, et al. Genetic effects on gene expression across human tissues. *Nature*. 2017 11;550(7675):204–13.
73. Seunggeun Lee, with contributions from Larisa Miropolsky and Michael Wu (2015). SKAT: SNP-Set (Sequence) Kernel Association Test. R package version 1.1.2. <https://CRAN.R-project.org/package=SKAT>.
74. Gamazon ER, Wheeler HE, Shah KP, Mozaafari SV, Aquino-Michaels K, Carroll RJ, et al. A gene-based association method for mapping traits using reference transcriptome data. *Nat Genet*. 2015 Sep;47(9):1091–8.

75. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J R Stat Soc Ser B Methodol. 1995;57(1):289–300.

Supporting Information

Fig S1. Comparisons of deltaSVM association P values between the heart and the other tissues in this study. hrt, heart.

Fig S2. Comparisons of $-\log_{10}(P)$ differences of heart significant genes between deltaSVM weighting and equal weighting SKAT tests. hrt, heart; X, x-axis; Y, y-axis.

Fig S3. Aorta QQ plots, SBP.

Fig S4. Aorta QQ plots, DBP.

Fig S5. Tibial Artery QQ plots, SBP.

Fig S6. Tibial Artery QQ plots, DBP.

Fig S7. ENCSR000EOK QQ plots, SBP.

Fig S8. ENCSR000EOK QQ plots, DBP.

Fig S9. ENCSR000EOM QQ plots, SBP.

Fig S10. ENCSR000EOM QQ plots, DBP.

Fig S11. ENCSR000EPW QQ plots, SBP.

Fig S12. ENCSR000EPW QQ plots, DBP.

Fig S13. ENCSR785BDQ QQ plots, SBP.

Fig S14. ENCSR785BDQ QQ plots, DBP.

Table S1. Partitioned heritability results from baseline model for SBP.

Table S2. Partitioned heritability results from baseline model for DBP.

Table S3. deltaSVM performance results.

Table S4. SKAT and MetaXcan results for QT interval.

Table S5. Individual variants analyzed in kidney for C10orf32 and CYP17A1.

Table S6. DAVID enrichment analysis of artery-significant genes.

810 **Text S1. Supplementary methods, results, and references.**

811



