Integration of genetic fine-mapping and multi-omics data reveals candidate effector genes for hypertension

Stefan van Duijvenboden[1,2,3*], Julia Ramírez[1,4,5*], William J. Young[1,6], Kaya J. Olczak[1], Farah Ahmed[1], Mohammed J.A.Y. Alhammadi[7,], International Consortium of Blood Pressure[†], Christopher G. Bell[1‡], Andrew P. Morris[8,9‡], & Patricia B. Munroe[1,10‡]

[1]William Harvey Research Institute, Barts and the London Faculty of Medicine and Dentistry, Queen Mary University of London, London, EC1M 6BQ, U.K.
[2]Institute of Cardiovascular Science, University College London, London, U.K.
[3]Nuffield Department of Population Health, University of Oxford, Oxford, U.K.
[4]Aragon Institute of Engineering Research, University of Zaragoza, Spain.
[5]Centro de Investigación Biomédica en Red – Bioingeniería, Biomateriales y Nanomedicina, Spain.
[6]Barts Heart Centre, St Bartholomew's Hospital, London, EC1A 7BE, U.K.
[7]Khalifa University, Abu Dhabi, United Arab Emirates.
[8]Centre for Genetics and Genomics Versus Arthritis, Centre for Musculoskeletal Research, The University of Manchester, Manchester, U.K.
[9]National Institute of Health and Care Research, Manchester Biomedical Research Centre, Manchester University NHS Foundation Trust, Manchester Academic Health Science Centre, Manchester, U.K.
[10]National Institute of Health and Care Research, Barts Cardiovascular Biomedical Research Centre, Queen Mary University of London, London, EC1M 6BQ, U.K.

*These first authors contributed equally
[†]Membership of the consortium is listed in the Supplementary Material
[‡]These authors directed the study equally

Corresponding authors:
Prof Andrew Morris: email: andrew.morris-5@manchester.ac.uk
Prof Patricia Munroe: email: p.b.munroe@qmul.ac.uk

Abstract

Genome-wide association studies of blood pressure (BP) have identified >1000 loci, but the effector genes and biological pathways at these loci are mostly unknown. Using published association summary statistics, we conducted annotation-informed fine-mapping incorporating tissue-specific chromatin segmentation and colocalization to identify causal variants and candidate effector genes for systolic BP, diastolic BP, and pulse pressure. We observed 532 distinct signals associated with ≥2 BP traits and 84 with all three. For >20% of signals, a single variant accounted for >75% posterior probability, 65 were missense variants in known (*SLC39A8*, *ADRB2, DBH*) and previously unreported BP candidate genes (*NRIP1*, *MMP14*). In disease-relevant tissues, we colocalized >80 and >400 distinct signals for each BP trait with *cis*-eQTLs, and regulatory regions from promoter capture Hi-C, respectively. Integrating mouse, human disorder, gene expression and tissue abundance data and literature review, we provide consolidated evidence for 436 BP candidate genes for future functional validation and identify several new drug targets.

Introduction

Elevated blood pressure (BP) or hypertension affects over 1 billion people and is one of the most important risk-factors for cardiovascular disease (CVD), leading to significant mortality and morbidity worldwide[1]. It is estimated to cause more than 10 million deaths per year[2]. Approximately 95% of hypertension cases are referred to as primary or essential hypertension and genetics contributes up to 50% of BP variance[3], the remainder due to lifestyle influences. Genome-wide association studies (GWAS), bespoke targeted arrays (Cardio Metabochip) and Exome-array wide association studies (EAWAS) have been deployed across samples of diverse ancestries from consortia (International Consortium for BP), and large biobanks (UK Biobank[4], Million Veteran's Program[5], Biobank Japan[6], Korean Association Resource[7]). These studies have led to the identification of over 1,000 BP-associated loci, with both common and rare variant associations reported[8-13]. However, for most of these loci, the effector genes and relevant biological processes through which BP associations are mediated have yet to be characterised. Here, we use published GWAS meta-analysis summary statistics (n >757,000) of systolic BP (SBP), diastolic BP (DBP) and pulse pressure (PP)[8] to perform fine mapping of causal variants at BP loci. Through the integration of GWAS with tissue specific epigenomic annotations, colocalization of BP associations with expression quantitative loci (eQTLs) and protein quantitative loci (pQTL)and Hi-C promoter interaction data, we identify consolidated effector genes and causal pathways, and assess their potential for drug target identification or repurposing opportunities.

Material and Methods

Study data and detection of distinct association signals

We utilised summary statistics from previously reported GWAS meta-analyses of BP traits in up to 757,601 individuals of European ancestry from the International Consortium of Blood Pressure and UK Biobank[8] (ICBP+UKBB). Each contributing GWAS had been imputed up to reference panels from the 1000 Genomes Project[14,15] and/or Haplotype Reference Consortium[16]. After quality control, meta-analysis association summary statistics for SBP, DBP and PP were reported for 7,088,121,

7,160,657 and 7,088,842 single nucleotide variants (SNVs), respectively. An overview of the study design is provided in Figure S1.

We began by considering autosomal lead SNVs that have been reported at genome-wide significance (variable threshold according to study design) for SBP, DBP or PP in previously published GWAS of blood pressure traits, which we have collated and are summarised in the recent review by Magavern and colleagues[13]. We initially defined genomic regions as mapping 500kb up- and down-stream of each lead SNV. However, where genomic regions overlapped, they were combined as a single genomic region to account for potential linkage disequilibrium (LD) between previously reported lead SNVs. Genomic regions that did not attain genome-wide significance ($P < 5x10^{-8}$) in the ICBP+UKBB meta-analysis for any BP trait were not considered for downstream interrogation. We then performed approximate conditional analyses using GCTA-COJO[17] to detect distinct association signals at each genomic region for each BP trait separately, using European ancestry haplotypes from the 1000 Genomes Project (Phase 3, October 2014 release)[14] as a reference for LD. Within each genomic region, variants attaining genome-wide significance ($P < 5x10^{-8}$) in the joint GCTA-COJO model were selected as index SNVs for distinct association signals.

We next assessed the evidence that distinct association signals for SBP, DBP and PP were shared across multiple BP traits. At each genomic region, distinct association signals for two traits were considered to be the same if: (i) the index SNVs were the same for both traits; (ii) the index SNVs were colinear in the joint GCTA-COJO models for each trait after including the index SNV for the other trait in the model; or (iii) the P-value of the index SNV for one trait increased to $P > 0.05$ after including the index SNV for the other trait in the model, and the P-value of the index SNP for the other trait increased to $P > 0.0001$ for the corresponding reciprocal conditioning.

Enrichment of BP associations for genomic annotations

We used fGWAS[18] to identify genomic annotations enriched for SBP, DBP or PP association signals. We considered a total of 253 functional and regulatory annotations derived from: (i) genic regions (protein coding exons, 3' UTRs and 5' UTRs) as defined by the GENCODE Project[19]; and (ii) chromatin state predictions of promoters and enhancers across 125 tissues from the Roadmap Epigenome Consortium[20]

implemented in Epilogos (https://epilogos.altius.org/). For each BP trait separately, we used a forward-selection approach to derive a joint model of enriched annotations. At each iteration, we added the annotation to the joint fGWAS model that maximised the improvement in the penalised likelihood. We continued until no additional annotations improved the fit of the joint model (P<0.00020, Bonferroni correction for 253 annotations).

Fine-mapping distinct association signals for BP traits

For each trait, we began by approximating the Bayes' factor (BF), Λ_{ij}, in favour of association of the jth SNV at the ith distinct association signal using summary statistics from the ICBP+UKBB meta-analyses. Specifically,

$$\Lambda_{ij} = \exp\left[\frac{D_{ij} - lnK_{ij}}{2}\right],$$

where $D_{ij} = b_{ij}^2/v_{ij}$, and b_{ij} and v_{ij} are the allelic log-OR and corresponding variance, respectively, across K_{ij} contributing GWAS to the ICBP+UKBB meta-analysis (here $K_{ij} = 2$)[21]. At genomic regions with a single association signal, b_{ij} and v_{ij} were taken from the unconditional meta-analysis. However, for genomic regions with multiple association signals, b_{ij} and v_{ij} were taken from the joint GCTA-COJO model, conditioning on the index SNVs for all other signals at the locus. The posterior probability for the jth SNV at the ith distinct signal, was then given by $\pi_{ij} \propto \gamma_j \Lambda_{ij}$, where γ_j is the relative prior probability of causality for the jth SNV. We considered an annotation-informed prior model, for which

$\gamma_j = \exp\left[\sum_k \hat{\beta}_k z_{jk}\right]$,where the summation is over the enriched annotations, $\hat{\beta}_k$ is the estimated log-fold enrichment of the kth annotation from the final joint fGWAS model, and z_{jk} is an indicator variable taking the value 1 if the jth SNV maps to the kth annotation, and 0 otherwise. Finally, we derived a 99% credible set[22] for the ith distinct association signal by: (i) ranking all SNVs according to their posterior probability π_{ij}; and (ii) including ranked SNVs until their cumulative posterior probability attains or exceeds 0.99. For comparison, we also calculated the posterior probability for the jth SNV at the ith distinct signal under a uniform prior model for which $\gamma_j \propto 1$.

High Confidence SNV Gene Set Enrichment Analysis

Genomic Regions Enrichment of Annotations Tool (GREAT) v4.0.4[23] was used to explore the high confidence SNVs potential biological impact. The default GREAT association parameters for gene regulatory domains (Proximal 5 kb upstream, 1 kb downstream, plus Distal up to 1 Mb) were used and curated regulatory domains included. Input was SNV BED files for each of the three traits (SBP n = 208, DBP n = 224 and PP n = 158). GREAT analysis included gene ontology (GO) Biological Processes, Human Phenotype, Mouse Phenotype and Knockout data.

Functional annotation

We use Variant-effect predictor (VEP) analysis to identify missense variants and queried their overlap with high confidence causal variants from the credible set analysis (https://grch37.ensembl.org/Homo_sapiens/Tools/VEP)[24].

Transcription Factor Binding Motif Analysis

We used the Transcription Factor Affinity Prediction (TRAP) v3.0.5[25] multiple sequences option to explore any enrichment for Transcription Factor (TF) binding motifs within the high confidence non-coding variants for each of the three traits (SBP n= 178; DBP n=187; and PP n =137). Sequences around each non-coding SNV were expanded to +/-10bp (via AWK) and the FASTA sequence extracted (hg19) via the BEDtools v2.30.0 command getfasta[26]. The Transfac 2010.1 Vertebrate matrix set was interrogated with human_promoter set as background model and the results were required to pass a Benjamini-Hochberg multiple-testing correction.

Colocalization with eQTLs in BP-relevant tissues

We performed a Bayesian statistical procedure to assess whether our annotation informed GWAS fine-mapping colocalized with eQTL signals. We selected eQTL tissues relevant for blood pressure from the publicly available eQTL results from GTEx version 8[27]. The tissue selection was informed by tissue enrichment analysis from prior GWAS (adipose, adrenal gland, artery and heart[8,12]) and biological mechanisms known to regulate BP (kidney cortex, nerve, and brain). The annotation informed BF in favour of association of the jth SNV at the ith distinct association signal was defined as:

$$\Lambda_{ij}^{*} = \pi_{ij} \sum_{j} \Lambda_{ij}$$

In this expression, π_{ij} is the annotation-informed posterior probability, and Λ_{ij} is the BF defined above. GWAS results were lifted from hg19 to hg38 using the UCSC liftOver tool[28] to allow direct comparison with the hg38 eQTL data. We undertook colocalization using the annotation informed BF using the COLOC software package in R[29], only for those signals for which a 99% credible set variant was the lead eQTL SNV.

Single-cell RNA-seq dataset analysis

We used the CellsxGene single cell dataset[30] to explore patterns of single cell RNA-seq expression data for BP effector genes indicated from our colocalised eQTLs. We used the following human datasets: fetal adrenal tissue[31]; adult kidney[32]; adult heart[33]; adult brain including: cerebellum, cortex, hypothalamus, hippocampus, substantial nigra and adult adipose including subcutaneous and visceral adipose (donors were healthy or type 2 diabetic, with BMI range 23-60). From the obtained single cell mRNA expression data, we calculated cell-specific expression for each gene as the ratio of each cell type expression to the total expression across all cell-types. This analysis was conducted separately for each tissue. Genes with a relative expression of more than 75% were selected for presentation of cell-type specific expression.

Long-range chromatin interaction (Hi–C) analyses

We identified potential target genes of regulatory SNVs using long-range chromatin interaction (Hi–C) data from tissues and cell types relevant for blood pressure regulation (adrenal gland, left and right heart ventricles, hippocampus, and cortex)[34]. Hi–C data is corrected for genomic biases and distance using the Hi–C Pro and Fit-Hi-C pipelines according to Schmitt et al. (40 kb resolution—correction applied to interactions with 50 kb- 5Mb span)[35]. We selected the most significant promoter interactions for all potential regulatory SNPs (RegulomeDB score ≤3) that were included in the 99% credible sets and report the interactors with the SNPs of highest regulatory potential to annotate the loci.

Colocalization with pQTLs in plasma

Using the same Bayesian statistical approach as performed for eQTL colocalization, we assessed whether our annotation informed GWAS fine-mapping colocalized with

cis-pQTL results using plasma protein concentration summary statistics from a study using 4,907 aptamers (SomaScan v4 assay) in 35,559 Icelanders[36]. Colocalization was performed using the annotation informed BF using the R COLOC software package[29], only for signals for which a 99% credible set variant was the lead cis-pQTL SNV.

Collation of evidence for effector BP genes

A full list of candidate genes for each BP trait was collated from the results of our fine-mapping pipeline and computational approaches. A gene was indicated for a signal if there was support from a coding and high confidence variant in the gene at the locus, or if the gene was indicated from eQTL or pQTL colocalization or Hi-C analyses. To refine the list of candidate genes we next collated additional information for each gene using data from GeneCards (https://genealacart.genecards.org)[37]. This included the following: 1) a mouse model from Mouse Genome Informatics (MGI) which has a cardiovascular or renal phenotype. 2) A cardiovascular, vascular or renal phenotype described for the candidate gene in the Human Phenotype Ontology database 3) Differential RNA expression of the candidate gene in the GTEx database in cardiovascular, vascular or renal tissues, only genes with fold changes >4 in a tissue were selected. 4) Differential protein abundance of the candidate gene based on 69 integrated normal proteomics datasets in HIPED (the Human Integrated Protein Expression Database). Genes with a fold change value of >6 and protein abundance value of >0.1 PPM in an anatomical site were selected. 5) The consistent tissue and target gene results from EpiMap (Table S18). The consolidated effector candidate genes for each BP trait were selected if there were at least 2 additional lines of evidence.

Consolidated effector gene pathway analysis

We used the Gene2Function analysis tool in FUMA (v1.4.0) to perform geneset enrichment and identify significantly associated GO terms and pathways[38]. Hypergeometric tests were performed to test if genes were over-represented in any predefined gene set and multiple testing correction was performed per category. The gene sets used are from MsigDB, WikiPathways and genes from the GWAS-catalog. The analysis included the consolidated effector genes only. The analysis was conducted for all BP traits and we report results with adjusted p-values <0.05.

Redundant GO terms were removed using the Reduce and Visualize Gene Ontology (REVIGO) web application[39]. REVIGO uses a hierarchical clustering method to remove highly similar terms, incorporating enrichment *P*-values in the selection process. Default settings (dispensability cut off <0.7) were used in this analysis.

Druggability of consolidated effector genes

To identify candidate druggable targets, a look-up was performed in a previously published database of the druggable genome developed by Finan et al.[40] This list contains protein-coding genes catagorised into three tiers: Tier 1 are targets of approved drugs and some drugs in clinical development, including targets of small molecules and biotherapeutics; Tier 2 are proteins closely related to drug targets or associated with drug-like compounds (≥50% shared protein sequence identity); Tier 3 includes extracellular proteins and members of key drug target families in Tier 1 (e.g., G protein-coupled receptors). To identify potential opportunities for drug repurposing, a look-up of each BP consolidated effector gene was performed in Tier 1 to identify existing drug targets (https://www.genome.jp/kegg/genes.html). Primary targets of antihypertensives were also identified using the KEGG drug database (https://www.genome.jp/kegg/drug/). The open targets database was subsequently interrogated to identify disease associations with each gene, to identify potential overlap that could indicate promising drug targets. Target, drug and disease association data was downloaded from the platform (https://platform.opentargets.org/downloads). Open targets calculates association scores to capture the data type (e.g., gene level) and source, to aggregate evidence for an association, by calculating the harmonic sum using a weighted vector of data source scores. This sum is divided by the maximum theoretical value, resulting in a score between 0 and 1. To identify enrichment of candidate effector genes in clinical indication categories and potentially re-positional drugs, we utilised the Genome for REPositioning drugs (GREP) software[41]. GREP performs a series of Fisher's exact tests, to identify enrichment of a gene-set in genes targeted by a drug in a clinical indication category (Anatomical Therapeutic Chemical Classification System [ATC] or International Classification of Diseases 10 [ICD10] diagnostic codes).

Results

Identification of BP loci and signals

We considered a total of 650 genomic regions encompassing previously reported lead SNVs for SBP, DBP or PP (**Methods**). Of these, lead SNVs at 606 loci attained genome-wide significance ($P < 5x10^{-8}$) for at least one BP trait and were considered for fine-mapping (Table S1). Through approximate cross-trait conditional analyses (**Methods**), we partitioned BP associations at the 606 genomic regions into a total of 1,850 distinct signals that were associated with at least one BP trait at genome-wide significance (Figure 1, Table S2). Of these signals, 532 were associated with at least two BP traits (333 with SBP and DBP, 267 with SBP and PP, and 100 with DBP and PP), and 84 were associated with all three traits. The only discordancy in direction of effect was for 17 of the 100 signals shared across DBP and PP, where the DBP increasing allele was the PP decreasing allele).

The cross-trait approximate conditional analyses revealed several genomic regions with complex patterns of associations with SBP, DBP and PP (Table S2). For six genomic regions, more than 20 distinct signals of association were observed for at least one BP trait. The most complex associations were observed across: (i) a 6.4Mb region of chromosome 17, encompassing previously reported loci that include *PLCD3*, *GOSR2*, *HOXB7*, *ZNF652*, and *PHB1* (locus ID 576, 37 distinct signals); (ii) a 5.8Mb region of chromosome 10, encompassing previously reported loci that include *PAX2*, *CYP17A1*, *NT5C2*, and *OBFC1* (locus ID 403, 34 distinct signals); and (iii) the major histocompatibility complex region of chromosome 6 (5.7Mb, locus ID 251, 32 distinct signals) that encompasses previously reported loci that include *PRRC2A*, *ABHD16A*, and *HLA-DQB1*.

Fine-mapping and genomic annotation reveals high-confidence causal variants for BP traits

Previous studies have demonstrated that improved localisation of causal variants driving association signals for complex human traits can be attained by integrating GWAS summary statistics with genomic annotation[42]. By mapping SNVs to functional and regulatory annotations from GENCODE[19,43] and the Roadmap Epigenomics Consortium[20] (**Methods**), we observed significant joint enrichment ($P < 2.0 \times 10^{-4}$,

Bonferroni correction for 253 annotations) for BP associations mapping to protein coding exons and 3' UTRs, enhancers in heart and adrenal gland, and promoters in adipose and heart (Table S3, Figure S2).

For each distinct signal, we then derived credible sets of variants that together account for 99% of the posterior probability (π) of driving the BP trait association under an annotation-informed prior model of causality in which SNVs mapping to the genomic annotations in the globally enriched signatures for SBP, DBP and PP are upweighted (**Methods**). The median 99% credible set size was 20 variants for SBP and DBP, and 22 for PP (Table S4 – S6). For 208 (24%), 224 (24.8%) and 159 (22.9%) SBP, DBP and PP signals, respectively, a single SNV accounted for more than 75% of the posterior probability of driving the BP association under the annotation-informed prior, which we defined as "high-confidence" for causality (Figure 2, Table S7- S9).

High-confidence SNVs are enriched for BP-related phenotypes

We used the Genomic Regions Enrichment of Annotations Tool (GREAT) v4.0.4[23] (**Methods**) to explore the potential biological impact of all high confidence SNVs through their enrichment within trait-related genomic regions including *cis*-regulatory elements (CREs). We explored SNVs separately for the three BP traits and physiologically consistent enrichment results were identified for these location data for Gene Ontology Biological Processes (e.g., circulatory system processes, regulation of BP), Human Phenotype (e.g., abnormality systemic blood pressure, abnormality of vasculature), Mouse Phenotype and Knockout data (e.g., abnormal blood vessel morphology, increased systemic arterial blood pressure) (Figure S3 and Table S10-S12).

Missense variants implicate causal candidate genes

We identified 65 high-confidence missense variants for BP association signals (Table S11 and S12). Among these, 20 were driving the same association signal for two BP traits, and one (*RGL3* p.Pro162His) was driving the same association signal for all three BP traits (Table 1). *RGL3* is not well characterised, but several missense variants in the gene have been previously identified in BP EAWAS[12]. In our study, three distinct association signals are driven by *HFE* missense variants; two are common, i.e. minor allele frequency (MAF) ≥5% (*HFE* p.His63Asp and *HFE* p.Cys282Tyr), and one is a

low frequency variant, i.e.MAF <5% (*HFE* p. Ser65Cys).. These variants are associated with predisposition to hereditary hemochromatosis, of which, portal hypertension and restrictive diastolic function are recognised phenotypes[44].

Fifteen missense variants were identified to have a posterior probability of >99.9% of driving distinct BP association signals. These variants implicate several well characterised BP genes (including *SLC39A8* p.Ala391Thr;*ADRB2* p.Gly16Arg and p.Thr164Ile; and *DBH* p.Arg549Cys). The variants p.Thr164Ile in *ADRB2* and p.Arg549Cys in *DBH* are both of low allelic frequency. The results also highlight less well-established candidate genes including *NRIP1, MMP14* and *PLCB3* (the *MMP14* and *PLCB3* missense variants have MAF <5%). *NRIP1* is a regulator of the mineralocorticoid receptor, and *MMP14* is an endopeptidase with a key role in degrading components of the extracellular matrix and regulation of blood vessel stability[45]. EAWAS[12] have identified missense variants associated with BP traits in *PLCB3*, a gene that encodes an enzyme involved in intracellular signal transduction found to be increased in a mouse model of hypertension and hypertrophy[46].

Several high-confidence missense variants implicate genes associated with kidney traits/disorders including *NCOA7* p.Ser399Ala, *LAMB2* p.Ala1765Thr (MAF <5%) and *NPHS2* p.Arg229Gln (MAF <5%). *NCOA7* encodes the nuclear receptor coactivator 7, a vacuolar proton pumping ATPase (V-ATPase) interacting protein. It is highly expressed in the kidney with knock out mice observed to have lower BP[47]. *LAMB2* encodes beta chain isoform laminin, and mutations in this gene cause Pierson syndrome (OMIM# 609049), a congenital nephrotic syndrome, in which the phenotype includes hypertension[48]. Mutations in *NPHS2* cause steroid-resistant nephrotic syndrome[49] and prior work has indicated a rare missense variant association with BP[12].

Non-coding BP association signals map to trait-related transcription factor binding sites

Whilst the identified high-confidence missense variants have directly interpretable effects, the majority of the posterior probability of causality for BP trait associations maps to non-coding sequence. To explore these high-confidence non-coding SNVs, we first sought evidence for enrichment of transcription factor binding site (TFBS)

motifs. We interrogated sets of sequences obtained by expanding 10bp either side of these SNVs for each of the three traits (see **Methods**). This identified significant enrichment for 7 SBP, 10 DBP, and 5 PP TFBS motifs that were partially overlapping (Table S13). The motif for *PAX2* was significant across all three traits (top corrected *P* = 2.8 x 10^{-25} for DBP), with this transcription factor being involved in nephron development, as well as implicated in monogenic renal abnormalities[50].

Effector genes identified using gene expression in BP-relevant tissues

To gain further insight into mechanisms through which non-coding association signals are mediated, where identification of the cognate effector gene is challenging[42], we integrated genetic fine-mapping data with expression quantitative trait loci (*cis*-eQTL) in disease relevant bulk tissues from the GTEx Consortium[27]. The tissues included were adipose, adrenal gland, artery, kidney cortex, heart, nerve, and brain. We observed convincing support for colocalization with eQTLs (**Methods)** for 96 SBP, 107 DBP and 84 PP signals (Table S14). In total, 54 (56%), 58 (54%), and 41 (49%) of the signals colocalized with an eQTL for a single gene in at least one tissue. Across all traits, there was a total of 135 genes with tissue-specific colocalizations, of which 55 (41%) were in arterial tissues, 35 (26%) were in nerve or brain tissues, 21 (16%) were in adipose tissue, 13 (10%) were in heart, 9 (7%) were in adrenal, and 2 (1%) were in kidney.

Of the signals associated with all three BP traits, nine colocalized with an eQTL for a single effector gene. These were: *AGT* (brain cerebellum), *ARHGAP24* (tibial artery and aorta), *ARHGAP42* (tibial artery, aorta, and adipose), *CHD13* (aorta), *lncRNA CTD-2336O2.1* (brain tissues), *FES* (tibial artery), *FGF5* (kidney cortex), *IGFBP3* (heart left ventricle), and *JPH2* (adrenal gland). Three genes (*AGT*, *ARHGAP42* and *IGFBP3*) have known or supporting data for a role in BP regulation. *AGT* encodes angiotensinogen, a substrate of the renin-angiotensinogen system – a key regulatory pathway[51]. *ARHGAP42* is selectively expressed in smooth muscle cells and modulates vascular resistance, and a knockout *Arhgap42* mouse model demonstrates salt mediated hypertension[52]. *IGFBP3*, which encodes the insulin growth factor binding protein 3, has data supporting association with BP and CVD phenotypes, and a knockout mouse model has increased ventricular wall thickness and shortened ST segment[53]. It also modulates insulin growth factor 1 (IGF-1) bioactivity with potential

regulation of vascular tone *in-vivo* through NO release[54]. Additionally, there is a high-confidence missense variant implicating *IGFBP3,* highlighting distinct associations mediated by the same gene but through different underlying biological processes. Other colocalized effector genes demonstrate links to cardiovascular phenotypes (*FES*[55], *FGF5*[56], and *JPH2*[57]) but have not yet been functionally characterised but demonstrate links to cardiovascular phenotypes.

We observed many individual loci with several distinct signals for each BP trait that colocalized with eQTLs for different genes. The genomic region on chromosome 12 encompassing *HDAC7*, *H1-7*, *CCDC65*, *PRKAG1*, *FAM186B*, *CERS5*, and *DIP2B*, spans 3.5Mb, and includes 11 signals for SBP, 12 for DBP and 5 for PP. Three signals colocalized with eQTLs and indicate two effector genes. One signal (associated with both SBP and DBP) colocalized with an eQTL for *CACNB3* (adipose, tibial nerve and artery), which encodes a regulatory beta subunit of the voltage-dependent calcium channel. The regulatory subunit of the voltage-gated calcium channel gives rise to L-type calcium currents[58]. A *CACNB3* knock-out mouse model has a cardiovascular phenotype that includes abnormal vascular smooth muscle cell hypertrophy, increased heart weight and increased SBP and DBP[59]. A second signal (associated with DBP) colocalized with an eQTL for *lncRNA RP4-605O3.4* (heart left ventricle), and a third signal (associated with SBP) colocalized with an eQTL in brain and heart left ventricle for this predicted gene.

A genomic region on chromosome 17 spanning 6.4Mb, which encompasses associations reported in several previous BP GWAS[8,12,60,61], includes 19 SBP, 16 DBP and 15 PP signals (Locus ID 576, Table S2). Colocalization of signals with eQTLs implicates six effector genes (*DCAKD*, *NMT1*, *lncRNA RP11-6N17.4*, *PNPO*, *PRR15L* and *ZNF652*). Three independent signals colocalized with eQTLs for *NMT1* in brain. The *NMT1* gene encodes N-myristoyltransferase, which catalyses the transfer of myristate from CoA to proteins, and there is no clear association with cardiovascular disease. However, the MalaCards database indicates an association with Patent Foramen Ovale, a common post-natal defect of cardiac atrial septation[62]. One DBP signal colocalized with an eQTL for *DCAKD* in adipose and nerve tissues. PP signals colocalized with eQTLs for *RP11-6N17.4* and *PNPO* in brain tissues. *PNPO* encodes pyridoxamine 5'-phosphate oxidase, an enzyme in the rate limiting step in vitamin B6

synthesis. Deficiency of PNPO primarily results in seizures, with many systemic symptoms including cardiac abnormalities[63]. We also observed a SBP signal that colocalized with an eQTL for *PRR15L* in tibial artery, and a signal associated with both with SBP and DBP that colocalized with an eQTL for *ZNF652* in adipose tissue.

At a second genomic region on chromosome 17 encompassing *MRC2*, *ACE*, *PECAM1*, and *MILR1*, we observed four signals for SBP, three for DBP and four for PP (Locus ID 580, Table S2), of which three signals colocalized with different genes across multiple tissues (Figure 3). One SBP signal colocalizes with an eQTL for *MRC2* in tibial artery. *MRC2* encodes the mannose receptor C type 2 and plays a role in extracellular matrix remodelling[64]. A signal associated with both SBP and DBP colocalized with an eQTL for *ACE* in kidney, adipose, and brain tissues. *ACE* encodes the angiotensin-converting enzyme, a central component of the renin–angiotensin–aldosterone system[65]. A third SBP signal colocalized with an eQTL for two genes across several tissues: *DDX5* (arteries, brain and tibial nerves) and *CEP95* (tibial nerve and arteries). These genes have little prior association with cardiovascular phenotypes. *DDX5* encodes DEAD-Box Helicase 5, which is thought to be a coregulator of transcription or splicing, and recent data indicates a role in smooth muscle cell protection and neointimal hyperplasia[66]. Homozygous *Ddx5* knockout mice die at embryonic day 11.5 and demonstrate blood vessel abnormalities. There is little information on *CEP95*, which encodes centrosomal protein 95, although differential gene expression was observed in spontaneously hypertensive rats[67].

Exploratory analysis of single cell datasets for BP effector genes

We performed an exploratory analysis investigating cell type specificity for effector genes indicated from eQTL analysis utilising single cell datasets from the adrenal, kidney, heart, brain and adipose tissues (**Methods**). The results per effector gene per tissue are provided in Figure S4. There were several genes across each tissue which had a relative expression >75% in a particular cell type compared to other cells in that organ, thus demonstrating potentially strong cell-type specific expression. The genes indicated were: *lncRNA RP11-179B2.2* in neurons in the brain hippocampus, *HSPB7*, *CLCNKA*, *ACE* in neurons in the brain cortex, *PCOLCE-ASI* in neurons in the brain cerebellum (as opposed to other non-neuronal cell types present in cortical tissue,

such as glia); *lncRNA RP11-373D23.3* in fibroadipogenic progenitor cell, *PAQR8* in lymphocyte, *ACHE* in mesothelial cells all from adipose tissues; *ACHE* in fibroadipogenic progenitor cell in adipose visceral omentum; and *JPH2*, *FHL3*, *lncRNA CTB-30L5.1* in cortical cells of the adrenal gland. There were no genes in the heart, kidney, or brain substantia nigra where cell-type specific expression exceeded the 75% threshold we selected.

Identification of effector genes using promotor-centered long-range chromatin interactions in disease relevant tissues

To explore possible long-range enhancer influence on specific target genes, we integrated genetic fine-mapping data with potential functional CREs identified to target the promoters of well-annotated protein-coding genes via long-range chromatin interactions (capture Hi-C data from Jung et al.[34]). Promoter interactions and candidate genes were identified for 629 signals at 366 genomic regions (RegulomeDB Score ≤ 3) across adrenal gland, dorsolateral prefrontal cortex, hippocampus, aorta, left ventricle, right ventricle, and fat (Table S15). We observed signals at 13 genomic regions that included 99% credible set variants with regulatory potential across SBP and DBP, for which several potential target genes of the regulatory variants were indicated. At five signals, one gene was indicated in a single tissue: *ACTRT2* (dorsolateral prefrontal cortex), *ARMC4* (right ventricle), *ncRNA AP001024.1* (hippocampus), *TBX3* (aorta) and *YES1* (hippocampus). At other genomic regions, many genes in one tissue were indicated: *HOXA5*, *HOXA6* and *HOXA3* (adrenal gland), *ncRNA RP11691N7.6*, *SELENOH*, *TIMM10*, *CLP1*, *YPEL4*, *ZDHHC5*, *FAM111A*, *lncRNA AP001350.1*, *GLYATL2*, *GLYATL1* (hippocampus, dorsolateral prefrontal cortex), and *ABHD17C*, *MESD* (two brain tissues). At three signals, more than one gene in more than one tissue were highlighted, such as *ADGRA2*, , *DDHD2*, *FGFR1*, *PLPP5*, , *LETM2*, *TACC1* (two brain tissues, adrenal gland, aorta, and right ventricle). Candidate genes at three signals have existing functional data supporting an association with BP or cardiovascular traits: *HOXA3*, *ADM* and *TBX3*[53,68,69].

We next explored whether signals that colocalized with eQTLs for effector genes overlapped with those implicated by Hi-C predicted promoter interactions. We focused on the 80 signals that have support for colocalization with eQTLs in relevant tissues

across all traits (Table S16). For 34 signals, the effector gene indicated by Hi-C was the same as that identified via colocalization with the eQTLs, and for 15 of these signals the same tissue was implicated (Table S16). The 15 candidate genes were: *AKR1B1*, *ASAP2, COL27A1, IRF5, MAP1B, MRPS6, MXD3, RAD52, RERE, RNF130, SLC5A3*, *SLC20A2, TNS3, TRIOBP*, and *USP36*. A review of the 15 genes indicates knock-out mouse models of three effector genes (*COL27A1, RERE*, and *SLC20A2*) have cardiovascular abnormalities, but these genes have not previously been highlighted as potential candidate genes for hypertension (Table S17).

To explore our Hi-C predicted promoter interactions more broadly, we additionally probed our results to see whether there was also support for these potential CREs to target the same effector gene through a completely different prediction methodology from the recent EpiMap analysis[70]. This method is based on an active chromatin correlation with target gene expression. Several physiologically relevant candidate effector transcripts were highlighted, where the two methods predicted the same target for the same SNV in the same tissue or organ (See Table S18). Potential targets included genes previously identified as highly plausible trait-related candidates from previous analyses, including *CLIC4*, *TNS1*, and *FERMT2*[71]. Target genes with presently unknown potential roles in SBP and DBP pathophysiology were also identified. These included two active CREs found in brain-related tissue, which would be of interest to explore for additional activity in potentially more physiologically relevant non-assayed tissues: *SHMT1*, the serine hydroxymethyltransferase 1 enzyme involved in folic acid metabolism associated, although inconsistently, with hypertension-related stroke[72] and *PLXNB2*, the Plexin-B2 transmembrane receptor that has an identified role in the developing kidney[73]. For PP this comprised, amongst others, some interesting target genes including *MYH11*, the smooth muscle myosin heavy chain 11 gene, with CRE activity in aortic tissue. Mutations within this gene lead to an autosomal dominant aortic aneurysm and dissection disorder (*AAT4*, OMIM #132900) with altered aortic stiffness[74,75]. Also, *COL6A3*, which encodes the alpha-3 chain of type VI collagen, was a target identified in heart tissue. This collagen gene is important in the developing mammalian heart[76] and is a causative gene in monogenic myopathy and dystonia diseases (OMIM: #158810; #254090; #616411)[77].

Effector genes identified using pQTLs in plasma

We integrated genetic fine-mapping data with cis-pQTL using summary statistics from a published study of plasma protein concentrations in 35,559 Icelanders[36]. Support for colocalization was identified for 10 SBP, 6 DBP and 11 PP signals corresponding to 10, 6 and 10 unique proteins, respectively. Across all BP traits, there were 16 unique signals at 16 proteins (Table S19). For two proteins (angiotensin and matrix-remodelling associated protein 7 [*MXRAP7*]), the encoding gene was also a significant eQTL finding. For four proteins (Tyrosine-protein kinase Fgr, Collagen alpha-3 (VI) chain, Plexin-B2 and *MXRAP7*), the encoding gene was a significant finding in the long-range chromatin interaction analyses.

Consolidated effector gene evidence

Using complementary fine mapping and computational approaches (high-confidence missense, colocalised eQTLs, Hi-C interactions and pQTLs) we identified 959 candidate genes for SBP, 904 candidate genes for DBP and 774 candidate genes for PP with at least one line of evidence (Tables S20 – S22).

We next looked for additional supportive evidence for each of these candidate genes by combining information from i) mouse, and ii) human cardiovascular and renal phenotypes, iii) the consistent target and tissue EpiMap findings, and differential iv) gene and v) protein abundance across cardiovascular tissues (Methods). We selected as consolidated effector genes those which had two or more additional lines of evidence. In total, 215 SBP, 205 DBP and 202 PP genes were identified (Tables S23-S25), which together reflect 436 unique genes.

To gain insights into the biological role of the consolidated effector genes for each BP trait we performed gene-set enrichment analyses. We found significant enrichment for 310, 270, and 245 gene ontology (GO) biological processes for SBP, DBP, and PP respectively (following removal of redundant processes, see **Methods**). There were 264 unique GO ID terms across the three traits, with 111, 77 and 76 unique to SBP, DBP and PP respectively. In total, 172 pathways were associated with SBP and DBP, 148 with SBP and PP, and 140 with DBP and PP. Moreover, 119 pathways were associated with all three BP traits (Tables S26 – S28). Some of the pathways associated with all three BP traits included: circulatory system development, embryo development, tube development, regulation of cell differentiation, urogenital system

development and renal system development – all processes previously highlighted as important in BP control. The most significant SBP unique processes included: heart development (P = 4.22 x 10^{-12}), positive regulation of signaling (P = 8.77 x 10^{-10}) and positive regulation of gene expression (P = 1.95 x 10^{-9}); for DBP, epithelium development (P = 6.45 x 10^{-14}), embryonic organ development (P = 3.14 x 10^{-11}) and epithelial cell proliferation (P = 3.08 x 10^{-7}), for PP, tube morphogenesis (P = 9.57 x 10^{-18}), muscle cell proliferation (P = 1.3 x 10^{-11}) and response to growth factor (P = 1.67 x 10^{-7}).

Drug target identification and repositioning opportunities

We assessed the druggability of the consolidated effector genes for each BP trait via the druggable genome dataset from Finan et al.[40] (Table S29, see **Methods**). We observed DBP to have a greater number of candidate effector genes that encode proteins that are the main drug targets for anti-hypertensive medications (*ACE*, *ADRA1A*, *ADRB1* and *NR3C2*), compared with SBP (*ACE*) and PP (none). For several effector genes that are targets of existing drugs, there was support as potential therapeutic targets for hypertension (e.g., *AKR1B1, PDE3A* and *MAP2K1*). *AKR1B1* (aldo-keto reductase family 1 member B) is a target of aldose reductase inhibitors that have been investigated for use in diabetes and also have effects on BP[78]. *PDE3A* (phosphodiesterase 3A) is a target for hypertension with bracydactyly, a rare autosomal dominant disorder and there is recent data indicating several common variant associations also in the general population[12,79]. *PDE3A* is targeted by several existing drugs, including Cilostazol (peripheral vascular disease), Levosimendan for intravenous therapy for acutely decompensated heart failure and Enoximone (pulmonary hypertension). There are no data currently indicating the use of *PDE3A* inhibitors for hypertension, however a recent study suggests activation of *PDE3A* in the heart may protect it from hypertrophy and failure[80]. *MAP2K1* is a target of anti-neoplastic agents (*MAP2K1* is altered in 1% of lung and head and neck squamous cell carcinomas)[81,82].The *MAPK* pathway is well recognised in BP control and p38-MAPK inhibition has been considered previously as a therapeutic target (which *MAP2K* activates).

To further ascertain drug repositioning opportunities, we tested for enrichment of these consolidated effector genes in clinical indication categories. We observed significant

enrichment of gene sets for cardiovascular and renal conditions (Table S30), with the results supporting the findings from interrogation of the Finan et al. druggable genome database.

Discussion

Strongly replicated human genetic associations with BP traits have been identified over the last decade but there is a lag in effector gene identification. In this study, we have used a robust contemporary fine mapping pipeline to advance from these initial broadly associated genomic regions to the identification of hundreds of candidate previously-unreported effector genes. These consolidated candidates are now excellent targets for future focused functional validation.

We were able to localise approximately a quarter of all associations across all three BP traits to a single causal variant with >75% posterior probability. Of these high-confidence SNVs, 65 were missense variants, including 20 identified for two BP traits, and one in *RGL3* for all three traits. For the high confidence non-coding and potentially *cis*-regulatory variants, we employed pathogenic tissue-specific expression and chromatin conformation to identify their target genes. Of these SNVs, ~100 per trait colocalized with *cis*-eQTLs. Plausible effector genes included the well-known Angiotensin (*AGT*) and angiotensin converting enzyme (*ACE*), also more recently described genes from GWAS with functional data including sodium/potassium-transporting ATPase subunit beta-1 (*ATP1B1*) and Rho GTPase activating protein 42 (*ARHGAP42*). Other possible but less well functionally evaluated genes identified through eQTL analysis included *CDH13*, *FES*, *FGF5*, and *JPH2*.

We identified many loci with multiple complex signals within the same genomic region affecting different genes in different tissue types. Also, of note, whilst we observed high-confidence missense variants in kidney genes, as well as an enrichment for non-coding variants to overlap a nephron developmental TFBS, we identified only a very small proportion of eQTL colocalizations in this tissue (*FGF5* and *ACE* and the lncRNA *AC021218.2*). This may reflect reduced power due to the relatively smaller numbers in GTEx for kidney than other tissues[27]. Using over 400 human kidneys and the same

ICBP+UKBB GWAS dataset, Eales and colleagues reported nearly 31% of BP associated variants contained kidney eSNPs[83]. We were not able to directly compare or incorporate results from this larger kidney dataset into our pipeline as summary statistics are not available for colocalization analysis. With the big difference in kidney eQTLs between datasets, the results strongly emphasise the importance of access to larger tissue banks for robust identification of all possible effector genes.

Focusing on the overlap between the eQTL and Hi-C results in disease relevant tissues, there was a subset of 15 target genes identified in the same tissue. Of these, the genes *COL27A1*, *RERE*, and *SLC20A2* also had additional evidence from mouse data. Support for target genes consistently predicted across multiple methods are the most robust[84]. We also explored overlap with the recent EpiMap[70], highlighting, amongst others, *MYH11* and *COL6A3* as strong effector candidates for PP. In total, our pipeline identified consolidated evidence effector genes for ~25% of BP association signals (215 genes for 865 SBP signals; 205 genes for 904 DBP signals; and 202 genes for 697 PP signals). Of the consolidated BP genes, 13% were identified to be drug targets, and several of these have good support for potential repurposing for BP control.

Our study applied rigorous multiple evidence methodology employing differing datasets from the Evangelou et al. to derive a consolidated effector gene list.[8] This impacted the identification of potential therapeutic targets as we were not able to replicate in our analysis the five candidate genes (*CA7, CACNA1C*, *CACNB4, PKD2L1* and *SLC12A2*) reported in their paper as the target of anti-hypertensive drug classes. However, *PKD2L1* did show individual Hi-C evidence as bring the effector gene at the locus.

The main strength of this work is that it combines robust GWAS associations, derived from a powerfully large dataset, with comprehensive genetic annotation and tissue-specific epigenomic maps derived from the Epigenomics Roadmap consortium. In the exploration of the putative functional non-coding variants, a further strength was that we were able to benefit from the expanded GTEx dataset[27], publicly available promoter capture Hi-C data in pathogenically relevant tissues[34], as well as exploration of target gene prediction overlap with EpiMap[70]. These analyses identified many biologically

plausible effector genes. Our work builds and advances on the initial Evangelou et al. findings, as here we perform formal fine-mapping analysis, as well as eQTL colocalisation. Furthermore, unlike previous work, our integrative analysis has enabled us to delineate a distinct list of potential effector genes.

Current weaknesses are the lack of population diversity in our GWAS dataset, as these are comprised of associations from European ancestry individuals only. Consequently, our results will be missing population-specific findings, as have been identified in other common diseases[85]. Whilst multi-ancestry BP sequencing studies have indicated the strong ancestry specific nature of rare variants[86], most common variants may be shared[87]. However, due to LD and MAF complexity, identified variants will need detailed exploration at a locus-by-locus basis across continental groups[88]. Furthermore, this lack of diversity is not only limited to the genetic findings. The epigenomic maps, whilst being derived from a breadth of cell-types giving good representation of strong tissue-specific regulatory differences, are within each cell-type drawn from very small numbers. Therefore, they lack detail regarding potential population variation in these functional units[89]. Another weakness is that whilst benefitting from dense genotyping and imputation of common SNVs, this is not exhaustive in capturing all the potential phenotypically associated genetic variation within each locus. This will miss the possible impact of rare SNVs, as well as any poorly tagged larger variants (copy number variants, short tandem repeats, inversions, etc.). Furthermore, these large variants may themselves facilitate functional epigenomic variation[90]. Future exploration of the phased interplay of genetic and epigenetic allelic elements by advancing long-read technologies will help to fill in these gaps[91]. We did make use of currently available single cell expression data sets to explore higher resolution cell-type specificity, which is not able to be resolved with bulk tissue analyses. However, whilst this revealed some intriguing findings, which would be of interest to explore further experimentally, we also acknowledge that these resources currently have technical and power constraints, but are rapidly evolving, and will become more comprehensive over time.

In conclusion, we have identified plausible causal genetic variants and effector genes enriched in BP pathways. Their investigation through experimental biosystems will not

only improve functional understanding of the biology of BP and its pathogenesis, but also potentially enable novel preventative and therapeutic opportunities.

Figure Legends

Figure 1. Overlap of 1,850 distinct signals attaining genome-wide significant evidence of association with SBP, DBP and PP in meta-analysis of BP GWAS in up to 757,601 individuals of European ancestry.

(a) Venn diagram showing the number of signals shared across BP traits. Sharing of signals across traits is much more common between SBP and DBP or SBP and PP, with just 16 associations shared between only DBP and PP. (b) Comparison of allelic effect sizes on SBP, DBP and PP for the index SNV at the 532 distinct association signals that are shared across multiple BP traits. The effect has been aligned to the SBP or PP increasing allele for the signal. Blue points correspond to the 448 association signals that are shared across exactly two BP traits, whilst red points correspond to the 84 association signals that are shared across all three BP traits. When signals are shared between SBP and PP, the direction of effect of the index SNV on the traits is always concordant.

Figure 2. Distinct BP association signals.

(a) Summary of distinct association signals for blood pressure traits. SBP: A single signal at 277 genomic regions and at least two at 180; DBP: A single signal at 262 genomic regions, and at least two at 188; PP: A single signal at 265 genomic regions, and at least two at 144. (b) Distribution of the posterior probability of causality of the variants in credible sets. SBP, systolic blood pressure; DBP, diastolic blood pressure; PP, pulse pressure.

Figure. 3. Colocalization between GWAS signals for systolic blood pressure and multi-tissue expression data at locus ID 580 on chromosome 17.

The top panel shows the unconditional GWAS data of the genomic region at chromosome 17 (60.2Mb – 62.9Mb) for systolic blood pressure. The lower four panels show the log annotation informed Bayesian factors of the conditional GWAS signal (blue, left axis) and gene expression data from GTEx eQTL data (red, right axis). Three distinct annotation informed signals colocalized with gene expression data from: *MRC2* (second panel) in tibial artery tissue, *ACE* (third panel) in kidney cortex tissue, and *CEP95* and *DDX5* in aortic artery tissue (bottom panels).

The x-axis shows the physical position on the chromosome (Mb) and the y-axes show the log annotation informed Bayesian factor from the GWAS (left axis) and the gene expression data (right axis). The intensity of the color indicates the linkage disequilibrium with respect to the sentinel GWAS SNP (blue) or top eQTL SNP (red).

Consortia

International Consortium of Blood Pressure (ICBP) members.

Evangelos Evangelou, Helen R Warren, He Gao, Georgios Ntritsos, Niki Dimou, Tonu Esko, Reedik Mägi, Lili Milani, Peter Almgren, Thibaud Boutin, Stéphanie Debette, Jun Ding, Franco Giulianini, Elizabeth G Holliday, Anne U Jackson, Ruifang Li-Gao, Wei-Yu Lin, Jian'an Luan, Massimo Mangino, Christopher Oldmeadow, Bram Peter Prins, Yong Qian, Muralidharan Sargurupremraj, Nabi Shah, Praveen Surendran, Sébastien Thériault, Niek Verweij, Sara M Willems, Jing-Hua Zhao, Philippe Amouyel, John Connell, Renée de Mutsert, Alex SF Doney, Martin Farrall, Cristina Menni, Andrew D Morris, Raymond Noordam, Guillaume Paré, Neil R Poulter, Denis C Shields, Alice Stanton, Simon Thom, Gonçalo Abecasis, Najaf Amin, Dan E Arking, Kristin L Ayers, Caterina M Barbieri, Chiara Batini, Joshua C Bis, Tineka Blake, Murielle Bochud, Michael Boehnke, Eric Boerwinkle, Dorret I Boomsma, Erwin P Bottinger, Peter S Braund, Marco Brumat, Archie Campbell, Harry Campbell, Aravinda Chakravarti, John C Chambers, Ganesh Chauhan, Marina Ciullo, Massimiliano Cocca, Francis Collins, Heather J Cordell, Gail Davies, Martin H de Borst, Eco J de Geus, Ian J Deary, Joris Deelen, Fabiola Del Greco M, Cumhur Yusuf Demirkale, Marcus Dörr, Georg B Ehret, Roberto Elosua, Stefan Enroth, A Mesut Erzurumluoglu, Teresa Ferreira, Mattias Frånberg, Oscar H Franco, Ilaria Gandin, Paolo Gasparini, Vilmantas Giedraitis, Christian Gieger, Giorgia Girotto, Anuj Goel, Alan J Gow, Vilmundur Gudnason, Xiuqing Guo,Ulf Gyllensten, Anders Hamsten, Tamara B Harris, Sarah E Harris, Catharina A Hartman, Aki S Havulinna, Andrew A Hicks, Edith Hofer, Albert Hofman, Jouke-Jan Hottenga, Jennifer E Huffman, Shih-Jen Hwang, Erik Ingelsson, Alan James, Rick Jansen, Marjo-Riitta Jarvelin, Roby Joehanes, Åsa Johansson, Andrew D Johnson, Peter K Joshi, Pekka Jousilahti, J Wouter Jukema, Antti Jula, Mika Kähönen, Sekar Kathiresan, Bernard D Keavney, Kay-Tee Khaw, Paul Knekt, Joanne Knight, Ivana Kolcic, Jaspal S Kooner, Seppo Koskinen, Kati Kristiansson, Zoltan Kutalik, Maris Laan, Marty Larson, Lenore J Launer, Benjamin Lehne, Terho Lehtimäki, David CM Liewald, Li Lin, Lars Lind, Cecilia M Lindgren, YongMei Liu, Ruth JF Loos, Lorna M Lopez, Yingchang Lu, Leo-Pekka Lyytikäinen, Anubha Mahajan, Chrysovalanto Mamasoula, Jaume Marrugat, Jonathan Marten, Yuri Milaneschi, Anna Morgan, Andrew P Morris, Alanna C Morrison, Peter J Munson, Mike A Nalls, Priyanka

Nandakumar, Christopher P Nelson, Teemu Niiranen, Ilja M Nolte, Teresa Nutile, Albertine J Oldehinke, Ben A Oostra, Paul F O'Reilly, Elin Org, Sandosh Padmanabhan, Walter Palmas, Aarno Palotie, Alison Patti, Brenda WJH Penninx, Markus Perola, Annette Peters, Ozren Polasek, Peter P Pramstaller, Quang Tri Nguyen, Olli T Raitakari, Rainer Rettig, Kenneth Rice, Paul M Ridker, Janina S Ried, Harriëtte Riese, Samuli Ripatti, Antonietta Robino, Lynda M Rose, Jerome I Rotter, Igor Rudan, Daniela Ruggiero, Yasaman Saba, Cinzia F Sala, Veikko Salomaa, Nilesh J Samani Antti-Pekka Sarin, Reinhold Schmidt, Helena Schmidt, Nick Shrine, David Siscovick, Albert V Smith, Harold Snieder, Siim Sõber, Rossella Sorice, John M Starr, David J Stott, David P Strachan, Rona J Strawbridge, Johan Sundström, Morris A Swertz, Kent D Taylor, Alexander Teumer, Martin D Tobin, Maciej Tomaszewski, Daniela Toniolo, Michela Traglia, Stella Trompet, Jaakko Tuomilehto, Christophe Tzourio, André G Uitterlinden, Ahmad Vaez, Peter J van der Most, Cornelia M van Duijn, Germaine C Verwoert, Veronique Vitart, Uwe Völker, Peter Vollenweider Dragana Vuckovic, Hugh Watkins, Sarah H Wild, Gonneke Willemsen, James F Wilson, Alan F Wright, Jie Yao, Tatijana Zemunik, Weihua Zhang, John R Attia, Adam S Butterworth, Daniel I Chasman, David Conen, Francesco Cucca, John Danesh, Caroline Hayward, Joanna MM Howson, Markku Laakso, Edward G Lakatta, Claudia Langenberg, Olle Melander, Dennis O Mook-Kanamori, Colin NA Palmer, Lorenz Risch, Robert A Scott, Rodney J Scott, Peter Sever, Tim D Spector, Pim van der Harst, Nicholas J Wareham, Eleftheria Zeggini, Daniel Levy, Patricia B Munroe, Christopher Newton-Cheh, Morris J Brown, Andres Metspalu, Bruce M. Psaty, Louise V Wain, Paul Elliott, Mark J Caulfield

Declarations of interest

The authors declare no competing interests.

Acknowledgements

This research was supported by the NIHR Barts Cardiovascular Research Centre (NIHR203330) and the NIHR Manchester Biomedical Research Centre (NIHR203308). J.R. acknowledges funding from the European Union's Horizon 2020 Research and Innovation Programme under the Marie Sklodowska-Curie grant agreement number 786833, from the European Union-NextGenerationEU, and

from Grant PID2021-128972OA-I00 funded by MCIN/AEI/ 10.13039/501100011033. W.J.Y. recognises the National Institute for Health Research (NIHR) Integrated Academic Training programme, which supports his Academic Clinical Lectureship post. Graphical Abstract created with Biorender.com and *EpiMap image adapted from Boix et al. [70]

Author Contributions

S.vD., J.R., W.J.Y., C.G.B., A.P.M., P.B.M designed the study. S.vD., J.R., W.J.Y., K.J.O., F.A., M.J.A.Y.A., C.G.B.., A.P.M., P.B.M. performed analyses. S.vD., J.R., W.J.Y., C.G.B., A.P.M., P.B.M. drafted the manuscript, and all authors provided critical revisions.

Web resources

CellxGene: https://cellxgene.cziscience.com/datasets
Epigenomics Roadmap Chromatin Segmentation: https://egg2.wustl.edu/roadmap/data/byFileType/chromhmmSegmentations/ChmmModels/coreMarks/jointModel/final/
Epilogos: https://epilogos.altius.org/
EpiMap: https://cboix.shinyapps.io/epimap_vis/
FUMA: https://fuma.ctglab.nl/
GeneCards: https://genealacart.genecards.org
GnomAD: https://gnomad.broadinstitute.org
GTEx project (v8) portal: https://gtexportal.org/home/
KEGG database: https://www.genome.jp/kegg
OMIM: https://www.omim.org/
Open Targets: https://genetics.opentargets.org/
Promoter Capture Hi-C: http://www.3div.kr
VEP: https://grch37.ensembl.org/Homo_sapiens/Tools/VEP)

Data and code Availability

The published article includes all datasets generated or analyzed during this study.

References

1. Forouzanfar, M.H., Liu, P., Roth, G.A., Ng, M., Biryukov, S., Marczak, L., Alexander, L., Estep, K., Hassen Abate, K., Akinyemiju, T.F., et al. (2017). Global Burden of Hypertension and Systolic Blood Pressure of at Least 110 to 115 mm Hg, 1990-2015. JAMA *317*, 165-182. 10.1001/jama.2016.19043.
2. G. B. D. Risk Factors Collaborators (2020). Global burden of 87 risk factors in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet *396*, 1223-1249. 10.1016/S0140-6736(20)30752-2.
3. Kolifarhood, G., Daneshpour, M., Hadaegh, F., Sabour, S., Mozafar Saadati, H., Akbar Haghdoust, A., Akbarzadeh, M., Sedaghati-Khayat, B., and Khosravi, N. (2019). Heritability of blood pressure traits in diverse populations: a systematic review and meta-analysis. J Hum Hypertens *33*, 775-785. 10.1038/s41371-019-0253-4.
4. Sudlow, C., Gallacher, J., Allen, N., Beral, V., Burton, P., Danesh, J., Downey, P., Elliott, P., Green, J., Landray, M., et al. (2015). UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age. PLOS Medicine *12*, e1001779. 10.1371/journal.pmed.1001779.
5. Gaziano, J.M., Concato, J., Brophy, M., Fiore, L., Pyarajan, S., Breeling, J., Whitbourne, S., Deen, J., Shannon, C., Humphries, D., et al. (2016). Million Veteran Program: A mega-biobank to study genetic influences on health and disease. J Clin Epidemiol *70*, 214-223. 10.1016/j.jclinepi.2015.09.016.
6. Kanai, M., Akiyama, M., Takahashi, A., Matoba, N., Momozawa, Y., Ikeda, M., Iwata, N., Ikegawa, S., Hirata, M., Matsuda, K., et al. (2018). Genetic analysis of quantitative traits in the Japanese population links cell types to complex human diseases. Nature Genetics *50*, 390-400. 10.1038/s41588-018-0047-6.
7. Kato, N., Takeuchi, F., Tabara, Y., Kelly, T.N., Go, M.J., Sim, X., Tay, W.T., Chen, C.-H., Zhang, Y., Yamamoto, K., et al. (2011). Meta-analysis of genome-wide association studies identifies common variants associated with blood pressure variation in east Asians. Nature Genetics *43*, 531-538. 10.1038/ng.834.
8. Evangelou, E., Warren, H.R., Mosen-Ansorena, D., Mifsud, B., Pazoki, R., Gao, H., Ntritsos, G., Dimou, N., Cabrera, C.P., Karaman, I., et al. (2018). Genetic analysis of over 1 million people identifies 535 new loci associated with blood pressure traits. Nat Genet *50*, 1412-1425. 10.1038/s41588-018-0205-x.
9. Giri, A., Hellwege, J.N., Keaton, J.M., Park, J., Qiu, C., Warren, H.R., Torstenson, E.S., Kovesdy, C.P., Sun, Y.V., Wilson, O.D., et al. (2019). Trans-ethnic association study of blood pressure determinants in over 750,000 individuals. Nat Genet *51*, 51-62. 10.1038/s41588-018-0303-9.
10. Sakaue, S., Kanai, M., Tanigawa, Y., Karjalainen, J., Kurki, M., Koshiba, S., Narita, A., Konuma, T., Yamamoto, K., Akiyama, M., et al. (2021). A cross-population atlas of genetic associations for 220 human phenotypes. Nat Genet *53*, 1415-1424. 10.1038/s41588-021-00931-x.
11. Ehret, G.B., Ferreira, T., Chasman, D.I., Jackson, A.U., Schmidt, E.M., Johnson, T., Thorleifsson, G., Luan, J., Donnelly, L.A., Kanoni, S., et al. (2016). The genetics of blood pressure regulation and its target organs from association studies in 342,415 individuals. Nat Genet *48*, 1171-1184. 10.1038/ng.3667.

12. Surendran, P., Feofanova, E.V., Lahrouchi, N., Ntalla, I., Karthikeyan, S., Cook, J., Chen, L., Mifsud, B., Yao, C., Kraja, A.T., et al. (2020). Discovery of rare variants associated with blood pressure regulation through meta-analysis of 1.3 million individuals. Nat Genet *52*, 1314-1332. 10.1038/s41588-020-00713-x.
13. Magavern, E.F., Warren, H.R., Ng, F.L., Cabrera, C.P., Munroe, P.B., and Caulfield, M.J. (2021). An Academic Clinician's Road Map to Hypertension Genomics: Recent Advances and Future Directions MMXX. Hypertension *77*, 284-295. 10.1161/HYPERTENSIONAHA.120.14535.
14. The 1000 Genomes Project Consortium, Abecasis, G.R., Auton, A., Brooks, L.D., DePristo, M.A., Durbin, R.M., Handsaker, R.E., Kang, H.M., Marth, G.T., and McVean, G.A. (2012). An integrated map of genetic variation from 1,092 human genomes. Nature *491*, 56-65. 10.1038/nature11632.
15. The 1000 Genomes Project Consortium (2015). A global reference for human genetic variation. Nature *526*, 68-74. 10.1038/nature15393.
16. McCarthy, S., Das, S., Kretzschmar, W., Delaneau, O., Wood, A.R., Teumer, A., Kang, H.M., Fuchsberger, C., Danecek, P., Sharp, K., et al. (2016). A reference panel of 64,976 haplotypes for genotype imputation. Nat Genet *48*, 1279-1283. 10.1038/ng.3643.
17. Yang, J., Ferreira, T., Morris, A.P., Medland, S.E., Genetic Investigation of, A.T.C., Replication, D.I.G., Meta-analysis, C., Madden, P.A., Heath, A.C., Martin, N.G., et al. (2012). Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. Nat Genet *44*, 369-375, S361-363. 10.1038/ng.2213.
18. Pickrell, J.K. (2014). Joint analysis of functional genomic data and genome-wide association studies of 18 human traits. Am J Hum Genet *94*, 559-573. 10.1016/j.ajhg.2014.03.004.
19. Harrow, J., Frankish, A., Gonzalez, J.M., Tapanari, E., Diekhans, M., and Kokocinski, F. (2012). GENCODE: the reference human genome annotation for The ENCODE Project. Genome Res. *22*. 10.1101/gr.135350.111.
20. Roadmap Epigenomics Consortium, Kundaje, A., Meuleman, W., Ernst, J., Bilenky, M., Yen, A., Heravi-Moussavi, A., Kheradpour, P., Zhang, Z., Wang, J., et al. (2015). Integrative analysis of 111 reference human epigenomes. Nature *518*, 317-330. 10.1038/nature14248.
21. Kass, R.E., and Raftery, A.E. (1995). Bayes Factors. Journal of the American Statistical Association *90*, 773-795. 10.1080/01621459.1995.10476572.
22. Maller, J.B., McVean, G., Byrnes, J., Vukcevic, D., Palin, K., Su, Z., Howson, J.M.M., Auton, A., Myers, S., Morris, A., et al. (2012). Bayesian refinement of association signals for 14 loci in 3 common diseases. Nature Genetics *44*, 1294-1301. 10.1038/ng.2435.
23. McLean, C.Y., Bristor, D., Hiller, M., Clarke, S.L., Schaar, B.T., Lowe, C.B., Wenger, A.M., and Bejerano, G. (2010). GREAT improves functional interpretation of cis-regulatory regions. Nature biotechnology *28*, 495-501. 10.1038/nbt.1630.
24. McLaren, W., Pritchard, B., Rios, D., Chen, Y., Flicek, P., and Cunningham, F. (2010). Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor. Bioinformatics *26*, 2069-2070. 10.1093/bioinformatics/btq330.
25. Thomas-Chollier, M., Hufton, A., Heinig, M., O'Keeffe, S., Masri, N.E., Roider, H.G., Manke, T., and Vingron, M. (2011). Transcription factor binding

predictions using TRAP for the analysis of ChIP-seq data and regulatory SNPs. Nat Protoc *6*, 1860-1869. 10.1038/nprot.2011.409.
26. Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841-842. 10.1093/bioinformatics/btq033.
27. GTEx Consortium (2020). The GTEx Consortium atlas of genetic regulatory effects across human tissues. Science *369*, 1318-1330. 10.1126/science.aaz1776.
28. Hinrichs, A.S., Karolchik, D., Baertsch, R., Barber, G.P., Bejerano, G., Clawson, H., Diekhans, M., Furey, T.S., Harte, R.A., Hsu, F., et al. (2006). The UCSC Genome Browser Database: update 2006. Nucleic Acids Res *34*, D590-598. 10.1093/nar/gkj144.
29. Giambartolomei, C., Vukcevic, D., Schadt, E.E., Franke, L., Hingorani, A.D., Wallace, C., and Plagnol, V. (2014). Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. PLoS Genet *10*, e1004383. 10.1371/journal.pgen.1004383.
30. Chan Zuckerberg Initiative. CZ CELLxGENE Discover. Date accessed: 6/7/2023. https://cellxgene.cziscience.com/
31. Cao, J., O'Day, D.R., Pliner, H.A., Kingsley, P.D., Deng, M., Daza, R.M., Zager, M.A., Aldinger, K.A., Blecher-Gonen, R., Zhang, F., et al. (2020). A human cell atlas of fetal gene expression. Science *370*. 10.1126/science.aba7721.
32. Stewart, B.J., Ferdinand, J.R., Young, M.D., Mitchell, T.J., Loudon, K.W., Riding, A.M., Richoz, N., Frazer, G.L., Staniforth, J.U.L., Vieira Braga, F.A., et al. (2019). Spatiotemporal immune zonation of the human kidney. Science *365*, 1461-1466. 10.1126/science.aat5031.
33. Litvinukova, M., Talavera-Lopez, C., Maatz, H., Reichart, D., Worth, C.L., Lindberg, E.L., Kanda, M., Polanski, K., Heinig, M., Lee, M., et al. (2020). Cells of the adult human heart. Nature *588*, 466-472. 10.1038/s41586-020-2797-4.
34. Jung, I., Schmitt, A., Diao, Y., Lee, A.J., Liu, T., Yang, D., Tan, C., Eom, J., Chan, M., Chee, S., et al. (2019). A compendium of promoter-centered long-range chromatin interactions in the human genome. Nat Genet *51*, 1442-1449. 10.1038/s41588-019-0494-8.
35. Schmitt, A.D., Hu, M., Jung, I., Xu, Z., Qiu, Y., Tan, C.L., Li, Y., Lin, S., Lin, Y., Barr, C.L., and Ren, B. (2016). A Compendium of Chromatin Contact Maps Reveals Spatially Active Regions in the Human Genome. Cell Rep *17*, 2042-2059. 10.1016/j.celrep.2016.10.061.
36. Ferkingstad, E., Sulem, P., Atlason, B.A., Sveinbjornsson, G., Magnusson, M.I., Styrmisdottir, E.L., Gunnarsdottir, K., Helgason, A., Oddsson, A., Halldorsson, B.V., et al. (2021). Large-scale integration of the plasma proteome with genetics and disease. Nat Genet *53*, 1712-1721. 10.1038/s41588-021-00978-w.
37. Stelzer, G., Rosen, N., Plaschkes, I., Zimmerman, S., Twik, M., Fishilevich, S., Stein, T.I., Nudel, R., Lieder, I., Mazor, Y., et al. (2016). The GeneCards Suite: From Gene Data Mining to Disease Genome Sequence Analyses. Curr Protoc Bioinformatics *54*, 1.30.31-31.30.33. 10.1002/cpbi.5.
38. Watanabe, K., Taskesen, E., van Bochoven, A., and Posthuma, D. (2017). Functional mapping and annotation of genetic associations with FUMA. Nat Commun *8*, 1826. 10.1038/s41467-017-01261-5.
39. Supek, F., Bosnjak, M., Skunca, N., and Smuc, T. (2011). REVIGO summarizes and visualizes long lists of gene ontology terms. PLoS One *6*, e21800. 10.1371/journal.pone.0021800.

40. Finan, C., Gaulton, A., Kruger, F.A., Lumbers, R.T., Shah, T., Engmann, J., Galver, L., Kelley, R., Karlsson, A., Santos, R., et al. (2017). The druggable genome and support for target identification and validation in drug development. Sci Transl Med *9*. 10.1126/scitranslmed.aag1166.
41. Sakaue, S., and Okada, Y. (2019). GREP: genome for REPositioning drugs. Bioinformatics *35*, 3821-3823. 10.1093/bioinformatics/btz166.
42. Schaid, D.J., Chen, W., and Larson, N.B. (2018). From genome-wide associations to candidate causal variants by statistical fine-mapping. Nat Rev Genet *19*, 491-504. 10.1038/s41576-018-0016-z.
43. Bernstein, B.E., Birney, E., Dunham, I., Green, E.D., Gunter, C., and Snyder, M. (2012). An integrated encyclopedia of DNA elements in the human genome. Nature *489*, 57-74. 10.1038/nature11247.
44. Milman, N.T., Schioedt, F.V., Junker, A.E., and Magnussen, K. (2019). Diagnosis and Treatment of Genetic HFE-Hemochromatosis: The Danish Aspect. Gastroenterology Res *12*, 221-232. 10.14740/gr1206.
45. Sounni, N.E., Dehne, K., van Kempen, L., Egeblad, M., Affara, N.I., Cuevas, I., Wiesen, J., Junankar, S., Korets, L., Lee, J., et al. (2010). Stromal regulation of vessel stability by MMP14 and TGFbeta. Dis Model Mech *3*, 317-332. 10.1242/dmm.003863.
46. Wu, J., Zhang, C., Liu, C., Zhang, A., Li, A., Zhang, J., and Zhang, Y. (2019). Aortic constriction induces hypertension and cardiac hypertrophy via (pro)renin receptor activation and the PLC-beta3 signaling pathway. Mol Med Rep *19*, 573-580. 10.3892/mmr.2018.9653.
47. Merkulova, M., Paunescu, T.G., Nair, A.V., Wang, C.Y., Capen, D.E., Oliver, P.L., Breton, S., and Brown, D. (2018). Targeted deletion of the Ncoa7 gene results in incomplete distal renal tubular acidosis in mice. Am J Physiol Renal Physiol *315*, F173-F185. 10.1152/ajprenal.00407.2017.
48. Mohney, B.G., Pulido, J.S., Lindor, N.M., Hogan, M.C., Consugar, M.B., Peters, J., Pankratz, V.S., Nasr, S.H., Smith, S.J., Gloor, J., et al. (2011). A novel mutation of LAMB2 in a multigenerational mennonite family reveals a new phenotypic variant of Pierson syndrome. Ophthalmology *118*, 1137-1144. 10.1016/j.ophtha.2010.10.009.
49. Larionov, A., Dahlke, E., Kunke, M., Zanon Rodriguez, L., Schiessl, I.M., Magnin, J.L., Kern, U., Alli, A.A., Mollet, G., Schilling, O., et al. (2019). Cathepsin B increases ENaC activity leading to hypertension early in nephrotic syndrome. J Cell Mol Med *23*, 6543-6553. 10.1111/jcmm.14387.
50. Bower, M., Salomon, R., Allanson, J., Antignac, C., Benedicenti, F., Benetti, E., Binenbaum, G., Jensen, U.B., Cochat, P., DeCramer, S., et al. (2012). Update of PAX2 mutations in renal coloboma syndrome and establishment of a locus-specific database. Hum Mutat *33*, 457-466. 10.1002/humu.22020.
51. Xu, Y., Rong, J., and Zhang, Z. (2021). The emerging role of angiotensinogen in cardiovascular diseases. J Cell Physiol *236*, 68-78. 10.1002/jcp.29889.
52. Carney, E.F. (2017). Hypertension: Role of ARHGAP42 in hypertension. Nat Rev Nephrol *13*, 134. 10.1038/nrneph.2017.13.
53. Koscielny, G., Yaikhom, G., Iyer, V., Meehan, T.F., Morgan, H., Atienza-Herrero, J., Blake, A., Chen, C.K., Easty, R., Di Fenza, A., et al. (2014). The International Mouse Phenotyping Consortium Web Portal, a unified point of access for knockout mice and related phenotyping data. Nucleic Acids Res *42*, D802-809. 10.1093/nar/gkt977.

54. Tsukahara, H., Gordienko, D.V., Tonshoff, B., Gelato, M.C., and Goligorsky, M.S. (1994). Direct demonstration of insulin-like growth factor-I-induced nitric oxide production by endothelial cells. Kidney Int *45*, 598-604. 10.1038/ki.1994.78.
55. Karamanavi, E., McVey, D.G., van der Laan, S.W., Stanczyk, P.J., Morris, G.E., Wang, Y., Yang, W., Chan, K., Poston, R.N., Luo, J., et al. (2022). The FES Gene at the 15q26 Coronary-Artery-Disease Locus Inhibits Atherosclerosis. Circ Res *131*, 1004-1017. 10.1161/CIRCRESAHA.122.321146.
56. Seo, H.-R., Jeong, H.E., Joo, H.J., Choi, S.-C., Park, C.-Y., Kim, J.-H., Choi, J.-H., Cui, L.-H., Hong, S.J., Chung, S., and Lim, D.-S. (2016). Intrinsic FGF2 and FGF5 promotes angiogenesis of human aortic endothelial cells in 3D microfluidic angiogenesis system. Scientific Reports *6*, 28832. 10.1038/srep28832.
57. Reynolds, J.O., Quick, A.P., Wang, Q., Beavers, D.L., Philippen, L.E., Showell, J., Barreto-Torres, G., Thuerauf, D.J., Doroudgar, S., Glembotski, C.C., and Wehrens, X.H.T. (2016). Junctophilin-2 gene therapy rescues heart failure by normalizing RyR2-mediated Ca2+ release. International Journal of Cardiology *225*, 371-380. 10.1016/j.ijcard.2016.10.021.
58. Collin, T., Lory, P., Taviaux, S., Courtieu, C., Guilbault, P., Berta, P., and Nargeot, J. (1994). Cloning, chromosomal location and functional expression of the human voltage-dependent calcium-channel beta 3 subunit. Eur J Biochem *220*, 257-262. 10.1111/j.1432-1033.1994.tb18621.x.
59. Bult, C.J., Blake, J.A., Smith, C.L., Kadin, J.A., Richardson, J.E., and Mouse Genome Database, G. (2019). Mouse Genome Database (MGD) 2019. Nucleic Acids Res *47*, D801-D806. 10.1093/nar/gky1056.
60. Newton-Cheh, C., Johnson, T., Gateva, V., Tobin, M.D., Bochud, M., Coin, L., Najjar, S.S., Zhao, J.H., Heath, S.C., Eyheramendy, S., et al. (2009). Genome-wide association study identifies eight loci associated with blood pressure. Nat Genet *41*, 666-676. 10.1038/ng.361.
61. International Consortium for Blood Pressure Genome-Wide Association, S., Ehret, G.B., Munroe, P.B., Rice, K.M., Bochud, M., Johnson, A.D., Chasman, D.I., Smith, A.V., Tobin, M.D., Verwoert, G.C., et al. (2011). Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. Nature *478*, 103-109. 10.1038/nature10405.
62. Rappaport, N., Twik, M., Plaschkes, I., Nudel, R., Iny Stein, T., Levitt, J., Gershoni, M., Morrey, C.P., Safran, M., and Lancet, D. (2017). MalaCards: an amalgamated human disease compendium with diverse clinical and genetic annotation and structured search. Nucleic Acids Res *45*, D877-D887. 10.1093/nar/gkw1012.
63. Wilson, M.P., Plecko, B., Mills, P.B., and Clayton, P.T. (2019). Disorders affecting vitamin B(6) metabolism. Journal of inherited metabolic disease *42*, 629-646. 10.1002/jimd.12060.
64. Wienke, D., Davies, G.C., Johnson, D.A., Sturge, J., Lambros, M.B., Savage, K., Elsheikh, S.E., Green, A.R., Ellis, I.O., Robertson, D., et al. (2007). The collagen receptor Endo180 (CD280) Is expressed on basal-like breast tumor cells and promotes tumor growth in vivo. Cancer Res *67*, 10230-10240. 10.1158/0008-5472.CAN-06-3496.
65. Imig, J.D. (2004). ACE Inhibition and Bradykinin-Mediated Renal Vascular Responses: EDHF Involvement. Hypertension *43*, 533-535. 10.1161/01.HYP.0000118054.86193.ce.

66. Fan, Y., Chen, Y., Zhang, J., Yang, F., Hu, Y., Zhang, L., Zeng, C., and Xu, Q. (2019). Protective Role of RNA Helicase DEAD-Box Protein 5 in Smooth Muscle Cell Proliferation and Vascular Remodeling. Circ Res *124*, e84-e100. 10.1161/CIRCRESAHA.119.314062.
67. Seda, O., Liska, F., Pravenec, M., Vernerova, Z., Kazdova, L., Krenova, D., Zidek, V., Sedova, L., Krupkova, M., and Kren, V. (2017). Connexin 50 mutation lowers blood pressure in spontaneously hypertensive rat. Physiol Res *66*, 15-28. 10.33549/physiolres.933432.
68. Chao, J., Jin, L., Lin, K.F., and Chao, L. (1997). Adrenomedullin gene delivery reduces blood pressure in spontaneously hypertensive rats. Hypertens Res *20*, 269-277. 10.1291/hypres.20.269.
69. Yan, Y., Liu, F., Dang, X., Zhou, R., and Liao, B. (2021). TBX3 induces biased differentiation of human induced pluripotent stem cells into cardiac pacemaker-like cells. Gene Expr Patterns *40*, 119184. 10.1016/j.gep.2021.119184.
70. Boix, C.A., James, B.T., Park, Y.P., Meuleman, W., and Kellis, M. (2021). Regulatory genomic circuitry of human disease loci by integrative epigenomics. Nature *590*, 300-307. 10.1038/s41586-020-03145-z.
71. Warren, H.R., Evangelou, E., Cabrera, C.P., Gao, H., Ren, M., Mifsud, B., Ntalla, I., Surendran, P., Liu, C., Cook, J.P., et al. (2017). Genome-wide association analysis identifies novel blood pressure loci and offers biological insights into cardiovascular risk. Nat Genet *49*, 403-415. 10.1038/ng.3768.
72. Huo, Y., Li, J., Qin, X., Huang, Y., Wang, X., Gottesman, R.F., Tang, G., Wang, B., Chen, D., He, M., et al. (2015). Efficacy of folic acid therapy in primary prevention of stroke among adults with hypertension in China: the CSPPT randomized clinical trial. JAMA *313*, 1325-1335. 10.1001/jama.2015.2274.
73. Perala, N., Jakobson, M., Ola, R., Fazzari, P., Penachioni, J.Y., Nymark, M., Tanninen, T., Immonen, T., Tamagnone, L., and Sariola, H. (2011). Sema4C-Plexin B2 signalling modulates ureteric branching in developing kidney. Differentiation *81*, 81-91. 10.1016/j.diff.2010.10.001.
74. Zhu, L., Vranckx, R., Khau Van Kien, P., Lalande, A., Boisset, N., Mathieu, F., Wegman, M., Glancy, L., Gasc, J.M., Brunotte, F., et al. (2006). Mutations in myosin heavy chain 11 cause a syndrome associating thoracic aortic aneurysm/aortic dissection and patent ductus arteriosus. Nat Genet *38*, 343-349. 10.1038/ng1721.
75. Kuang, S.Q., Kwartler, C.S., Byanova, K.L., Pham, J., Gong, L., Prakash, S.K., Huang, J., Kamm, K.E., Stull, J.T., Sweeney, H.L., and Milewicz, D.M. (2012). Rare, nonsynonymous variant in the smooth muscle-specific isoform of myosin heavy chain, MYH11, R247C, alters force generation in the aorta and phenotype of smooth muscle cells. Circ Res *110*, 1411-1422. 10.1161/CIRCRESAHA.111.261743.
76. Klewer, S.E., Krob, S.L., Kolker, S.J., and Kitten, G.T. (1998). Expression of type VI collagen in the developing mouse heart. Dev Dyn *211*, 248-255. 10.1002/(SICI)1097-0177(199803)211:3<248::AID-AJA6>3.0.CO;2-H.
77. Zech, M., Lam, D.D., Francescatto, L., Schormair, B., Salminen, A.V., Jochim, A., Wieland, T., Lichtner, P., Peters, A., Gieger, C., et al. (2015). Recessive mutations in the alpha3 (VI) collagen gene COL6A3 cause early-onset isolated dystonia. Am J Hum Genet *96*, 883-893. 10.1016/j.ajhg.2015.04.010.
78. El-Bassossy, H., Badawy, D., Neamatallah, T., and Fahmy, A. (2016). Ferulic acid, a natural polyphenol, alleviates insulin resistance and hypertension in

fructose fed rats: Effect on endothelial-dependent relaxation. Chem Biol Interact *254*, 191-197. 10.1016/j.cbi.2016.06.013.
79. Maass, P.G., Aydin, A., Luft, F.C., Schachterle, C., Weise, A., Stricker, S., Lindschau, C., Vaegler, M., Qadri, F., Toka, H.R., et al. (2015). PDE3A mutations cause autosomal dominant hypertension with brachydactyly. Nat Genet *47*, 647-653. 10.1038/ng.3302.
80. Ercu, M., Mucke, M.B., Pallien, T., Marko, L., Sholokh, A., Schachterle, C., Aydin, A., Kidd, A., Walter, S., Esmati, Y., et al. (2022). Mutant Phosphodiesterase 3A Protects From Hypertension-Induced Cardiac Damage. Circulation *146*, 1758-1778. 10.1161/CIRCULATIONAHA.122.060210.
81. Jain, A.P., Patel, K., Pinto, S., Radhakrishnan, A., Nanjappa, V., Kumar, M., Raja, R., Patil, A.H., Kumari, A., Manoharan, M., et al. (2019). MAP2K1 is a potential therapeutic target in erlotinib resistant head and neck squamous cell carcinoma. Sci Rep *9*, 18793. 10.1038/s41598-019-55208-5.
82. Olzinski, A.R., McCafferty, T.A., Zhao, S.Q., Behm, D.J., Eybye, M.E., Maniscalco, K., Bentley, R., Frazier, K.S., Milliner, C.M., Mirabile, R.C., et al. (2005). Hypertensive target organ damage is attenuated by a p38 MAPK inhibitor: role of systemic blood pressure and endothelial protection. Cardiovasc Res *66*, 170-178. 10.1016/j.cardiores.2004.12.021.
83. Eales, J.M., Jiang, X., Xu, X., Saluja, S., Akbarov, A., Cano-Gamez, E., McNulty, M.T., Finan, C., Guo, H., Wystrychowski, W., et al. (2021). Uncovering genetic mechanisms of hypertension through multi-omic analysis of the kidney. Nat Genet *53*, 630-637. 10.1038/s41588-021-00835-w.
84. Gazal, S., Weissbrod, O., Hormozdiari, F., Dey, K.K., Nasser, J., Jagadeesh, K.A., Weiner, D.J., Shi, H., Fulco, C.P., O'Connor, L.J., et al. (2022). Combining SNP-to-gene linking strategies to identify disease genes and assess disease omnigenicity. Nat Genet *54*, 827-836. 10.1038/s41588-022-01087-y.
85. Washington, C., 3rd, Dapas, M., Biddanda, A., Magnaye, K.M., Aneas, I., Helling, B.A., Szczesny, B., Boorgula, M.P., Taub, M.A., Kenny, E., et al. (2022). African-specific alleles modify risk for asthma at the 17q12-q21 locus in African Americans. Genome Med *14*, 112. 10.1186/s13073-022-01114-x.
86. Kelly, T.N., Sun, X., He, K.Y., Brown, M.R., Taliun, S.A.G., Hellwege, J.N., Irvin, M.R., Mi, X., Brody, J.A., Franceschini, N., et al. (2022). Insights From a Large-Scale Whole-Genome Sequencing Study of Systolic Blood Pressure, Diastolic Blood Pressure, and Hypertension. Hypertension *79*, 1656-1667. 10.1161/HYPERTENSIONAHA.122.19324.
87. Wang, Y., Guo, J., Ni, G., Yang, J., Visscher, P.M., and Yengo, L. (2020). Theoretical and empirical quantification of the accuracy of polygenic scores in ancestry divergent populations. Nat Commun *11*, 3865. 10.1038/s41467-020-17719-y.
88. Wojcik, G.L., Graff, M., Nishimura, K.K., Tao, R., Haessler, J., Gignoux, C.R., Highland, H.M., Patel, Y.M., Sorokin, E.P., Avery, C.L., et al. (2019). Genetic analyses of diverse populations improves discovery for complex traits. Nature *570*, 514-518. 10.1038/s41586-019-1310-4.
89. Breeze, C.E., Beck, S., Berndt, S.I., and Franceschini, N. (2022). The missing diversity in human epigenomic studies. Nat Genet *54*, 737-739. 10.1038/s41588-022-01081-4.
90. Bell, C.G., Gao, F., Yuan, W., Roos, L., Acton, R.J., Xia, Y., Bell, J., Ward, K., Mangino, M., Hysi, P.G., et al. (2018). Obligatory and facilitative allelic variation

in the DNA methylome within common disease-associated loci. Nat Commun *9*, 8. 10.1038/s41467-017-01586-1.

91. Battaglia, S., Dong, K., Wu, J., Chen, Z., Najm, F.J., Zhang, Y., Moore, M.M., Hecht, V., Shoresh, N., and Bernstein, B.E. (2022). Long-range phasing of dynamic, tissue-specific and allele-specific regulatory elements. Nat Genet *54*, 1504-1513. 10.1038/s41588-022-01188-8.
92. Surendran, P., Drenos, F., Young, R., Warren, H., Cook, J.P., Manning, A.K., Grarup, N., Sim, X., Barnes, D.R., Witkowska, K., et al. (2016). Trans-ancestry meta-analyses identify rare and common variants associated with blood pressure and hypertension. Nat Genet *48*, 1151-1161. 10.1038/ng.3654.
93. Liu, C., Kraja, A.T., Smith, J.A., Brody, J.A., Franceschini, N., Bis, J.C., Rice, K., Morrison, A.C., Lu, Y., Weiss, S., et al. (2016). Meta-analysis identifies common and rare variants influencing blood pressure and overlapping with metabolic trait loci. Nat Genet *48*, 1162-1170. 10.1038/ng.3660.
94. Backman, J.D., Li, A.H., Marcketta, A., Sun, D., Mbatchou, J., Kessler, M.D., Benner, C., Liu, D., Locke, A.E., Balasubramanian, S., et al. (2021). Exome sequencing and analysis of 454,787 UK Biobank participants. Nature *599*, 628-634. 10.1038/s41586-021-04103-z.

Table 1| High-confidence missense variants for blood pressure association signals (PP>0.95)

Signal ID	Index SNV	Missense variant	Canonical transcript	Annotation	Chr	Position	Polyphen	SIFT	Trait	p-value	Posterior probability (%)
1_6	rs262695	rs262695	ENST00000545087.1	*AL590822.1* p.Cys78Arg	1	2,144,788	N/A	N/A	SBP	9.40E-14	96.2
108_1	rs1047891	rs1047891	ENST00000430249.2	*CPS1* p.Thr1412Asn	2	211,540,507	benign	tolerated	SBP	1.40E-14	98.5
									DBP	8.20E-14	99.5
132_2	rs74951356	rs74951356[a]	ENST00000418109.1	*LAMB2* p.Ala1765Thr	3	49,158,763	benign	tolerated	DBP	5.00E-09	95.8
158_3	rs61762319	rs61762319[a]	ENST00000460393.1	*MME* p.Met8Val	3	154,801,978	benign	deleterious	SBP	1.50E-09	99.8
170_6	rs2498323	rs2498323	ENST00000382774.3	*HGFAC* p.Arg644Gln	4	3,451,109	possibly damaging	tolerated	PP	7.40E-13	100
191_3	rs13107325	rs13107325	ENST00000394833.2	*SLC39A8* p.Ala391Thr	4	103,188,709	possibly damaging	tolerated	SBP	4.20E-53	100
221_1	rs2307111	rs2307111	ENST00000428202.2	*POC5* p.His36Arg	5	75,003,678	benign	tolerated	DBP	1.60E-22	97.6
237_7	rs1800888	rs1800888[a]	ENST00000305988.4	*ADRB2* p.Thr164Ile	5	148,206,885	benign	tolerated	DBP	7.40E-13	100
249_1	rs1800730	rs1800730	ENST00000357618.5	*HFE* p.Ser65Cys	6	26,091,185	probably damaging	deleterious	SBP	2.00E-09	96.4
249_3	rs1800562	rs1800562[a]	ENST00000357618.5	*HFE* p.Cys282Tyr	6	26,093,141	probably damaging	deleterious	DBP	2.10E-37	96.4
251_7	rs41543814	rs41543814	ENST00000376228.5	*HLA-C* p.Ala97Thr	6	31,239,430	benign	tolerated (LC)	DBP	1.50E-19	100
251_10	rs2844573	rs2308655	ENST00000412585.2	*HLA-B* p.Cys349Ser	6	31,322,303	benign	tolerated (LC)	PP	1.40E-12	99.2
252_1	rs3176336	rs2395655	ENST00000448526.2	*CDKN1A* p.Asp28Gly	6	36,645,696	benign	tolerated (LC)	PP	4.70E-13	98.5
255_1	rs78648104	rs78648104	ENST00000008391.3	*TFAP2D* p.Phe74Leu	6	50,683,009	benign	tolerated	SBP	2.40E-15	99.9

272_1	rs6919947	rs6919947	ENST00000368357.3	*NCOA7* p.Ser399Ala	6	126,210,395	benign	tolerated (LC)	SBP	4.90E-17	100
300_3	rs2854746	rs2854746	ENST00000381083.4	*IGFBP3* p.Ala32Gly	7	45,960,645	benign	tolerated	DBP	5.00E-11	97.9
313_2	rs11556924	rs11556924	ENST00000358303.4	*ZC3HC1* p.Arg363His	7	129,663,496	probably damaging	deleterious	DBP	1.50E-26	98.3
365_2	rs76452347	rs76452347	ENST00000354323.2	*HRCT1* p.Arg63Trp	9	35,906,471	possibly damaging	deleterious (LC)	SBP	7.10E-14	100
379_2	rs6271	rs6271[a,b]	ENST00000393056.2	*DBH* p.Arg549Cys	9	136,522,274	possibly damaging	tolerated	SBP	1.20E-19	97.6
394_1	rs2236295	rs2236295	ENST00000373783.1	*ADO* p.Gly25Trp	10	64,564,892	possibly damaging	tolerated	SBP	2.80E-22	96.7
402_3	rs2274224	rs2274224[b]	ENST00000371380.3	*PLCE1* p.Arg1575Pro	10	96,039,597	benign	tolerated	SBP	9.00E-57	96.6
417_10	rs10770059 (SBP)	rs415895	ENST00000318950.6	*SWAP70* p.Gln505Glu	11	9,769,562	benign	tolerated	SBP	5.00E-47	96.5
432_4	rs117874826	rs117874826[a,b]	ENST00000540288.1	*PLCB3* p.Glu564Ala	11	64,027,666	benign	deleterious	SBP	2.30E-11	100
434_1	rs36027301	rs36027301	ENST00000265686.3	*TCIRG1* p.Arg56Trp	11	67,809,268	probably damaging	deleterious	SBP	6.00E-11	97.6
447_7	rs573455	rs573455	ENST00000278935.3	*CEP164* p.Gln1119Arg	11	117,267,884	benign	tolerated	PP	8.70E-34	100
463_15	rs1126930	rs1126930[a]	ENST00000316299.5	*PRKAG1* p.Thr98Ser	12	49,399,132	benign	tolerated	PP	4.60E-14	95.6
499_1	rs17880989	rs17880989	ENST00000311852.6	*MMP14* p.Met355Ile	14	23,313,633	benign	deleterious	DBP	3.20E-12	100
572_1	rs704	rs704	ENST00000226218.4	*VTN* p.Thr400Met	17	26,694,861	benign	tolerated	SBP	1.90E-08	96.5
582_6	rs34587622	rs34587622	ENST00000427177.1	*SEPT9* p.Pro145Leu	17	75,398,498	benign	tolerated (LC)	SBP	6.20E-14	99.9
606_9	rs167479	rs167479	ENST00000393423.3	*RGL3* p.Pro162His	19	11,526,765	probably damaging	deleterious	SBP	8.70E-69	100
610_2	rs45522544	rs45522544	ENST00000357324.6	*ATP13A1* p.Glu556Lys	19	19,765,499	benign	tolerated	DBP	3.10E-08	100
616_4	rs34093919	rs34093919	ENST00000308370.7	*LTBP4* p.Asp752Asn	19	41,117,300	possibly damaging	deleterious	PP	2.80E-14	97.4

616_7	rs1800470	rs1800470	ENST00000221930.5	*TGFB1* p.Pro10Leu	19	41,858,921	N/A	tolerated (LC)	PP	1.90E-15	99
617_2	rs7412	rs7412	ENST00000252486.4	*APOE* p.Arg176Cys	19	45,412,079	probably damaging	deleterious	SBP	2.00E-14	100
623_3	rs35761929	rs35761929	ENST00000254958.5	*JAG1* p.Pro871Arg	20	10,622,501	benign	deleterious	DBP	2.60E-18	99.1
636_1	rs2229742	rs2229742	ENST00000400202.1	*NRIP1* p.Arg448Gly	21	16,339,172	probably damaging	deleterious	SBP	7.40E-16	100
									PP	2.30E-11	99.7

SNV, single nucleotide variant; Chr, chromosome; SIFT, Sorting Intolerant from Tolerant algorithm which predicts the effect of coding variants on protein function; Polyphen, Polymorphism Phenotyping tool predicts possible impact of an amino acid substitution on the structure and function of a human protein; Posterior probability, the SNV's accounted posterior probability of driving the blood pressure association under the annotation-informed prior. [a] *indicates a low frequency variant (our data and in non-Finnish Europeans (https://gnomad.broadinstitute.org).* [b] *indicates supporting evidence for this gene from exome or EAWAS studies* [12,92-94]

2 .