Genome-wide association analysis identifies novel blood pressure loci and offers biological insights into cardiovascular risk

Elevated blood pressure is the leading heritable risk factor for cardiovascular disease worldwide. We report genetic association of blood pressure (systolic, diastolic, pulse pressure) among UK Biobank participants of European ancestry with independent replication in other cohorts, and robust validation of 107 independent loci. We also identify new independent variants at 11 previously reported blood pressure loci. In combination with results from a range of *in silico* functional analyses and wet bench experiments, our findings highlight new biological pathways for blood pressure regulation enriched for genes expressed in vascular tissues and identify potential therapeutic targets for hypertension. Results from genetic risk score models raise the possibility of a precision medicine approach through early lifestyle intervention to offset the impact of blood pressure–raising genetic variants on future cardiovascular disease risk.

Elevated blood pressure (BP) is a strong, heritable[1–4] and modifiable driver of risk for stroke and coronary artery disease and is a leading cause of global mortality and morbidity[5,6]. At the time of analysis, genome-wide association study (GWAS) meta-analyses, and analyses of custom or exome content, had identified and replicated genetic variants of mostly modest or weak effect on BP at over 120 loci[7–11]. Here we report association analyses between BP traits and genetic variants among ~140,000 participants in UK Biobank, a prospective cohort study of 500,000 men and women aged 40–69 years with extensive baseline phenotypic measurements, stored biological samples[12] and follow-up by electronic health record linkage[13]. We undertake independent replication in large international consortia and other cohorts, providing robust validation of our findings and new biological insights into BP regulation.

Our study design is summarized in **Figure 1**. Briefly, data are available for 152,249 UK Biobank participants genotyped using a customized array (including GWAS and exome content) and with genome-wide imputation based on 1000 Genomes and UK10K sequencing data[14]. (Further details on the UK Biobank imputation are available at the UK Biobank website.) After quality measures and exclusions (Online Methods), we study 140,886 unrelated individuals of European ancestry with two seated clinic BP measurements using the Omron HEM-7015IT device (**Supplementary Table 1**). We carry out GWAS analyses of systolic (SBP) and diastolic (DBP) blood pressure and of pulse pressure (PP) using single-variant linear regression under an additive model, based on ~9.8 million single-nucleotide variants (SNVs) with minor allele frequency (MAF) ≥1% and imputation quality score (INFO) >0.1. For SNVs with $P < 1 \times 10^{-6}$, we take forward for replication the sentinel SNV (that is, the SNV with the lowest P value) at each locus, defined by linkage disequilibrium (LD) $r^2 \geq 0.2$, within a 1-Mb interval. We similarly analyze exome content for variants with MAF ≥0.01%, including rare variants, taking into replication the sentinel SNV ($P < 1 \times 10^{-5}$) from loci that are non-overlapping ($r^2 < 0.2$) with the GWAS findings. Overall, we take sentinel SNVs from 240 loci into replication: 218 from GWAS and 22 from exome analysis ($r^2 < 0.2$ and >500 kb from previously reported BP SNVs at the time of analysis and not annotated to previously reported BP-associated genes; **Supplementary Table 2**).

The replication resources comprise individuals of European ancestry from a large BP meta-analysis consortium (ICBP cohorts listed in the **Supplementary Note**) and further cohorts with 1000 Genomes data for GWAS (**Supplementary Table 3**) and two large BP exome consortia. We use $P < 5 \times 10^{-8}$ to denote genome-wide significance in the combined (discovery and replication) meta-analyses, with $P < 0.01$ for support in the replication data alone and concordant direction of effect. Additionally, we take forward for replication potential secondary signals at 51 previously reported BP loci at the time of analysis (excluding the human leukocyte antigen (HLA) region).

To better understand the functional consequences of our findings, we carry out a series of *in silico* investigations and experimental analysis of gene expression in relevant vascular tissue for selected putative functional SNVs (**Supplementary Fig. 1**).

RESULTS

Genetic variants at novel loci at time of analysis

Of the 240 loci taken forward to replication, we validated 107 loci at $P < 5 \times 10^{-8}$, of which 102 derived from the GWAS analysis were replicated and subjected to meta-analysis in a total of 330,956 individuals (**Tables 1–3** and **Supplementary Figs. 2a–c** and **3a**), and a further 5 loci from the exome analysis in a total of 422,604 individuals (**Tables 1–3**, **Supplementary Fig. 3b** and **Supplementary Tables 4–6**). Thirty-two of these validated loci are novel findings. Since the time of analysis, the remaining 75 loci have also been reported in another study[15], although at least 53 of these were previously unvalidated (**Tables 1–3**);

A full list of authors and affiliations appears at the end of the paper.

Accepted 14 December 2016

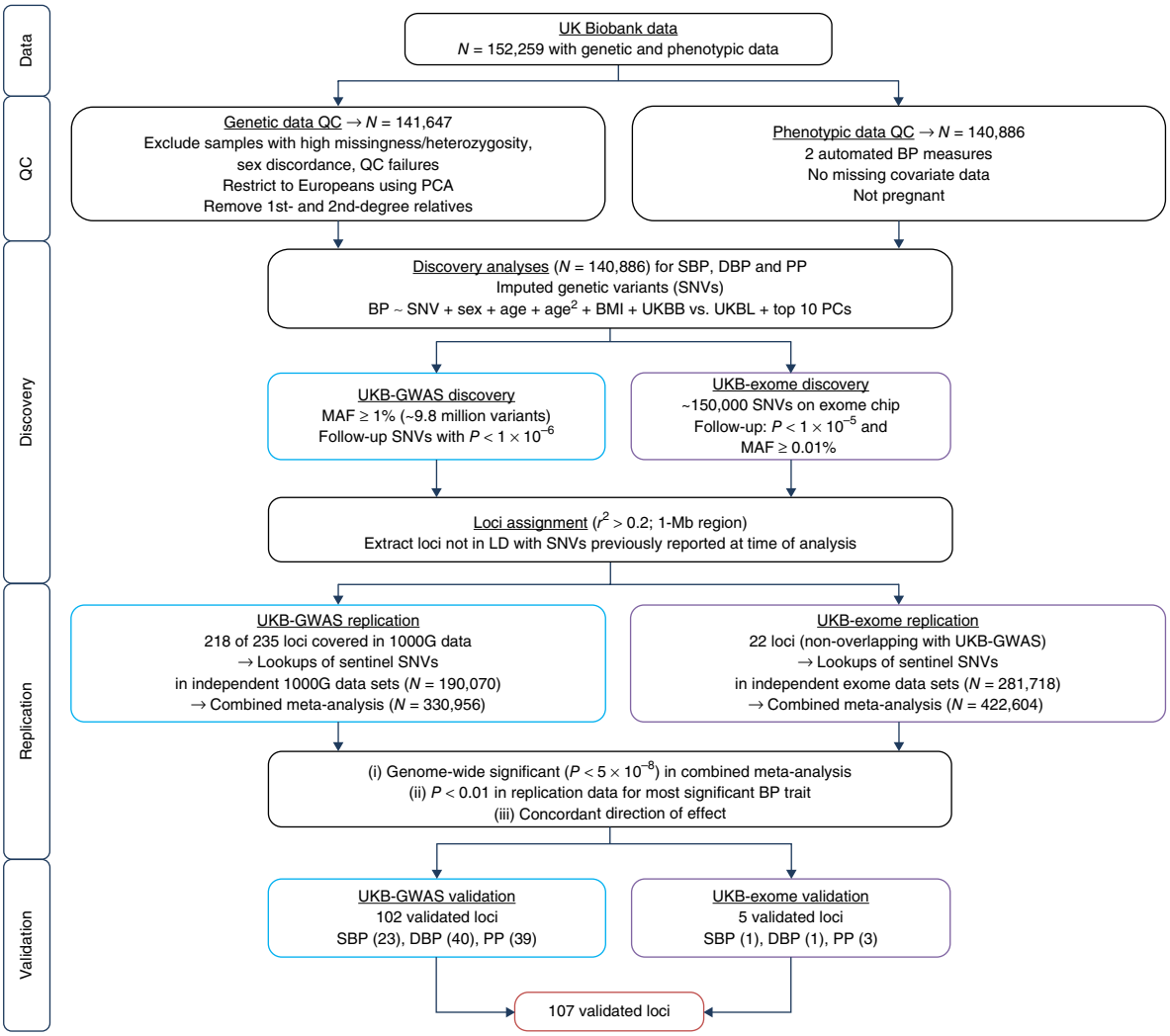


Figure 1 Study design schematic for discovery and validation of loci. *N*, sample size; QC, quality control; PCA, principal-component analysis; BP: blood pressure; SBP, systolic blood pressure; DBP, diastolic blood pressure; PP, pulse pressure; SNVs, single-nucleotide variants; BMI, body mass index; UKB, UK Biobank; UKBL, UK BiLEVE; MAF, minor allele frequency; LD, linkage disequilibrium; 1000G, 1000 Genomes; UKBBvsUKBL, a binary indicator variable for UK Biobank versus UK BiLEVE to adjust for the different genotyping chips.

hence, we have now validated these loci for the first time. We therefore present results for all 107 validated loci in our study. Most SNVs also showed association with hypertension in the UK Biobank data: for example, 93 of the 107 validated sentinel SNVs were nominally significant ($P < 0.01$) (**Supplementary Table 7**).

Of the 107 validated loci, 24 were reported for association with SBP as the primary trait (most significant from combined meta-analysis), 41 were reported for DBP and 42 were reported for PP, although many loci were associated with more than one BP trait (**Supplementary Fig. 4**). For example, in the combined meta-analysis, 24 validated loci were associated with both SBP and DBP, 11 were associated with SBP and PP, 1 was associated with DBP and PP, and 4 loci (*NADK-CPSF3L*, *GTF2B*, *METTL21A*-AC079767.3 and *PAX2*) were associated with all three traits at genome-wide significance (**Fig. 2**).

After conditional analysis on the sentinel SNV at each locus, we identified an independent validated secondary SNV at 5 of the 107 loci (**Supplementary Tables 8a** and **9**). In comparison with previously reported SNVs at the time of analysis, the contribution of our validated loci increases the percentage of trait variance explained by ~1%, for example, to 3.56% for SBP.

We report signals at known hypertension drug targets, including the *ACE* (angiotensin-converting enzyme) locus (rs4308, $P = 6.8 \times 10^{-14}$, ACE inhibitors), *CACNA2D2* (rs743757, $P = 2.4 \times 10^{-10}$, calcium channel blockers), *MME* (rs143112823 in the *RP11-439C8.2* locus, $P = 1.4 \times 10^{-14}$, omapatrilat), *ADRA2B* (rs2579519 in the *GPAT2–FAHD2CP* locus, $P = 4.8 \times 10^{-12}$, beta blockers), *SLC14A2* (rs7236548, $P = 2.0 \times 10^{-18}$, nifedipine) and *PDE5A* (phosphodiesterase 5A; rs66887589, $P = 3.4 \times 10^{-15}$, sildenafil).

Additionally, we evaluated our validated SNVs, where available, in cohorts of non-European ancestry[9–11], while recognizing that these analyses are likely underpowered (**Supplementary Table 10**). We found concordance in direction of effect ($P < 0.05$) for GWAS SNVs for all three BP traits among individuals of East Asian ancestry and for DBP for individuals of South Asian ancestry, also for exome SNVs among individuals of Hispanic ancestry, pointing to trans-ancestry effects for many of the BP-associated variants.

A PhenoScanner[16] search showed that 27 of our 107 validated sentinel SNVs (or proxies; $r^2 \geq 0.8$) exhibited genome-wide significant associations with other traits (**Supplementary Fig. 5a**), including coronary artery disease and myocardial infarction (where BP is likely

Table 1 Loci validated with SBP as the primary trait: combined meta-analysis results from GWAS and exome data for each sentinel variant

Locus[a]	Chr.	Position[b]	rsID	EA	EAF	*N*	*β*	s.e.	*P*	Note[c]
GWAS										
NADK–CPSF3L	1	1,685,921	rs139385870	D	0.50	281,890	–0.352	0.05	1.3×10^{-12}	GIU
CELA2A	1	15,798,197	rs3820068	A	0.81	310,776	0.425	0.06	1.1×10^{-12}	GIU
GTF2B	1	89,360,158	rs10922502	A	0.62	323,666	–0.382	0.05	2.2×10^{-15}	GIU
FOSL2	2	28,635,740	rs7562	T	0.52	319,942	0.263	0.05	1.9×10^{-8}	
PRKD3	2	37,517,566	rs13420463	A	0.77	330,307	0.356	0.05	7.0×10^{-11}	GIU
METTL21A–AC079767.3	2	208,526,140	rs55780018	T	0.54	304,567	–0.391	0.05	5.9×10^{-16}	GIU
RYK	3	134,000,025	rs9859176	T	0.40	322,428	0.322	0.05	1.3×10^{-11}	G
NPNT	4	106,911,742	rs13112725	C	0.76	306,370	0.435	0.06	1.5×10^{-14}	GIU
TMEM161B	5	87,514,515	rs10059921	T	0.08	298,543	–0.526	0.09	4.0×10^{-9}	GIU
FBN2	5	127,868,199	rs6595838	A	0.30	328,401	0.344	0.05	7.6×10^{-12}	GIU
CASC15	6	22,130,601	rs6911827	T	0.45	326,471	0.296	0.05	2.0×10^{-10}	GIU
TFAP2D	6	50,683,009	rs78648104	T	0.92	305,426	–0.481	0.08	1.3×10^{-8}	
MKLN1	7	131,059,056	rs13238550	A	0.40	325,647	0.331	0.05	1.9×10^{-12}	
HIPK2	7	139,463,264	rs1011018	A	0.20	325,110	–0.329	0.06	1.5×10^{-8}	
ZFAT	8	135,612,745	rs894344	A	0.60	329,834	–0.258	0.05	3.2×10^{-8}	
PAX2	10	102,604,514	rs112184198	A	0.10	323,791	–0.659	0.08	3.6×10^{-18}	GIU
MCF2L	13	113,636,156	rs9549328	T	0.23	313,787	0.318	0.06	1.5×10^{-8}	GI
FERMT2	14	53,377,540	rs9888615	T	0.29	326,235	–0.318	0.05	3.5×10^{-10}	GIU
PPP2R5E	14	63,928,546	rs8016306	A	0.80	329,869	0.335	0.06	3.7×10^{-9}	
ABHD17C	15	81,013,037	rs35199222	A	0.45	323,407	0.322	0.05	5.2×10^{-12}	GI
CFDP1	16	75,331,044	rs11643209	T	0.42	309,242	–0.339	0.05	1.8×10^{-12}	GI
CRK	17	1,333,598	rs12941318	T	0.49	299,739	–0.269	0.05	2.5×10^{-8}	GIU
ACOX1	17	73,949,045	rs2467099	T	0.22	326,401	–0.307	0.06	3.3×10^{-8}	GIU
Exome										
SSPN	12	26,438,189	rs6487543	A	0.77	244,842	0.300	0.05	6.3×10^{-10}	

SBP, systolic blood pressure; chr., chromosome; EA, effect allele; EAF, effect allele frequency in UK Biobank; *β*, effect estimate; s.e., standard error of effect; *N*, total sample size analyzed.

[a]Named according to the nearest annotated gene(s). [b]Given with respect to Build 37. [c]Loci published since our analysis[15] are indicated for GERA (G), GERA + ICBP (HapMap) (GI) or GERA + ICBP (HapMap) + UKB (GIU) analyses.

on the causal pathway[17]), cardiovascular risk factors (for example, lipids, height and body mass index) and non-cardiovascular traits (for example, lung function, cancer and Alzheimer's disease).

Variants at previously reported loci at time of analysis

In conditional analyses, we identified 20 secondary SNVs (15 common, 1 rare and 4 low-frequency variants) that were conditionally independent of the BP-associated SNVs at 16 previously reported loci at the time of analysis (**Supplementary Tables 8b, 11** and **12**). Nine of these SNVs have been reported since the time of analysis; hence, we identify 14 new secondary SNVs at 11 loci. One rare variant (rs138582164, MAF = 0.1%) in the *CDH17* locus predicted to act as an exonic stop-gain mutation at the *GEM* gene was associated with a relatively large effect on PP (3.5 mm Hg per allele copy; **Supplementary Table 8b**). At three previously reported loci (*EBF1*, *PDE3A* and *JAG1*), we identified multiple independent secondary SNVs in addition to the previously reported SNVs (**Supplementary Table 11**).

The UK Biobank data showed support ($P < 0.01$) for 119 of 122 previously reported BP loci at the time of analysis (159 of 163 SNVs) for one or more BP traits (**Supplementary Fig. 2a–c** and **Supplementary Table 13**). We did not find support for one SNV (rs11066280, *RPL6–ALDH1*) identified from a GWAS of East Asian ancestry[18], which may indicate ancestry-specific effects. We compared the MAF and effect sizes in UK Biobank with published results of previously reported variants at the time of analysis (**Supplementary Fig. 6**), finding consistency of results between the two sources of data.

We also examined findings for low-frequency and rare gene mutations previously reported to be associated with monogenic hypertension disorders[19] and included on the UK Biobank gene array. Despite lack of power overall, the variant with the lowest *P* value (rs387907156, *KLH3*, MAF = 0.02%) had a seemingly large effect on BP: 8.2 mm Hg (standard error (s.e.) = 4.1, $P = 0.046$) per allele for SBP and 5.6 mm Hg (s.e. = 2.6, $P = 0.048$) for PP (**Supplementary Table 14**).

Functional analyses

We annotated the 107 validated loci to 212 genes (on the basis of LD $r^2 \geq 0.8$) and sought putative function from *in silico* analyses and gene expression experiments. Candidate genes with the strongest supporting evidence are indicated in the last column of **Supplementary Table 4** with an indication of the supporting data source. All genome-wide significant variants in LD ($r^2 > 0.8$) with the variants reported here, ranked by supporting evidence, are annotated in **Supplementary Table 15**. Of the 107 validated sentinel SNVs, 3 were indels; all other variants were SNPs. We identified nonsynonymous SNVs at 13 of the 107 validated loci (**Supplementary Table 16**), 3 of which were predicted to be damaging (ANNOVAR) in *TFAP2D* (rs78648104), *NOX4* (rs56061986) and *CCDC141* (rs17362588; reported to be associated with heart rate[20]) (**Supplementary Fig. 5a**). Beyond the coding regions, we identify 29 SNVs in 3′ UTRs that were predicted to significantly weaken or cause loss of microRNA (miRNA) regulation by altering the recognition motif in 7 genes and to strengthen or create target sites for miRNA binding in 13 genes (based on miRNASNP db; **Supplementary Table 16**).

From our expression quantitative trait locus (eQTL) analysis (Genotype-Tissue Expression (GTEx)), 59 of the 107 validated loci contained variants with eQTLs in at least one tissue (**Supplementary Table 17**); arterial tissue had the largest number of loci with eQTLs (**Supplementary Fig. 7**), with targeted *in silico* analysis showing six loci with eQTLs in arterial tissue (**Supplementary Table 16**). For example, the GTEx tibial artery eQTL in *SF3A3* (rs4360494) showed strong *in silico* supporting evidence, including an arterial

Table 2 Loci validated with DBP as the primary trait: combined meta-analysis results from GWAS and exome data for each sentinel variant

Locus[a]	Chr.	Position[b]	rsID	EA	EAF	*N*	β	s.e.	*P*	Note[c]
GWAS										
chr1mb25	1	25,030,470	rs6686889	T	0.25	322,575	0.185	0.03	3.6×10^{-9}	
DNM3	1	172,357,441	rs12405515	T	0.56	328,543	–0.165	0.03	1.4×10^{-9}	GIU
GPATCH2	1	217,718,789	rs12408022	T	0.26	320,983	0.198	0.03	2.4×10^{-10}	GIU
CDC42BPA	1	227,252,626	rs10916082	A	0.73	327,636	–0.177	0.03	8.4×10^{-9}	
WNT3A	1	228,191,075	rs2760061	A	0.47	312,761	0.230	0.03	2.1×10^{-16}	GIU
SDCCAG8	1	243,471,192	rs953492	A	0.46	325,253	0.220	0.03	7.4×10^{-16}	G
ADCY3	2	25,139,596	rs55701159	T	0.89	321,052	0.285	0.04	7.2×10^{-11}	
SLC8A1	2	40,567,743	rs4952611	T	0.58	309,395	–0.157	0.03	4.0×10^{-8}	
AC016735.1	2	43,167,878	rs76326501	A	0.91	318,127	0.419	0.05	3.6×10^{-18}	
GPAT2–FAHD2CP	2	96,675,166	rs2579519	T	0.63	311,557	–0.197	0.03	4.8×10^{-12}	
TEX41	2	145,646,072	rs1438896	T	0.30	329,278	0.234	0.03	2.0×10^{-15}	GIU
CCDC141	2	179,786,068	rs79146658	T	0.91	321,318	–0.311	0.05	2.4×10^{-10}	G
TMEM194B	2	191,439,591	rs7592578	T	0.19	304,672	–0.240	0.04	9.5×10^{-12}	
TNS1	2	218,668,732	rs1063281	T	0.60	315,354	–0.200	0.03	1.3×10^{-12}	GIU
CAMKV–ACTBP13	3	49,913,705	rs36022378	T	0.80	319,983	–0.202	0.03	4.7×10^{-9}	GIU
CACNA2D2	3	50,476,378	rs743757	C	0.14	328,836	0.245	0.04	2.4×10^{-10}	GIU
FAM208A	3	56,726,646	rs9827472	T	0.37	323,058	–0.177	0.03	4.3×10^{-10}	GIU
RP11–439C8.2	3	154,707,967	rs143112823	A	0.09	297,343	–0.403	0.05	1.4×10^{-14}	GIU
SENP2	3	185,317,674	rs12374077	C	0.35	327,513	0.163	0.03	9.2×10^{-9}	GIU
PDE5A	4	120,509,279	rs66887589	T	0.52	324,397	–0.215	0.03	3.4×10^{-15}	GIU
POC5	5	75,038,431	rs10078021	T	0.63	314,172	–0.164	0.03	1.3×10^{-8}	G
CPEB4	5	173,377,636	rs72812846	A	0.28	312,601	–0.209	0.03	2.2×10^{-11}	GIU
PKHD1	6	51,832,494	rs13205180	T	0.49	325,419	0.168	0.03	7.0×10^{-10}	GIU
PDE10A	6	166,178,451	rs147212971	T	0.06	296,010	–0.360	0.06	1.6×10^{-9}	GIU
SLC35F1	6	118,572,486	rs9372498	A	0.08	330,625	0.334	0.05	1.8×10^{-11}	GIU
SNX31	8	101,676,675	rs2978098	A	0.54	324,424	0.165	0.03	1.5×10^{-9}	
RP11–273G15.2	8	144,060,955	rs62524579	A	0.53	268,645	–0.175	0.03	3.8×10^{-9}	GIU
MTAP	9	21,801,530	rs4364717	A	0.55	327,173	–0.175	0.03	1.3×10^{-10}	
BDNF	11	27,728,102	rs11030119	A	0.31	330,002	–0.163	0.03	2.9×10^{-8}	GIU
MYEOV	11	69,079,707	rs67330701	T	0.09	276,760	–0.367	0.05	2.1×10^{-12}	GIU
RP11–321F6.1	15	66,869,072	rs7178615	A	0.37	318,076	–0.179	0.03	2.6×10^{-10}	
ADAMTS7	15	79,070,000	rs62012628	T	0.29	244,143	–0.238	0.03	5.1×10^{-12}	
chr15mb95	15	95,312,071	rs12906962	T	0.68	319,952	–0.221	0.03	5.6×10^{-14}	GIU
PPL	16	4,943,019	rs12921187	T	0.43	326,469	–0.174	0.03	2.5×10^{-10}	G
FBXL19	16	30,936,743	rs72799341	A	0.24	324,502	0.185	0.03	5.8×10^{-9}	GIU
CMIP	16	81,574,197	rs8059962	T	0.42	319,839	–0.170	0.03	1.3×10^{-9}	
ACE	17	61,559,625	rs4308	A	0.37	319,394	0.213	0.03	6.8×10^{-14}	GIU
MAPK4	18	48,142,854	rs745821	T	0.76	330,954	0.189	0.03	1.4×10^{-9}	
CCNE1	19	30,294,991	rs62104477	T	0.33	320,347	0.177	0.03	1.2×10^{-9}	GIU
PLCB1	20	8,626,271	rs6108168	A	0.25	327,368	–0.211	0.03	1.1×10^{-11}	
Exome										
MRAS	3	138,119,952	rs2306374	T	0.84	281,715	–0.184	0.03	7.4×10^{-9}	GIU

DBP, diastolic blood pressure; chr., chromosome; EA, effect allele; EAF, effect allele frequency in UK Biobank; β, effect estimate; s.e., standard error of effect; *N*, total sample size analyzed.

[a]Named according to the nearest annotated gene(s). Loci >5 kb away from the nearest gene are named by chromosome number and position, e.g., chr1mb25 is located on chromosome 1 at 25 Mb. [b]Given with respect to Build 37. [c]Loci published since our analysis[15] are indicated for GERA (G), GERA + ICBP (HapMap) (GI) or GERA + ICBP (HapMap) + UKB (GIU) analyses.

DNase I–hypersensitive site within which the major C allele removes a predicted AP-2-binding site (**Supplementary Fig. 8**). Hence, we prioritized this gene for *in vitro* functional analysis.

By considering all loci reported here (our 107 validated loci and previously reported loci at the time of analysis), our DEPICT analysis identified enrichment of expression across 31 tissues and cell types (**Supplementary Fig. 9** and **Supplementary Table 18**), with the greatest enrichment in arteries ($P = 1.9 \times 10^{-6}$, false discovery rate (FDR) < 1%). We used FORGE to investigate and identify significant (FDR, $P < 0.05$) cell-type-specific enrichment within DNase I–hypersensitive sites in a range of tissues, including dermal and lung microvascular endothelial cell types and cardiac fibroblasts (**Supplementary Fig. 10**). For a set of curated candidate regulatory SNVs from our 107 validated loci (**Supplementary Note**), widespread enrichment was found in microvascular endothelium, aortic smooth muscle, aortic fibroblasts, vascular epithelium, heart and skin (**Supplementary Fig. 10**). In addition, we identified significant enrichment of histone marks in a wide range of cell types, including strong enrichment for trimethylation of histone H3 at lysine 4 (H3K4me3; an activating modification found near promoters) in umbilical vein endothelial cells (HUVECs) (**Supplementary Fig. 11**). To explore expression specifically at the level of cardiovascular cell types, we used Fantom5 reference transcript expression data (Online Methods) to cluster the 212 genes annotated to our 107 validated loci according to tissue specificity (**Supplementary**

Table 3 Loci validated with PP as the primary trait: combined meta-analysis results from GWAS and exome data for the sentinel variant

Locus[a]	Chr.	Position[b]	rsID	EA	EAF	*N*	*β*	s.e.	*P*	Note[c]
GWAS										
chr1mb9	1	9,441,949	rs9662255	A	0.43	310,618	–0.207	0.03	1.9×10^{-10}	GIU
SF3A3	1	38,455,891	rs4360494	C	0.55	282,851	0.278	0.03	3.7×10^{-16}	G
RP4-710M16.1–PPAP2B	1	56,576,924	rs112557609	A	0.35	325,952	0.227	0.03	6.8×10^{-12}	
FGGY	1	59,653,742	rs3889199	A	0.71	329,486	0.351	0.03	1.8×10^{-24}	G
C2orf43	2	20,881,840	rs2289081	C	0.36	329,140	–0.223	0.03	5.5×10^{-12}	GI
PRKCE	2	46,363,336	rs11690961	A	0.88	327,847	0.340	0.05	3.9×10^{-12}	GIU
CEP68	2	65,283,972	rs74181299	T	0.62	324,224	0.230	0.03	9.6×10^{-13}	GIU
TCF7L1	2	85,491,365	rs11689667	T	0.54	330,634	0.176	0.03	1.7×10^{-8}	GIU
FN1	2	216,300,482	rs1250259	A	0.74	325,485	–0.314	0.04	8.7×10^{-19}	G
GATA2	3	128,201,889	rs62270945	T	0.03	279,925	0.607	0.10	1.8×10^{-9}	GIU
PALLD	4	169,717,148	rs1566497	A	0.42	320,948	0.236	0.03	1.9×10^{-13}	GI
chr4mb174	4	174,584,663	rs17059668	C	0.92	313,277	–0.332	0.06	2.8×10^{-8}	
LHFPL2	5	77,837,789	rs10057188	A	0.46	325,985	–0.205	0.03	6.7×10^{-11}	GIU
GJA1	6	121,781,390	rs11154027	T	0.47	316,708	0.207	0.03	1.1×10^{-10}	
ESR1	6	152,397,912	rs36083386	I	0.11	323,303	0.439	0.05	1.5×10^{-18}	G
FNDC1	6	159,699,125	rs449789	C	0.14	325,584	0.359	0.05	2.4×10^{-15}	GIU
THBS2	6	169,587,103	rs1322639	A	0.78	319,866	0.316	0.04	4.8×10^{-17}	G
SUGCT	7	40,447,971	rs76206723	A	0.10	328,162	–0.346	0.05	7.4×10^{-12}	GIU
SLC20A2	8	42,324,765	rs2978456	T	0.55	304,964	–0.188	0.03	1.2×10^{-8}	GIU
TRAPPC9	8	141,060,027	rs4454254	A	0.63	330,022	–0.261	0.03	5.1×10^{-16}	
SCAI	9	127,900,996	rs72765298	T	0.87	316,271	–0.374	0.05	2.7×10^{-14}	GI
KIAA1462	10	30,317,073	rs9337951	A	0.34	299,646	0.280	0.04	2.5×10^{-15}	G
ARHGAP12	10	32,082,658	rs10826995	T	0.71	327,373	–0.212	0.03	1.1×10^{-9}	GIU
PRDM11	11	45,208,141	rs11442819	I	0.11	326,483	–0.279	0.05	7.1×10^{-9}	GIU
NOX4	11	89,224,453	rs2289125	A	0.21	307,682	–0.377	0.04	9.1×10^{-22}	G
CEP164	11	117,283,676	rs8258	T	0.38	327,038	0.236	0.03	2.9×10^{-13}	G
CCDC41	12	94,880,742	rs139236208	A	0.10	291,244	–0.363	0.06	1.6×10^{-10}	G
RP11-61O1.1	14	98,587,630	rs9323988	T	0.63	327,551	–0.212	0.03	4.1×10^{-11}	GIU
VAC14	16	70,755,610	rs117006983	A	0.01	250,766	0.986	0.14	4.1×10^{-12}	
CDH13	16	83,045,790	rs7500448	A	0.75	321,958	0.329	0.04	1.1×10^{-19}	G
KIAA0753	17	6,473,828	rs7226020	T	0.56	303,389	–0.256	0.03	2.3×10^{-14}	GIU
TP53-SLC2A4	17	7,571,752	rs78378222	T	0.99	294,053	0.904	0.14	1.8×10^{-10}	GIU
KCNH4–HSD17B1	17	40,317,241	rs79089478	T	0.97	318,326	0.584	0.10	3.1×10^{-9}	
PYY	17	42,060,631	rs62080325	A	0.66	315,689	–0.186	0.03	4.0×10^{-8}	
MRC2	17	60,767,151	rs740698	T	0.56	311,450	–0.228	0.03	3.1×10^{-12}	
SLC14A2	18	43,097,750	rs7236548	A	0.18	330,075	0.352	0.04	2.0×10^{-18}	G
SLC24A3	20	19,465,907	rs6081613	A	0.28	315,546	0.263	0.04	1.6×10^{-13}	GIU
ARVCF	22	19,967,980	rs12628032	T	0.30	310,292	0.240	0.03	5.5×10^{-12}	GIU
XRCC6	22	42,038,786	rs73161324	T	0.05	267,722	0.496	0.07	2.8×10^{-11}	
Exome										
CD34	1	208,024,820	rs12731740	T	0.10	279,078	–0.249	0.04	1.1×10^{-8}	
ZNF638	2	71,627,539	rs3771371	T	0.57	280,285	–0.160	0.03	5.8×10^{-9}	GIU
CRACR2B	11	828,916	rs7126805	A	0.73	145,162	0.222	0.04	3.3×10^{-9}	

PP, pulse pressure; chr., chromosome; EA, effect allele; EAF, effect allele frequency in UK Biobank; *β*, effect estimate; s.e., standard error of effect; *N*, total sample size analyzed.
[a]Named according to the nearest annotated gene(s). [b]Given with respect to Build 37. [c]Loci published since our analysis[15] are indicated for GERA (G), GERA + ICBP (HapMap) (GI) or GERA + ICBP (HapMap) + UKB (GIU) analyses.

Fig. 12), with the significantly clustered genes forming four tissue-specific clusters, including a vascular smooth muscle cell (VSMC) and fibroblast cluster, an endothelial cell cluster (including probable endothelial cells in highly vascularized tissues) and a combined vascular cell cluster.

Additionally, Ingenuity pathway analysis and upstream transcriptional analysis showed enrichment of canonical pathways implicated in cardiovascular disease, including those targeted by antihypertensive drugs, such as the α-adrenergic, CXCR4, endothelin signaling and angiotensin receptor pathways (**Supplementary Table 19**). In keeping with vascular mediation of genetic influence, we identified diphenyleneiodonium, an inhibitor of flavin-containing oxidases, including NAD(P)H oxidase (NOX), which is reported to reverse endothelial dysfunction (and hypertension) in a rat model[21].

To identify long-range target genes of noncoding variants, we used chromatin interaction (Hi-C) data from HUVECs, as enhancers and silencers often form chromatin loops with their target promoter. In most loci, the strongest promoter interaction involved a gene in high LD with the SNV, but for 21 loci we found a distal potential target gene (**Supplementary Table 16**). Pathway analysis of the distal genes showed the greatest enrichment in regulators of cardiac hypertrophy.

We evaluated pleiotropy using the Genomic Regions Enrichment of Annotations Tool (GREAT) to study enrichment of mouse phenotype and human disease ontology terms across all loci reported here. These

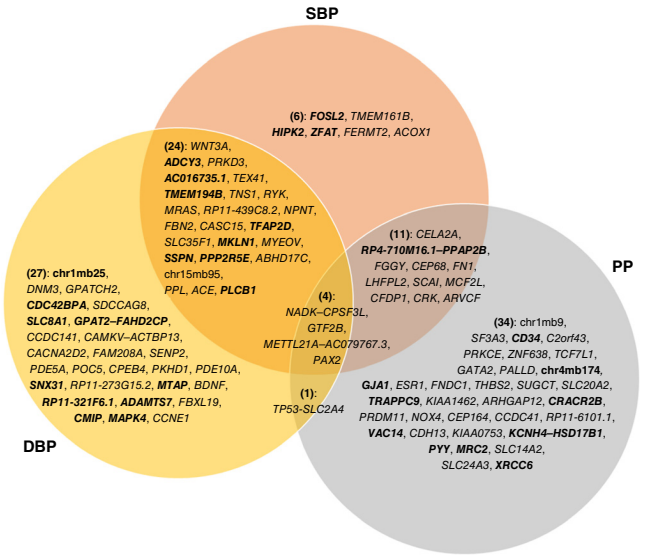


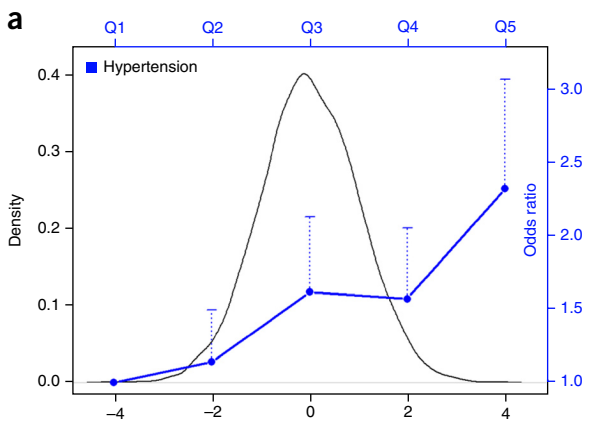
Figure 2 Venn diagram of the 107 validated loci from our study. This diagram shows concordance of significant associations across the three BP phenotypes for the 107 validated sentinel variants (**Tables 1**–**3**) from both the GWAS and exome analyses, according to genome-wide significance in the combined meta-analysis. The locus names labeled within the Venn diagram correspond to those in **Tables 1**–**3** and relate to the nearest annotated gene. The locus names in bold highlight the 32 loci that are reported for the first time in our study. SBP, systolic blood pressure; DBP, diastolic blood pressure; PP, pulse pressure.

highlighted cardiovascular system abnormalities and vascular disease as the most highly enriched terms (**Supplementary Fig. 5b,c**).

Collectively, evidence from eQTLs, DEPICT, DNase I–hypersensitive sites, histone marks, Hi-C data and ontological analyses indicates predominant vascular and cardiovascular tissue involvement for genes within the BP-associated loci.

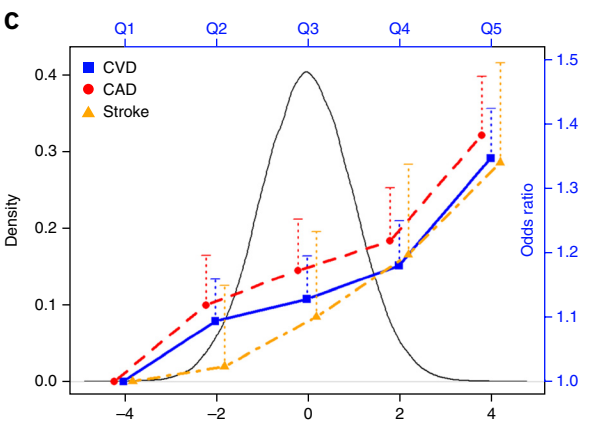
We also looked for association of our validated sentinel SNVs with metabolomic signatures. Three SNVs within the *NOX4*, *KCNH4* and *LHFPL2* loci showed significant associations (family-wise error rate < 5%) with lipoprotein subfractions from ^{1}H nuclear magnetic resonance (NMR) spectroscopy analysis of 2,000 Airwave study samples (**Supplementary Tables 20** and **21**). The results for these variants suggest a link between BP regulation and lipid metabolism. Eleven SNVs (including at the *LHFPL2* locus) showed association (family-wise error rate < 5%) with metabolites in blood or urine from the publicly available Metabolomics GWAS Server resource based on mass spectrometry[22,23] (**Supplementary Table 21**), including sugar acids, sphingolipids, fatty acids, glycerophospholipids, organic acids and benzene derivatives.

Several genes and variants with putative function were highlighted in our *in silico* analysis as having biological support (for example, eQTLs or nonsynonymous SNVs), and those with novelty and tractability to laboratory investigation (for example, expression in available tissue models) were prioritized. Sentinel variants in three genes that were highly significant in the combined meta-analysis (**Tables 2** and **3**) were selected for experimental testing and were successfully genotyped, each for at least 100 samples. We selected *ADAMTS7* because of strong biological support (for example, the mouse knockout phenotype), *SF3A3* because of eQTLs and *NOX4* because it contains a rare nonsynonymous SNV (**Supplementary Table 9**), in addition to common variant associations. We used qPCR to study the impact of these sentinel variants on gene expression in human VSMCs and endothelial cells (Online Methods). For *SF3A3*, the major C allele of variant rs4360494 associated with increased PP (0.278 (±0.03 (s.e.)) mm Hg, $P = 3.7 \times 10^{-16}$, $N = 307{,}682$) was associated with *SF3A3* expression in human VSMCs, although not in endothelial cells (**Supplementary Fig. 13a**) and the T allele of SNV rs62012628



b

(mm Hg)	Q1	Q2	Q3	Q4	Q5
SBP	136.4 (17.2)	138.7 (16.5)	142.5 (18.8)	142.5 (19.4)	145.7 (19.7)
DBP	82.9 (9.9)	84.3 (10.3)	85.9 (11.1)	86.0 (10.9)	87.6 (11.9)
PP	53.4 (11.6)	54.4 (11.3)	56.7 (12.3)	56.5 (12.4)	58.2 (12.4)



d

(count)	Q1	Q2	Q3	Q4	Q5
CVD	2,462	2,689	2,759	2,874	3,229
CAD	1,783	1,996	2,078	2,154	2,417
Stroke	581	597	640	695	776

Figure 3 Distribution of genetic risk score and its relationship with blood pressure, hypertension and cardiovascular disease outcomes. The GRS is based on all reported loci: both previously reported loci at the time of analysis and all validated BP variants from this study. (**a**) Distribution of GRS in Airwave and sex-adjusted odds ratios of hypertension in individuals of age 50+ years comparing each of the upper four GRS quintiles with the lowest quintile; dotted lines represent the upper 95% confidence intervals. (**b**) Mean BP (s.d.) in Airwave individuals of age 50+ years across GRS quintiles. (**c**) Distribution of GRS in UKB and sex-adjusted odds ratios of cardiovascular disease (CVD), coronary artery disease (CAD) and stroke comparing each of the upper four GRS quintiles with the lowest quintile; dotted lines represent the upper 95% confidence intervals. (**d**) Count of CVD, CAD and stroke (events and deaths) across GRS quintiles in UKB participants.

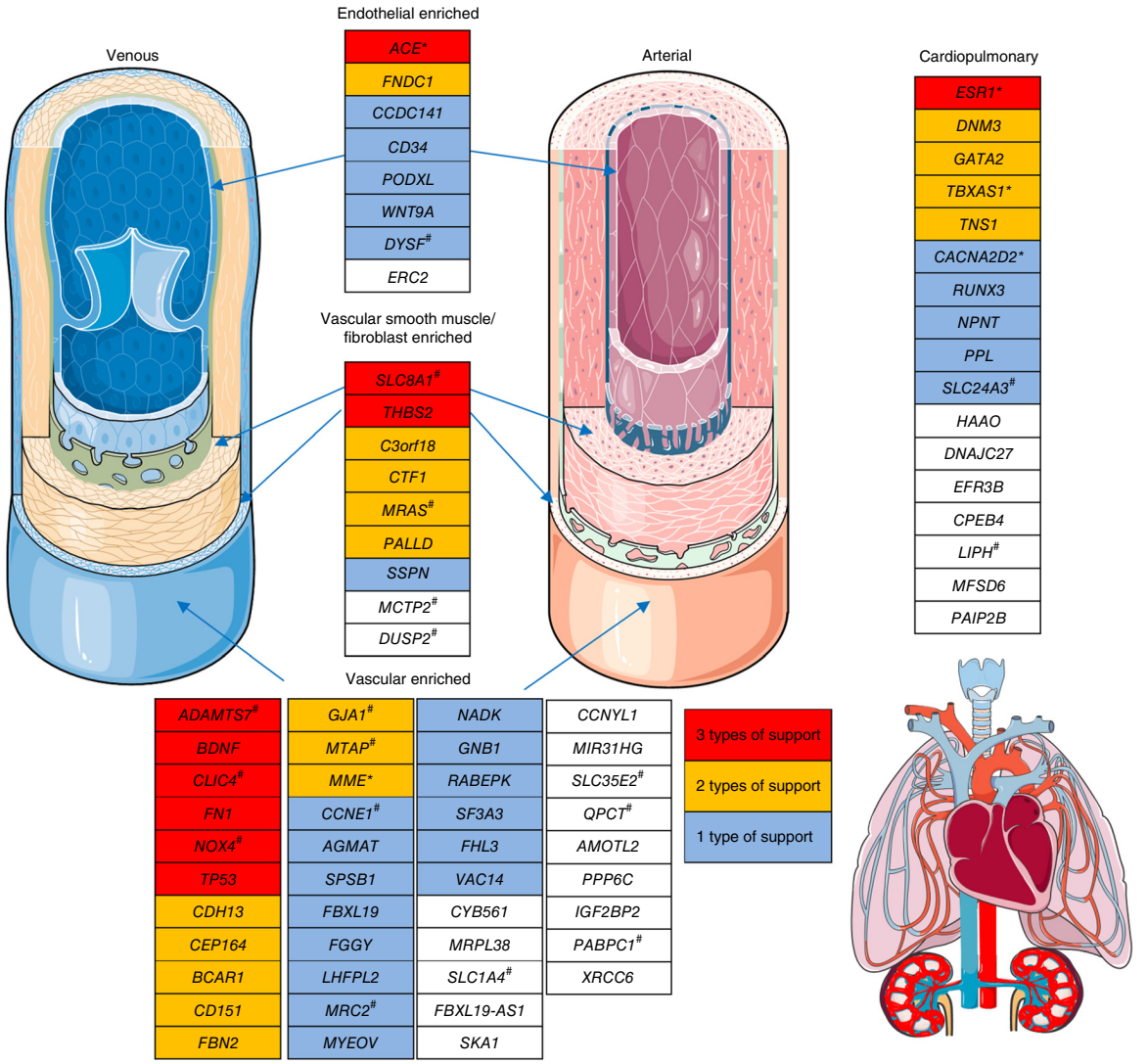


Figure 4 Summary of cardiovascular gene expression from validated loci. Genes are shown on the basis of their tissue expression and supporting evidence summarized in **Supplementary Table 16**, based on knockout (KO) phenotype and previously reported BP biology or a strong functional rationale: eQTL, nonsynonymous SNV or Hi-C data. Multiple lines of evidence indicate the central importance of the vasculature in BP regulation, and we thus highlight existing drugged (*) and druggable (#) targets among these genes. Illustrations used elements with permission from Servier Medical Art. We note that some druggable genes may carry a safety liability, such as *GJA1*, which has known association with QT interval[20].

in *ADAMTS7* associated with lower DBP (0.238 (±0.03) mm Hg, $P = 5.1 \times 10^{-12}$, $N = 244{,}143$) was associated with reduced *ADAMTS7* expression in human VSMCs (**Supplementary Fig. 13b**), while the minor A allele of SNV rs2289125 at the *NOX4* locus associated with lower PP (−0.377 (±0.04) mm Hg, $P = 9.1 \times 10^{-22}$, $N = 282{,}851$) correlated with increased *NOX4* expression in endothelial cells although not VSMCs (**Supplementary Fig. 13c**). Our study thus finds evidence for novel *cis*-eQTLs in *ADAMTS7* and *NOX4* in addition to validating the previously reported GTEx eQTL in *SF3A3*, and it supports the vascular expression of these genes.

Genetic risk score analyses

We created an unbiased genetic risk score (GRS) (**Supplementary Table 22**) to evaluate, in an independent cohort (Airwave; Online Methods), the impact of the combination of all loci reported here on BP levels and risk of hypertension. When compared with the lowest quintile of the distribution of the GRS, individuals >50 years old in the highest quintile had sex-adjusted mean SBP that was 9.3 mm Hg higher (95% confidence interval (CI) = 6.9 to 11.7 mm Hg, $P = 1.0 \times 10^{-13}$) and an over twofold higher risk of hypertension (odds ratio (OR) = 2.32, 95% CI = 1.76 to 3.06, $P = 2.8 \times 10^{-9}$) as compared with individuals in the lowest quintile (**Fig. 3** and **Supplementary Table 23**). Similar results were obtained from GRS associations with BP and hypertension within UK Biobank (**Supplementary Table 24**). In UK Biobank—based on self-reported health data, record linkage to Hospital Episode Statistics and mortality follow-up data (**Supplementary Table 25**)—we showed that the GRS was associated with increased risk of stroke, coronary heart disease and all cardiovascular outcomes; comparing the upper and lower quintiles of the GRS distribution, sex-adjusted odds ratios were 1.34 (95% CI = 1.20 to 1.49, $P = 1.5 \times 10^{-7}$), 1.38 (95% CI = 1.30 to 1.47, $P = 4.3 \times 10^{-23}$) and 1.35 (95% CI = 1.27 to 1.42, $P = 1.3 \times 10^{-25}$), respectively (**Fig. 3** and **Supplementary Table 26**). Results are also provided for incident-only cases (**Supplementary Table 27**).

DISCUSSION

A key attribute of this study is the combination of a large, single discovery sample with standardized BP measurement and dense 1000 Genomes and UK10K imputation, yielding a high-quality data set with ~9.8 million variants[14], taking advantage of major international consortia for parallel replication of common and low-frequency variants. In total, we include GWAS data from 330,956 individuals and exonic SNVs from a total of 422,604 individuals. This strategy resulted in the identification of 107 robustly validated loci for BP traits, including 32 loci that have not previously been reported, and at least 53 further loci were validated for the first time. Despite its size, our study is still underpowered to find low-frequency variants. Our findings are mostly common variants, with modest effect sizes that are similar to those for variants previously reported at the time of analysis (**Supplementary Fig. 14**). The lack of rare variant discovery could also be due to the challenge of detecting rare variants from imputed data. There may be greater potential for identifying rare variants from the future release of genetic data for all 500,000 UK Biobank participants.

Our findings point to new biology as well as highlighting gene regions in systems that have previously been implicated in the genetics of BP. Several of our validated loci affect atherosclerosis or vascular remodeling (*ADAMTS7*, *THBS2* and *CFDP1*) and exhibit locus pleiotropy in previous GWAS for coronary artery disease or carotid intimal media thickness[24–26] (**Fig. 4** and **Supplementary Fig. 5a**). In previous work, we have shown that expression of *ADAMTS7* is upregulated and increases VSMC migration in response to vascular injury in relation to a distinct coronary artery variant (rs3825807; not in LD with our sentinel SNV, $r^2 = 0.17$) (ref. 27). In endothelial cells, *ADAMTS7* encodes a metalloproteinase to cleave thrombospondin-1 encoded by *THBS2*, and this cleavage leads to reduced endothelial cell migration and has a role in neointimal repair in the vessel wall[27]. Our functional work indicates that the allele associated with lower DBP is also associated with lower *ADAMTS7* expression in human VSMCs; this fits with the mouse knockout that exhibits reduced atherosclerosis. *SF3A3* encodes a splicing factor with no previous links to BP other than our reported association and eQTL. At the *CFDP1* locus, our sentinel SNV is in high LD ($r^2 = 0.95$) with a variant previously associated with carotid intimal medial thickness. Collectively, our findings highlight a potential common mechanism among these genes in vascular remodeling that has previously been observed in small-resistance arteries in essential hypertension[28].

NADPH oxidase 4 (*NOX4*) has an established role in the endothelium, where it enhances vasodilatation and reduces BP *in vivo*[29]. This oxidase generates reactive oxygen species in the endothelium and may contribute to salt-sensitive hypertension in the kidney and the vasculature[30–32]. We found that the allele of the common variant at the *NOX4* locus correlates with increased tissue-specific *NOX4* expression in endothelial cells rather than VSMCs (**Supplementary Fig. 13c**). *NOX4* mediates endothelial cell apoptosis and facilitates vascular collagen synthesis, contributing to endothelial dysfunction and arterial stiffness, and may explain the association with PP[33,34].

We identify several loci containing genes involved in vascular signaling and second messenger systems, such as *PDE5A* and *PDE10A*[35–37]. The phosphodiesterase PDE5A hydrolyzes cyclic GMP and is inhibited by sildenafil, which leads to vasodilatation[38]. This finding fits with our previous discoveries of a role for gene loci encoding elements of the natriuretic peptide–nitric oxide pathway and guanylate cyclase signaling systems in BP regulation[18,39,40]. Our findings strengthen the case for evaluating the opportunity to repurpose PDE5A inhibitors for use in hypertension.

The importance of microvascular function is emphasized by the identification of solute carrier transporters such as *SLC14A2* encoding a urea transporter, which has previously been linked to autosomal dominant Streeten-type orthostatic hypotensive disorder[41] and BP response to nifedipine, a calcium channel blocker antihypertensive drug[42]. *SLC8A1* encodes a sodium–calcium exchanger expressed in cardiomyocytes that alters cardiac contractility and hypertrophy. This shows abnormal BP in *SLC8A1*-transgenic mice[43]. Variants at *SLC35F1* have previously been associated with resting heart rate and ventricular size, which could contribute to BP elevation[44].

We also identify loci that are involved in cardiovascular development (*GATA2*, *KIAA1462*, *FBN2*, *FN1* and *HAND2*) such as fibrillin 2 (*FBN2*), which overlaps in action with fibrillin 1 in development of the aortic matrix[45–49]. In addition, fibronectin expression is increased in hypertension and in atherosclerosis, but it may also have a role in the development of the heart[49–51].

Our analysis validates loci containing genes with previous physiological connection to BP such as *BDNF*, *FAM208A* and *CACNA2D2* (refs. 52–54). The neurotrophin brain-derived neurotrophic factor (BDNF) modulates angiotensin 11 in the brain to elevate BP in experimental models; higher serum levels correlate with reduced risk of cardiovascular disease and mortality[52]. In experimental models, FAM208A, which is thought to be a transcription factor, is a strong candidate for a QTL of BP[54]. The gene *CACNA2D2* encodes a subunit of the L-type calcium channel that is most abundantly expressed in the atrium and in neurons and may be a target for negatively chronotropic and inotropic calcium channel antagonists that reduce BP[55].

We examine long-range genomic interactions using Hi-C, whereby we find promoter regions that have a strong chromatin interaction with a novel SNV. One example is *EPAS1*, which is ~200 kb away from the SNV (rs11690961). It encodes hypoxia-inducible factor 2α, which affects catecholamine homeostasis and protects against heart failure; mutations in the gene are associated with pulmonary hypertension[56]. Another such gene is *INHBA*, 1.3 Mb away from the SNV (rs12531683), the product of which is elevated in pulmonary hypertension and contributes to vascular remodeling by inducing expression of endothelin-1 and plasminogen activator inhibitor-1 in pulmonary smooth muscle cells[57].

Our observation of BP that is 9–10 mm Hg higher at age 50 years and older when comparing the top and bottom fifths of the BP GRS distribution has potential clinical and public health implications. We stratified by age owing to a significant interaction of the GRS with age (P values ranging between 9.96×10^{-11} and 1.16×10^{-3} for interaction with continuous BP traits, $P = 0.012$ for hypertension). Measuring the GRS in early life increases the possibility of adopting an early precision medicine approach to offset genetic risk through lifestyle intervention (that is, reduced sodium intake, increased potassium intake, maintenance of optimal weight, low adult alcohol consumption and regular exercise)[58–60]. Studies of non-pharmacological approaches to BP control indicate that a reduction of 10 mm Hg or more in SBP is an achievable goal through lifestyle measures alone[61], while recent evidence suggests that favorable lifestyle may offset the cardiovascular sequelae associated with high genetic risk[62]. As the above data are observational, the extent to which adherence to lifestyle recommendations among individuals at high genetic risk might result in favorable outcomes remains uncertain; given the substantial effect of the GRS on BP by middle age, the potential for adopting early lifestyle intervention among individuals at high genetic risk, along with population-wide measures to lower BP, warrants further study.

Since the completion of our study, another BP GWAS using UK Biobank data has been published[15], as part of a larger single-stage

combined meta-analysis without replication; it reported a total of 316 loci, including 241 loci identified from the meta-analysis involving UK Biobank that were not tested for validation. Of the 107 validated loci reported in our study, 32 are discovered and validated for the first time in our analysis of UK Biobank. In addition, 75 sentinel SNVs are in LD ($r^2 \geq 0.2$) with the recently reported loci[15], and we validate at least 53 of these for the first time in our study (indicated by "GIU" in **Tables 1–3**). Furthermore, we note that 49 of the reported loci from the recent study[15] did not validate in our large independent replication resource.

In summary we describe 107 validated loci for BP, offering new biology, identifying potential new therapeutic targets and raising the possibility of a precision medicine approach to modify risk of hypertension and cardiovascular outcomes. Altogether, this represents a major advance in understanding of the genetic architecture of BP.

URLs. UK Biobank, https://www.ukbiobank.ac.uk/; genotype imputation and genetic association studies using UK Biobank data, http://biobank.ctsu.ox.ac.uk/crystal/refer.cgi?id=157020; UK Biobank Axiom array content summary, http://biobank.ctsu.ox.ac.uk/crystal/refer.cgi?id=146640; exome chip design, http://genome.sph.umich.edu/wiki/Exome_Chip_Design; Genotype-Tissue Expression (GTEx) database, http://www.gtexportal.org/; GREAT Enrichment, http://bejerano.stanford.edu/great; Ingenuity Pathway Analysis (IPA) software, http://www.qiagen.com/ingenuity; Chembl, http://www.ebi.ac.uk/chembl/; Drug Gene Interaction database, http://dgidb.genome.wustl.edu/; FORGE (accessed 16 August 2016), http://browser.1000genomes.org/Homo_sapiens/UserData/Forge?db=core; Fantom5 data (accessed 16 August 2016), http://fantom.gsc.riken.jp/5/; ENCODE DNase I data (wgEncodeAwgDnaseMasterSites; accessed 20 August 2016 using Table browser), http://hgdownload.cse.ucsc.edu/goldenpath/hg19/encodeDCC/wgEncodeAwgDnaseMasterSites/; ENCODE cell type data (accessed 20 August 2016), http://genome.ucsc.edu/ENCODE/cellTypes.html; Servier Medical Art, http://www.servier.fr/servier-medical-art.

METHODS

Methods, including statements of data availability and any associated accession codes and references, are available in the online version of the paper.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

H.R.W., C.P.C., M.R., M.R.B., P.B.M., M.B. and M.J.C. were funded by the National Institute for Health Research (NIHR) as part of the portfolio of translational research of the NIHR Biomedical Research Unit at Barts and The London School of Medicine and Dentistry. H.G. was funded by the NIHR Imperial College Health Care NHS Trust and Imperial College London Biomedical Research Centre. M.R. was a recipient of a grant from the China Scholarship Council (2011632047). B.M. holds an MRC eMedLab Medical Bioinformatics Career Development Fellowship, funded from award MR/L016311/1. J.M.M.H. was funded by the UK Medical Research Council (G0800270), British Heart Foundation (SP/09/002), UK National Institute for Health Research Cambridge Biomedical Research Centre, European Research Council (268834) and European Commission Framework Programme 7 (HEALTH-F2-2012-279233). B.K. holds a British Heart Foundation Personal Chair (CH/13/2/30154). N.J.S. holds a chair funded by the British Heart Foundation and is an NIHR Senior Investigator. F.D. was funded by the MRC Unit at the University of Bristol (MC_UU_12013/1-9). P. Surendran was funded by the UK Medical Research Council (G0800270). C.L. and A.K. were funded by NHLBI intramural funding. C.N.-C. was funded by the National Institutes of Health (HL113933, HL124262). P.v.d.H. was funded by ZonMw grant 90.700.441, Marie Sklodowska-Curie GF (call, H2020-MSCA-IF-2014; project ID, 661395). N.V. was supported by a Marie Sklodowska-Curie GF grant (661395) and ICIN-NHI. N.P. received funding from the UK National Institute for Health Research Biomedical Research Centre at Imperial College Healthcare NHS Trust and Imperial College London and also from his Senior Investigator Award. P. Sever was supported by the NIHR Biomedical Research Centre at Imperial College Healthcare NHS Trust and Imperial College London. S.T. was supported by the NIHR Biomedical Research Centre at Imperial College Healthcare NHS Trust and Imperial College London. P.F.O'R. received funding from the UK Medical Research Council (MR/N015746/1) and the Wellcome Trust (109863/Z/15/Z). I.K. was supported by the EU PhenoMeNal project (Horizon 2020, 654241). A.C. was funded by the National Institutes of Health (HL128782, HL086694). M.F. was supported by a Wellcome Trust core award (090532/Z/09/Z) and the BHF Centre of Research Excellence, Oxford (RE/13/1/30181). C.H. was funded by an MRC core grant for QTL in Health and Disease programme. Some of this work used the ALICE and SPECTRE High-Performance Computing Facilities at the University of Leicester. M.J.C. is a National Institute for Health Research (NIHR) senior investigator. P.E. is a National Institute for Health Research (NIHR) senior investigator and acknowledges support from the NIHR Biomedical Research Centre at Imperial College Healthcare NHS Trust and Imperial College London, and the NIHR Health Protection Research Unit in Health Impact of Environmental Hazards (HPRU-2012-10141). As director of the MRC-PHE Centre for Environment and Health, P.E. acknowledges support from the Medical Research Council and Public Health England (MR/L01341X/1). This work used the computing resources of the UK Medical Bioinformatics partnership–aggregation, integration, visualisation and analysis of large, complex data (UK MED-BIO), which is supported by the Medical Research Council (MR/L01632X/1). This research was supported by the British Heart Foundation (grant SP/13/2/30111). Project title: Large-Scale Comprehensive Genotyping of UK Biobank for Cardiometabolic Traits and Diseases: UK CardioMetabolic Consortium (UKCMC). This research has been conducted using the UK Biobank Resource under application number 236.

AUTHOR CONTRIBUTIONS

Central analysis: H.R.W., C.P.C., H.G., M.R.B., M.P.S.L., M.R., I.T., B.M., I.K., E.E. Writing of the manuscript: H.R.W., M.R.B., E.E., C.P.C., H.G., I.T., B.M., M.R., M.J.C., P.E. (with group leads, M.J.C., P.E.). Working group membership: M.J.C., H.R.W., E.E., I.T., P.B.M., L.V.W., N.J.S., M.T., J.M.M.H., M.D.T., I.N., B.K., H.G., M.R.B., C.P.C., J.S.K., P.E. (with co-chairs M.J.C., P.E.). Replication consortium contributor: (ICBP-1000G) G.B.E., L.V.W., D.L., A.C., M.J.C., M.D.T., P.F.O'R., J.K., H.S.; (CHD Exome+ Consortium) P. Surendran, R.C., D.S., J.M.M.H.; (ExomeBP Consortium) J.P.C., F.D., P.B.M.; (T2D-GENES Consortium and GoT2DGenes Consortium) C.M.L.; (CHARGE) G.B.E., C.L., A.T.K., D.L., C.N.-C., D.I.C.; (iGEN-BP) M.L., J.C.C., N.K., J.H., E.S.T., P.E., J.S.K., P.v.d.H. Replication study contributor: (Lifelines) N.V., P.v.d.H., H.S., M.A.S.; (GS:SFHS) J.M., C.H., D.P., S.P.; (EGCUT) T.E., M.A., R.M., A.M.; (PREVEND) P.v.d.H., N.V., R.T.G., S.J.L.B.; (ASCOT) H.R.W., M.J.C., P.B.M., P.S., N.P., A.S., D.S., S.T.; (BRIGHT) H.R.W., M.J.C., P.B.M., M.B., M.F., J.C.; (Airwave) H.G., E.E., M.P.S.L., I.K., I.T., P.E. All authors critically reviewed and approved the final version of the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare competing financial interests: details are available in the online version of the paper.

1. Muñoz, M. *et al.* Evaluating the contribution of genetics and familial shared environment to common disease using the UK Biobank. *Nat. Genet.* **48**, 980–983 (2016).
2. Feinleib, M. *et al.* The NHLBI twin study of cardiovascular disease risk factors: methodology and summary of results. *Am. J. Epidemiol.* **106**, 284–285 (1977).
3. Poulter, N.R., Prabhakaran, D. & Caulfield, M. Hypertension. *Lancet* **386**, 801–812 (2015).
4. Mongeau, J.G., Biron, P. & Sing, C.F. The influence of genetics and household environment upon the variability of normal blood pressure: the Montreal Adoption Survey. *Clin. Exp. Hypertens. A* **8**, 653–660 (1986).
5. Forouzanfar, M.H. *et al.* Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks in 188 countries, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet* **386**, 2287–2323 (2015).
6. Sundström, J. *et al.* Blood pressure–lowering treatment based on cardiovascular risk: a meta-analysis of individual patient data. *Lancet* **384**, 591–598 (2014).
7. Cabrera, C.P. *et al.* Exploring hypertension genome-wide association studies findings and impact on pathophysiology, pathways, and pharmacogenetics. *Wiley Interdiscip. Rev. Syst. Biol. Med.* **7**, 73–90 (2015).

8. Ehret, G.B. *et al.* The genetics of blood pressure regulation and its target organs from association studies in 342,415 individuals. *Nat. Genet.* **48**, 1171–1184 (2016).
9. Surendran, P. *et al.* Trans-ancestry meta-analyses identify rare and common variants associated with blood pressure and hypertension. *Nat. Genet.* **48**, 1151–1161 (2016).
10. Liu, C. *et al.* Meta-analysis identifies common and rare variants influencing blood pressure and overlapping with metabolic trait loci. *Nat. Genet.* **48**, 1162–1170 (2016).
11. Kato, N. *et al.* Trans-ancestry genome-wide association study identifies 12 genetic loci influencing blood pressure and implicates a role for DNA methylation. *Nat. Genet.* **47**, 1282–1293 (2015).
12. Elliott, P. & Peakman, T.C. The UK Biobank sample handling and storage protocol for the collection, processing and archiving of human blood and urine. *Int. J. Epidemiol.* **37**, 234–244 (2008).
13. Sudlow, C. *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.* **12**, e1001779 (2015).
14. Huang, J. *et al.* Improved imputation of low-frequency and rare variants using the UK10K haplotype reference panel. *Nat. Commun.* **6**, 8111 (2015).
15. Hoffmann, T.J. *et al.* Genome-wide association analyses using electronic health records identify new loci influencing blood pressure variation. *Nat. Genet.* **49**, 54–64 (2017).
16. Staley, J.R. *et al.* PhenoScanner: a database of human genotype–phenotype associations. *Bioinformatics* **32**, 3207–3209 (2016).
17. Ettehad, D. *et al.* Blood pressure lowering for prevention of cardiovascular disease and death: a systematic review and meta-analysis. *Lancet* **387**, 957–967 (2016).
18. Kato, N. *et al.* Meta-analysis of genome-wide association studies identifies common variants associated with blood pressure variation in east Asians. *Nat. Genet.* **43**, 531–538 (2011).
19. Munroe, P.B., Barnes, M.R. & Caulfield, M.J. Advances in blood pressure genomics. *Circ. Res.* **112**, 1365–1379 (2013).
20. den Hoed, M. *et al.* Identification of heart rate–associated loci and their effects on cardiac conduction and rhythm disorders. *Nat. Genet.* **45**, 621–631 (2013).
21. Hamilton, C.A., Brosnan, M.J., McIntyre, M., Graham, D. & Dominiczak, A.F. Superoxide excess in hypertension and aging: a common cause of endothelial dysfunction. *Hypertension* **37**, 529–534 (2001).
22. Shin, S.Y. *et al.* An atlas of genetic influences on human blood metabolites. *Nat. Genet.* **46**, 543–550 (2014).
23. Raffler, J. *et al.* Genome-wide association study with targeted and non-targeted NMR metabolomics identifies 15 novel loci of urinary human metabolic individuality. *PLoS Genet.* **11**, e1005487 (2015).
24. van Setten, J. *et al.* Genome-wide association study of coronary and aortic calcification implicates risk loci for coronary artery disease and myocardial infarction. *Atherosclerosis* **228**, 400–405 (2013).
25. McCarthy, J.J. *et al.* Large scale association analysis for identification of genes underlying premature coronary heart disease: cumulative perspective from analysis of 111 candidate genes. *J. Med. Genet.* **41**, 334–341 (2004).
26. van Meurs, J.B. *et al.* Common genetic loci influencing plasma homocysteine concentrations and their effect on risk of coronary artery disease. *Am. J. Clin. Nutr.* **98**, 668–676 (2013).
27. Pu, X. *et al.* ADAMTS7 cleavage and vascular smooth muscle cell migration is affected by a coronary-artery-disease-associated variant. *Am. J. Hum. Genet.* **92**, 366–374 (2013).
28. Rizzoni, D. & Agabiti-Rosei, E. Structural abnormalities of small resistance arteries in essential hypertension. *Intern. Emerg. Med.* **7**, 205–212 (2012).
29. Ray, R. *et al.* Endothelial Nox4 NADPH oxidase enhances vasodilatation and reduces blood pressure *in vivo*. *Arterioscler. Thromb. Vasc. Biol.* **31**, 1368–1376 (2011).
30. Touyz, R.M. & Montezano, A.C. Vascular Nox4: a multifarious NADPH oxidase. *Circ. Res.* **110**, 1159–1161 (2012).
31. Steppan, J., Barodka, V., Berkowitz, D.E. & Nyhan, D. Vascular stiffness and increased pulse pressure in the aging cardiovascular system. *Cardiol. Res. Pract.* **2011**, 263585 (2011).
32. Yan, F. *et al.* Nox4 and redox signaling mediate TGF-β-induced endothelial cell apoptosis and phenotypic switch. *Cell Death Dis.* **5**, e1010 (2014).
33. Chan, E.C. *et al.* Nox4 modulates collagen production stimulated by transforming growth factor β1 *in vivo* and *in vitro*. *Biochem. Biophys. Res. Commun.* **430**, 918–925 (2013).
34. Vasa-Nicotera, M. *et al.* miR-146a is modulated in human endothelial cell with aging. *Atherosclerosis* **217**, 326–330 (2011).
35. Tian, X. *et al.* Phosphodiesterase 10A upregulation contributes to pulmonary vascular remodeling. *PLoS One* **6**, e18136 (2011).
36. Takimoto, E. *et al.* Chronic inhibition of cyclic GMP phosphodiesterase 5A prevents and reverses cardiac hypertrophy. *Nat. Med.* **11**, 214–222 (2005).
37. Pérez, N.G. *et al.* Phosphodiesterase 5A inhibition induces Na^+/H^+ exchanger blockade and protection against myocardial infarction. *Hypertension* **49**, 1095–1103 (2007).
38. Oliver, J.J., Melville, V.P. & Webb, D.J. Effect of regular phosphodiesterase type 5 inhibition in hypertension. *Hypertension* **48**, 622–627 (2006).
39. Levy, D. *et al.* Genome-wide association study of blood pressure and hypertension. *Nat. Genet.* **41**, 677–687 (2009).
40. Newton-Cheh, C. *et al.* Genome-wide association study identifies eight loci associated with blood pressure. *Nat. Genet.* **41**, 666–676 (2009).
41. DeStefano, A.L. *et al.* Autosomal dominant orthostatic hypotensive disorder maps to chromosome 18q. *Am. J. Hum. Genet.* **63**, 1425–1430 (1998).
42. Hong, X. *et al.* Genetic polymorphisms of the urea transporter gene are associated with antihypertensive response to nifedipine GITS. *Methods Find. Exp. Clin. Pharmacol.* **29**, 3–10 (2007).
43. Takimoto, E. *et al.* Sodium calcium exchanger plays a key role in alteration of cardiac function in response to pressure overload. *FASEB J.* **16**, 373–378 (2002).
44. Ronaldson, P.T. & Davis, T.P. Targeting transporters: promoting blood–brain barrier repair in response to oxidative stress injury. *Brain Res.* **1623**, 39–52 (2015).
45. Carta, L. *et al.* Fibrillins 1 and 2 perform partially overlapping functions during aortic development. *J. Biol. Chem.* **281**, 8016–8023 (2006).
46. Kazenwadel, J. *et al.* Loss-of-function germline *GATA2* mutations in patients with MDS/AML or MonoMAC syndrome and primary lymphedema reveal a key role for GATA2 in the lymphatic vasculature. *Blood* **119**, 1283–1291 (2012).
47. Akashi, M., Higashi, T., Masuda, S., Komori, T. & Furuse, M. A coronary artery disease–associated gene product, JCAD/KIAA1462, is a novel component of endothelial cell–cell junctions. *Biochem. Biophys. Res. Commun.* **413**, 224–229 (2011).
48. Cakstina, I. *et al.* Primary culture of avian embryonic heart forming region cells to study the regulation of vertebrate early heart morphogenesis by vitamin A. *BMC Dev. Biol.* **14**, 10 (2014).
49. Wang, J., Karra, R., Dickson, A.L. & Poss, K.D. Fibronectin is deposited by injury-activated epicardial cells and is necessary for zebrafish heart regeneration. *Dev. Biol.* **382**, 427–435 (2013).
50. Dietrich, T. *et al.* ED-B fibronectin (ED-B) can be targeted using a novel single chain antibody conjugate and is associated with macrophage accumulation in atherosclerotic lesions. *Basic Res. Cardiol.* **102**, 298–307 (2007).
51. Stoynev, N. *et al.* Gene expression in peripheral blood of patients with hypertension and patients with type 2 diabetes. *J. Cardiovasc. Med. (Hagerstown)* **15**, 702–709 (2014).
52. Erdos, B., Backes, I., McCowan, M.L., Hayward, L.F. & Scheuer, D.A. Brain-derived neurotrophic factor modulates angiotensin signaling in the hypothalamus to increase blood pressure in rats. *Am. J. Physiol. Heart Circ. Physiol.* **308**, H612–H622 (2015).
53. Chan, S.H., Wu, C.W., Chang, A.Y., Hsu, K.S. & Chan, J.Y. Transcriptional upregulation of brain-derived neurotrophic factor in rostral ventrolateral medulla by angiotensin II: significance in superoxide homeostasis and neural regulation of arterial pressure. *Circ. Res.* **107**, 1127–1139 (2010).
54. Crespo, K., Ménard, A. & Deng, A.Y. Retinoblastoma-associated protein 140 as a candidate for a novel etiological gene to hypertension. *Clin. Exp. Hypertens.* **38**, 533–540 (2016).
55. Watanabe, Y. *et al.* Accumulation of common polymorphisms is associated with development of hypertension: a 12-year follow-up from the Ohasama study. *Hypertens. Res.* **33**, 129–134 (2010).
56. Gale, D.P., Harten, S.K., Reid, C.D., Tuddenham, E.G. & Maxwell, P.H. Autosomal dominant erythrocytosis and pulmonary arterial hypertension associated with an activating HIF2α mutation. *Blood* **112**, 919–921 (2008).
57. Yndestad, A. *et al.* Elevated levels of activin A in clinical and experimental pulmonary hypertension. *J. Appl. Physiol.* **106**, 1356–1364 (2009).
58. Sacks, F.M. *et al.* Effects on blood pressure of reduced dietary sodium and the Dietary Approaches to Stop Hypertension (DASH) diet. *N. Engl. J. Med.* **344**, 3–10 (2001).
59. Intersalt Cooperative Research Group. Intersalt: an international study of electrolyte excretion and blood pressure. Results for 24 hour urinary sodium and potassium excretion. *Br. Med. J.* **297**, 319–328 (1988).
60. Whelton, P.K. *et al.* Primary prevention of hypertension: clinical and public health advisory from The National High Blood Pressure Education Program. *J. Am. Med. Assoc.* **288**, 1882–1888 (2002).
61. Chan, Q. *et al.* An update on nutrients and blood pressure. *J. Atheroscler. Thromb.* **23**, 276–289 (2016).
62. Khera, A.V. *et al.* Genetic risk, adherence to a healthy lifestyle, and coronary disease. *N. Engl. J. Med.* **375**, 2349–2358 (2016).

Helen R Warren[1,2,60], Evangelos Evangelou[3,4,60], Claudia P Cabrera[1,2,60], He Gao[3,5,60], Meixia Ren[1,2,60], Borbala Mifsud[1,60], Ioanna Ntalla[1], Praveen Surendran[6], Chunyu Liu[7–9], James P Cook[10], Aldi T Kraja[11], Fotios Drenos[12,13], Marie Loh[3,14], Niek Verweij[15–18], Jonathan Marten[19], Ibrahim Karaman[3], Marcelo P Segura Lepe[3,20], Paul F O'Reilly[21], Joanne Knight[22], Harold Snieder[23], Norihiro Kato[24], Jiang He[25],

E Shyong Tai[26,27], M Abdullah Said[15], David Porteous[28], Maris Alver[29], Neil Poulter[30], Martin Farrall[31], Ron T Gansevoort[32], Sandosh Padmanabhan[33], Reedik Mägi[29], Alice Stanton[34], John Connell[35], Stephan J L Bakker[36], Andres Metspalu[29], Denis C Shields[37], Simon Thom[38], Morris Brown[1,2], Peter Sever[38], Tõnu Esko[16,29], Caroline Hayward[19], Pim van der Harst[15], Danish Saleheen[39–41], Rajiv Chowdhury[6], John C Chambers[3,42–44], Daniel I Chasman[45,46], Aravinda Chakravarti[47], Christopher Newton-Cheh[16–18], Cecilia M Lindgren[16,48,49], Daniel Levy[7,9], Jaspal S Kooner[43,50,51], Bernard Keavney[52,53], Maciej Tomaszewski[52,53], Nilesh J Samani[54,55], Joanna M M Howson[6], Martin D Tobin[56], Patricia B Munroe[1,2], Georg B Ehret[47,57], Louise V Wain[56], The International Consortium of Blood Pressure (ICBP) 1000G Analyses[58], The CHD Exome+ Consortium[59], The ExomeBP Consortium[59], The T2D-GENES Consortium[59], The GoT2DGenes Consortium[59], The Cohorts for Heart and Ageing Research in Genome Epidemiology (CHARGE) BP Exome Consortium[59], The International Genomics of Blood Pressure (iGEN-BP) Consortium[59], Michael R Barnes[1,2,60], Ioanna Tzoulaki[3–5,60], Mark J Caulfield[1,2,60] & Paul Elliott[3,5,60] for The UK Biobank CardioMetabolic Consortium BP working group

[1]William Harvey Research Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, UK. [2]National Institute for Health Research Barts Cardiovascular Biomedical Research Unit, Queen Mary University of London, London, UK. [3]Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, London, UK. [4]Department of Hygiene and Epidemiology, University of Ioannina Medical School, Ioannina, Greece. [5]MRC-PHE Centre for Environment and Health, Imperial College London, London, UK. [6]Cardiovascular Epidemiology Unit, Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK. [7]Population Sciences Branch, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, Maryland, USA. [8]Boston University School of Public Health, Boston, Massachusetts, USA. [9]National Heart, Lung, and Blood Institute's Framingham Heart Study, Framingham, Massachusetts, USA. [10]Department of Biostatistics, University of Liverpool, Liverpool, UK. [11]Division of Statistical Genomics, Department of Genetics and Center for Genome Sciences and Systems Biology, Washington University School of Medicine, St. Louis, Missouri, USA. [12]MRC Integrative Epidemiology Unit, School of Social and Community Medicine, University of Bristol, Bristol, UK. [13]Centre for Cardiovascular Genetics, Institute of Cardiovascular Science, University College London, London, UK. [14]Translational Laboratory in Genetic Medicine, Agency for Science, Technology and Research (A*STAR), Singapore. [15]University of Groningen, University Medical Center Groningen, Department of Cardiology, Groningen, the Netherlands. [16]Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, Cambridge, Massachusetts, USA. [17]Center for Human Genetic Research, Massachusetts General Hospital, Boston, Massachusetts, USA. [18]Cardiovascular Research Center, Massachusetts General Hospital, Boston, Massachusetts, USA. [19]MRC Human Genetics Unit, MRC Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, UK. [20]Bayer Pharma, Berlin, Germany. [21]Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK. [22]Data Science Institute, Lancester University, Lancaster, UK. [23]University of Groningen, University Medical Center Groningen, Department of Epidemiology, Groningen, the Netherlands. [24]Department of Gene Diagnostics and Therapeutics, Research Institute, National Center for Global Health and Medicine, Tokyo, Japan. [25]Department of Epidemiology, Tulane University School of Public Health and Tropical Medicine, New Orleans, Louisiana, USA. [26]Saw Swee Hock School of Public Health, National University of Singapore and National University Health System, Singapore. [27]Department of Medicine, Yong Loo Lin School of Medicine, National University of Singapore, Singapore. [28]Centre for Genomic and Experimental Medicine, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, UK. [29]Estonian Genome Center, University of Tartu, Tartu, Estonia. [30]Imperial Clinical Trials Unit, School of Public Health, Imperial College London, London, UK. [31]Division of Cardiovascular Medicine, Radcliffe Department of Medicine, Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK. [32]University of Groningen, University Medical Center Groningen, Department of Nephrology, Groningen, the Netherlands. [33]Institute of Cardiovascular and Medical Sciences, University of Glasgow, Glasgow, UK. [34]Molecular and Cellular Therapeutics, Royal College of Surgeons in Ireland, Dublin, Ireland. [35]Ninewells Hospital and Medical School, University of Dundee, Dundee, UK. [36]University of Groningen, University Medical Center Groningen, Department of Internal Medicine, Groningen, the Netherlands. [37]School of Medicine, Conway Institute, University College Dublin, Dublin, Ireland. [38]National Heart and Lung Institute, Imperial College London, London, UK. [39]Department of Biostatistics and Epidemiology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, USA. [40]Centre for Non-Communicable Diseases, Karachi, Pakistan. [41]Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK. [42]Ealing Hospital National Health Service (NHS) Trust, Southall, UK. [43]Imperial College Healthcare NHS Trust, London, UK. [44]Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore. [45]Division of Preventive Medicine, Brigham and Women's Hospital, Boston, Massachusetts, USA. [46]Harvard Medical School, Boston, Massachusetts, USA. [47]Center for Complex Disease Genomics, McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA. [48]Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK. [49]Big Data Institute at the Li Ka Shing Centre for Health Information and Discovery, University of Oxford, Oxford, UK. [50]Department of Cardiology, Ealing Hospital NHS Trust, Southall, UK. [51]National Heart and Lung Institute, Cardiovascular Sciences, Hammersmith Campus, Imperial College London, London, UK. [52]Division of Cardiovascular Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, UK. [53]Division of Medicine, Central Manchester NHS Foundation Trust, Manchester Academic Health Science Centre, Manchester, UK. [54]Department of Cardiovascular Sciences, University of Leicester, BHF Cardiovascular Research Centre, Glenfield Hospital, Leicester, UK. [55]NIHR Leicester Cardiovascular Biomedical Research Unit, Glenfield Hospital, Leicester, UK. [56]Department of Health Sciences, University of Leicester, Leicester, UK. [57]Cardiology, Department of Medicine, Geneva University Hospital, Geneva, Switzerland. [58]A full list of members and affiliations appears at the end of the paper. [59]A full list of members and affiliations appears in the **Supplementary Note**. [60]These authors contributed equally to this work. Correspondence should be addressed to P.E. (p.elliott@imperial.ac.uk) or M.J.C. (m.j.caulfield@qmul.ac.uk).

The International Consortium of Blood Pressure (ICBP) 1000G Analyses

Louise V Wain[56], Ahmad Vaez[23,61], Rick Jansen[62], Roby Joehanes[9,63], Peter J van der Most[23], A Mesut Erzurumluoglu[56], Paul O'Reilly[21], Claudia P Cabrera[1,2], Helen R Warren[1,2], Lynda M Rose[45], Germaine C Verwoert[64], Jouke-Jan Hottenga[65], Rona J Strawbridge[66,67], Tonu Esko[29,68,69], Dan E Arking[47], Shih-Jen Hwang[70,71], Xiuqing Guo[72], Zoltan Kutalik[73,74], Stella Trompet[75,76], Nick Shrine[56], Alexander Teumer[77,78], Janina S Ried[79], Joshua C Bis[80], Albert V Smith[81,82], Najaf Amin[83], Ilja M Nolte[23], Leo-Pekka Lyytikäinen[84,85], Anubha Mahajan[48], Nicholas J Wareham[86], Edith Hofer[87,88], Peter K Joshi[89], Kati Kristiansson[90], Michela Traglia[91], Aki S Havulinna[90], Anuj Goel[48,92], Mike A Nalls[93,94], Siim Sõber[95], Dragana Vuckovic[96,97], Jian'an Luan[86],

Fabiola Del Greco M[98], Kristin L Ayers[99], Jaume Marrugat[100], Daniela Ruggiero[101], Lorna M Lopez[102–104], Teemu Niiranen[90], Stefan Enroth[105], Anne U Jackson[106], Christopher P Nelson[54,55], Jennifer E Huffman[19], Weihua Zhang[3,50], Jonathan Marten[19], Ilaria Gandin[97], Sarah E Harris[28,102], Tatijana Zemonik[107], Yingchang Lu[108], Evangelos Evangelou[3,4], Nabi Shah[109,110], Martin H de Borst[36], Massimo Mangino[111,112], Bram P Prins[113], Archie Campbell[28,114], Ruifang Li-Gao[115], Ganesh Chauhan[116,117], Christopher Oldmeadow[118], Gonçalo Abecasis[119], Maryam Abedi[120], Caterina M Barbieri[91], Michael R Barnes[1,2], Chiara Batini[56], BIOS Consortium , Tineka Blake[56], Michael Boehnke[106], Erwin P Bottinger[108], Peter S Braund[54,55], Morris Brown[1,2], Marco Brumat[97], Harry Campbell[89], John C Chambers[3,43,50], Massimiliano Cocca[97], Francis Collins[121], John Connell[35], Heather J Cordell[122], Jeffrey J Damman[123], Gail Davies[102,124], Eco J de Geus[65], Renée de Mutsert[115], Joris Deelen[125], Yusuf Demirkale[126], Alex S F Doney[109], Marcus Dörr[78,127], Martin Farrall[48,92], Teresa Ferreira[48], Mattias Frånberg[66,67,128], He Gao[3], Vilmantas Giedraitis[129], Christian Gieger[130], Franco Giulianini[45], Alan J Gow[102,131], Anders Hamsten[66,67], Tamara B Harris[132], Albert Hofman[64,133], Elizabeth G Holliday[118], Marjo-Riitta Jarvelin[5,134–136], Åsa Johansson[105], Andrew D Johnson[9,137], Pekka Jousilahti[90], Antti Jula[90], Mika Kähönen[138,139], Sekar Kathiresan[16–18], Kay-Tee Khaw[41], Ivana Kolcic[140], Seppo Koskinen[90], Claudia Langenberg[86], Marty Larson[9], Lenore J Launer[132], Benjamin Lehne[3], David C M Liewald[102,124], Lifelines Cohort Study, Li Lin[57], Lars Lind[141], François Mach[57], Chrysovalanto Mamasoula[142], Cristina Menni[111], Borbala Mifsud[1], Yuri Milaneschi[143], Anna Morgan[97], Andrew D Morris[144], Alanna C Morrison[145], Peter J Munson[126], Priyanka Nandakumar[47], Quang Tri Nguyen[126], Teresa Nutile[101], Albertine J Oldehinkel[146], Ben A Oostra[83], Elin Org[29], Sandosh Padmanabhan[114,147], Aarno Palotie[148], Guillaume Paré[149], Alison Pattie[124], Brenda W J H Penninx[143], Neil Poulter[30], Peter P Pramstaller[98,150,151], Olli T Raitakari[152,153], Meixia Ren[1,154], Kenneth Rice[155], Paul M Ridker[45,46], Harriëtte Riese[146], Samuli Ripatti[148], Antonietta Robino[156], Jerome I Rotter[72,157], Igor Rudan[89], Yasaman Saba[158], Aude Saint Pierre[98,159], Cinzia F Sala[91], Antti-Pekka Sarin[148], Reinhold Schmidt[87], Rodney Scott[118,160,161], Marc A Seelen[36], Denis C Shields[37], David Siscovick[162], Rossella Sorice[101,163], Alice Stanton[34], David J Stott[33], Johan Sundström[141], Morris Swertz[164], Kent D Taylor[165,166], Simon Thom[167], Ioanna Tzoulaki[3], Christophe Tzourio[116,117,168], André G Uitterlinden[64,169], Understanding Society Scientific group, Uwe Völker[78,170], Peter Vollenweider[171], Sarah Wild[89], Gonneke Willemsen[65], Alan F Wright[19], Jie Yao[72], Sébastien Thériault[149], David Conen[172], Attia John[118,160,161], Peter Sever[38], Stéphanie Debette[116,117,173], Dennis O Mook-Kanamori[115,174], Eleftheria Zeggini[113], Tim D Spector[111], Pim van der Harst[15], Colin N A Palmer[109], Anne-Claire Vergnaud[3], Ruth J F Loos[87,108,175], Ozren Polasek[140], John M Starr[102,176], Giorgia Girotto[96,97], Caroline Hayward[19,114], Jaspal S Kooner[38,43,50], Cecila M Lindgren[48,69], Veronique Vitart[19], Nilesh J Samani[54,55], Jaakko Tuomilehto[177–180], Ulf Gyllensten[105], Paul Knekt[90], Ian J Deary[102,124], Marina Ciullo[101,163], Roberto Elosua[100], Bernard D Keavney[52], Andrew A Hicks[98], Robert A Scott[86], Paolo Gasparini[96,97], Maris Laan[95,181], YongMei Liu[182], Hugh Watkins[48,92], Catharina A Hartman[146], Veikko Salomaa[90], Daniela Toniolo[91], Markus Perola[89,148,183], James F Wilson[19,89], Helena Schmidt[158,184], Jing Hua Zhao[86], Terho Lehtimäki[84,85], Cornelia M van Duijn[83], Vilmundur Gudnason[81,82], Bruce M Psaty[80,185–187], Annette Peters[79], Rainer Rettig[188], Alan James[189,190], J Wouter Jukema[75], David P Strachan[191], Walter Palmas[192], Andres Metspalu[29], Erik Ingelsson[193,194], Dorret I Boomsma[65], Oscar H Franco[64], Murielle Bochud[73], Christopher Newton-Cheh[16–18,69], Patricia B Munroe[1,2], Paul Elliott[5], Daniel I Chasman[45,46], Aravinda Chakravarti[47], Joanne Knight[22], Andrew P Morris[10,48], Daniel Levy[7,71], Martin D Tobin[56], Harold Snieder[23], Mark J Caulfield[1,2] & Georg B Ehret[47,57]

[61]Research Institute for Primordial Prevention of Non-Communicable Disease, Isfahan University of Medical Sciences, Isfahan, Iran. [62]Department of Psychiatry, VU University Medical Center, Neuroscience Campus Amsterdam, Amsterdam, the Netherlands. [63]Hebrew SeniorLife, Harvard Medical School, Boston, Massachusetts, USA. [64]Department of Epidemiology, Erasmus MC, Rotterdam, the Netherlands. [65]Department of Biological Psychology, Vrije Universiteit, Amsterdam, EMGO+ institute, VU University Medical Center, Amsterdam, the Netherlands. [66]Cardiovascular Medicine Unit, Department of Medicine Solna, Karolinska Institutet, Stockholm, Sweden. [67]Centre for Molecular Medicine, Karolinska Universitetsjukhuset, Solna, Sweden. [68]Division of Endocrinology, Children's Hospital, Boston, Massachusetts, USA. [69]Broad Institute of Harvard and MIT, Cambridge, Massachusetts, USA. [70]Population Science Branch, Division of Intramural Research, National Heart, Lung, and Blood Institute, US National Institutes of Health, Bethesda, Maryland, USA. [71]Framingham Heart Study, Framingham, Massachusetts, USA. [72]Institute for Translational Genomics and Population Sciences, Department of Pediatrics, LABioMed at Harbor-UCLA Medical Center, Torrance, California, USA. [73]Institute of Social and Preventive Medicine, Lausanne University Hospital, Lausanne, Switzerland. [74]Swiss Institute of Bioinformatics, Lausanne, Switzerland. [75]Department of Cardiology, Leiden University Medical Center, Leiden, the Netherlands. [76]Department of Gerontology and Geriatrics, Leiden University Medical Center, Leiden, the Netherlands. [77]Institute for Community Medicine, University Medicine Greifswald, Greifswald, Germany. [78]DZHK (German Centre for Cardiovascular Research), partner site Greifswald, Greifswald, Germany. [79]Institute of Epidemiology II, Helmholtz Zentrum München, Neuherberg, Germany.

[80]Cardiovascular Health Research Unit, Department of Medicine, University of Washington, Seattle, Washington, USA. [81]Icelandic Heart Association, Kopavogur, Iceland. [82]Faculty of Medicine, University of Iceland, Reykjavik, Iceland. [83]Genetic Epidemiology Unit, Department of Epidemiology, Erasmus MC, Rotterdam, the Netherlands. [84]Department of Clinical Chemistry, Fimlab Laboratories, Tampere, Finland. [85]Department of Clinical Chemistry, University of Tampere School of Medicine, Tampere, Finland. [86]MRC Epidemiology Unit, University of Cambridge School of Clinical Medicine, Institute of Metabolic Science, Cambridge Biomedical Campus, Cambridge, UK. [87]Clinical Division of Neurogeriatrics, Department of Neurology, Medical University Graz, Graz, Austria. [88]Institute of Medical Informatics, Statistics and Documentation, Medical University Graz, Graz, Austria. [89]Centre for Global Health Research, Usher Institute of Population Health Sciences and Informatics, University of Edinburgh, Edinburgh, UK. [90]Department of Health, National Institute for Health and Welfare (THL), Helsinki, Finland. [91]Division of Genetics and Cell Biology, San Raffaele Scientific Institute, Milan, Italy. [92]Division of Cardiovascular Medicine, Radcliffe Department of Medicine, University of Oxford, Oxford, UK. [93]Laboratory of Neurogenetics, National Institute on Aging, US National Institutes of Health, Bethesda, Maryland, USA. [94]Kelly Services, Rockville, Maryland, USA. [95]Human Molecular Genetics Research Group, Institute of Molecular and Cell Biology, University of Tartu, Tartu, Estonia. [96]Experimental Genetics Division, Sidra Medical and Research Center, Doha, Qatar. [97]Department of Medical, Surgical and Health Sciences, University of Trieste, Trieste, Italy. [98]Center for Biomedicine, European Academy Bozen/Bolzano (EURAC), Bolzano, Italy (affiliated institute of the University of Lübeck). [99]Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, New York, USA. [100]Cardiovascular Epidemiology and Genetics, IMIM, Barcelona, Spain. [101]Institute of Genetics and Biophysics A. Buzzati-Traverso, CNR, Naples, Italy. [102]Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, Edinburgh, UK. [103]Department of Psychiatry, Royal College of Surgeons in Ireland, Education and Research Centre, Beaumont Hospital, Dublin, Ireland. [104]University College Dublin, UCD Conway Institute, Centre for Proteome Research, UCD, Belfield, Dublin, Ireland. [105]Department of Immunology, Genetics and Pathology, Uppsala Universitet, Science for Life Laboratory, Uppsala, Sweden. [106]Department of Biostatistics and Center for Statistical Genetics, University of Michigan, Ann Arbor, Michigan, USA. [107]Department of Biology, Faculty of Medicine, University of Split, Split, Croatia. [108]Charles Bronfman Institute for Personalized Medicine, Icachn School of Medicine at Mount Sinai, New York, New York, USA. [109]Medical Research Institute, University of Dundee, Ninewells Hospital and Medical School, Dundee, UK. [110]Department of Pharmacy, COMSATS Institute of Information Technology, Abbottabad, Pakistan. [111]Department of Twin Research and Genetic Epidemiology, King's College London, London, UK. [112]National Institute for Health Research Biomedical Research Centre, London, UK. [113]Department of Human Genetics, Wellcome Trust Sanger Institute, Cambridge, UK. [114]Generation Scotland, Centre for Genomic and Experimental Medicine, University of Edinburgh, Edinburgh, UK. [115]Department of Clinical Epidemiology, Leiden University Medical Center, Leiden, the Netherlands. [116]INSERM U1219, Bordeaux Population Health Center, Bordeaux, France. [117]Bordeaux University, Bordeaux, France. [118]Hunter Medical Research Institute, New Lambton, New South Wales, Australia. [119]Center for Statistical Genetics, Department of Biostatistics, SPH II, Ann Arbor, Michigan, USA. [120]Department of Genetics and Molecular Biology, Isfahan University of Medical Sciences, Isfahan, Iran. [121]Medical Genomics and Metabolic Genetics Branch, National Human Genome Research Institute, US National Institutes of Health, Bethesda, Maryland, USA. [122]Institute of Genetic Medicine, Newcastle University, Newcastle-upon-Tyne, UK. [123]Department of Pathology, Amsterdam Medical Center, Amsterdam, the Netherlands. [124]Department of Psychology, University of Edinburgh, Edinburgh, UK. [125]Department of Molecular Epidemiology, Leiden University Medical Center, Leiden, the Netherlands. [126]Center for Information Technology, US National Institutes of Health, Bethesda, Maryland, USA. [127]Department of Internal Medicine B, University Medicine Greifswald, Greifswald, Germany. [128]Department of Numerical Analysis and Computer Science, Stockholm University, Stockholm, Sweden. [129]Department of Public Health and Caring Sciences, Geriatrics, Uppsala, Sweden. [130]Helmholtz Zentrum München, Deutsches Forschungszentrum für Gesundheit und Umwelt, Neuherberg, Germany. [131]Department of Psychology, School of Life Sciences, Heriot-Watt University, Edinburgh, UK. [132]Intramural Research Program, Laboratory of Epidemiology, Demography, and Biometry, National Institute on Aging, Bethesda, Maryland, USA. [133]Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, Massachusetts, USA. [134]Center For Life-Course Health Research, University of Oulu, Oulu, Finland. [135]Biocenter Oulu, University of Oulu, Oulu, Finland. [136]Unit of Primary Care, Oulu University Hospital, Oulu, Finland. [137]National Heart, Lung, and Blood Institute, Cardiovascular Epidemiology and Human Genomics Branch, Bethesda, Maryland, USA. [138]Department of Clinical Physiology, Tampere University Hospital, Tampere, Finland. [139]Department of Clinical Physiology, University of Tampere School of Medicine, Tampere, Finland. [140]Department of Public Health, Faculty of Medicine, University of Split, Split, Croatia. [141]Department of Medical Sciences, Cardiovascular Epidemiology, Uppsala University, Uppsala, Sweden. [142]Institute of Health and Society, Newcastle University, Newcastle-upon-Tyne, UK. [143]Department of Psychiatry, EMGO Institute for Health and Care Research, VU University Medical Center, Amsterdam, the Netherlands. [144]School of Molecular, Genetic and Population Health Sciences, University of Edinburgh, Medical School, Edinburgh, UK. [145]Department of Epidemiology, Human Genetics and Environmental Sciences, School of Public Health, University of Texas Health Science Center at Houston, Houston, Texas, USA. [146]University of Groningen, University Medical Center Groningen, Interdisciplinary Center Psychopathology and Emotion Regulation (IPCE), Groningen, the Netherlands. [147]British Heart Foundation Glasgow Cardiovascular Research Centre, Institute of Cardiovascular and Medical Sciences, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow, UK. [148]Institute for Molecular Medicine Finland (FIMM), University of Helsinki, Helsinki, Finland. [149]Department of Pathology and Molecular Medicine, McMaster University, Hamilton, Ontario, Canada. [150]Department of Neurology, General Central Hospital, Bolzano, Italy. [151]Department of Neurology, University of Lübeck, Lübeck, Germany. [152]Department of Clinical Physiology and Nuclear Medicine, Turku University Hospital, Turku, Finland. [153]Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, Turku, Finland. [154]Department of Cardiology, Fujian Provincial Hospital, Fujian Medical University, Fuzhou, China. [155]Department of Biostatistics, University of Washington, Seattle, Washington, USA. [156]Institute for Maternal and Child Health, IRCCS Burlo Garofolo, Trieste, Italy. [157]Institute for Translational Genomics and Population Sciences, Department of Medicine, LABioMed at Harbor-UCLA Medical Center, Torrance, California, USA. [158]Institute of Molecular Biology and Biochemistry, Centre for Molecular Medicine, Medical University of Graz, Graz, Austria. [159]INSERM U1078, Etablissement Français du Sang, Brest, France. [160]Faculty of Health, University of Newcastle, Callaghan, New South Wales, Australia. [161]School of Biomedical Sciences and Pharmacy, John Hunter Hospital, New Lambton, New South Wales, Australia. [162]New York Academy of Medicine, New York, New York, USA. [163]IRCCS Neuromed, Pozzilli, Italy. [164]University of Groningen, University Medical Center Groningen, Department of Genetics, Groningen, the Netherlands. [165]Institute for Translational Genomics and Population Sciences, Los Angeles BioMedical Research Institute at Harbor-UCLA Medical Center, Torrance, California, USA. [166]Division of Genetic Outcomes, Department of Pediatrics, Harbor-UCLA Medical Center, Torrance, California, USA. [167]International Centre for Circulatory Health, Imperial College London, London, UK. [168]Department of Public Health, Bordeaux University Hospital, Bordeaux, France. [169]Department of Internal Medicine, Erasmus MC, Rotterdam, the Netherlands. [170]Interfaculty Institute for Genetics and Functional Genomics, University Medicine Greifswald, Greifswald, Germany. [171]Department of Internal Medicine, Lausanne University Hospital, CHUV, Lausanne, Switzerland. [172]Population Health Research Institute, McMaster University, Hamilton, Ontario, Canada. [173]Department of Neurology, Bordeaux University Hospital, Bordeaux, France. [174]Department of Public Health and Primary Care, Leiden University Medical Center, Leiden, the Netherlands. [175]Mindich Child Health Development Institute, Icahn School of Medicine at Mount Sinai, New York, New York, USA. [176]Alzheimer Scotland Dementia Research Centre, University of Edinburgh, Edinburgh, UK. [177]Diabetes Prevention Unit, National Institute for Health and Welfare, Helsinki, Finland. [178]South Ostrobothnia Central Hospital, Seinäjoki, Finland. [179]Red RECAVA Grupo RD06/0014/0015, Hospital Universitario La Paz, Madrid, Spain. [180]Centre for Vascular Prevention, Danube University Krems, Krems, Austria. [181]Institute of Biomedicine and Translational Medicine, University of Tartu, Tartu, Estonia. [182]Division of Public Health Sciences, Wake Forest School of Medicine, Winston-Salem, USA. [183]Estonian Genome Center, University of Tartu, Tartu, Estonia. [184]Department of Neurology, Medical University Graz, Graz, Austria. [185]Department of Epidemiology University of Washington, Seattle, Washington, USA. [186]Department of Health Services, University of Washington, Seattle, Washington, USA. [187]Group Health Research Institute, Group Health, Seattle, Washington, USA. [188]Institute of Physiology, University Medicine Greifswald, Karlsburg, Germany. [189]Department of Pulmonary Physiology and Sleep, Sir Charles Gairdner Hospital, Nedlands, Western Australia, Australia. [190]School of Medicine and Pharmacology, University of Western Australia, Crawley, Western Australia, Australia. [191]Population Health Research Institute, St George's, University of London, London, UK. [192]Department of Medicine, Columbia University Medical Center, New York, New York, USA. [193]Department of Medical Sciences, Molecular Epidemiology and Science for Life Laboratory, Uppsala University, Uppsala, Sweden. [194]Division of Cardiovascular Medicine, Department of Medicine, Stanford University School of Medicine, Stanford, California, USA.

ONLINE METHODS

UK Biobank data. Our GWAS analysis is performed using data from the interim release of the first ~150,000 UK Biobank (UKB) participants (**Supplementary Note**): ~100,000 individuals from UK Biobank genotyped at ~800,000 SNVs with a custom Affymetrix UK Biobank Axiom array and ~50,000 individuals genotyped with a custom Affymetrix UK BiLEVE Axiom array from the UK BiLEVE study[63], a subset of UKB. SNVs were imputed centrally by UKB using a merged UK10K sequencing + 1000 Genomes imputation reference panel. Information on UK Biobank array design and protocols is available on the UK Biobank website.

Quality control. Following quality control procedures already carried out centrally by UKB, we excluded discordant SNVs and samples with quality control failures, sex discordance and high heterozygosity/missingness. We further restricted our data to a subset of individuals of European ancestry. By applying *k*-means clustering to the principal-component analysis (PCA) data, a total of N = 145,315 Europeans remained (**Supplementary Fig. 15**). We used kinship data to exclude first- and second-degree relatives, with N = 141,647 unrelated individuals remaining. Finally, we restricted our data to non-pregnant individuals with two automated BP measurements available, resulting in a maximum of N = 140,886 individuals for analysis (**Supplementary Note**).

Phenotypic data. After calculating the mean SBP and DBP values from the two BP measurements, we adjusted for medication use by adding 15 and 10 mm Hg to SBP and DBP, respectively, for individuals reported to be taking BP-lowering medication (21.4% of individuals)[64]. PP was calculated as SBP minus DBP, according to the medication-adjusted traits. Hypertension, used in secondary analyses, was defined as (i) SBP ≥140 mm Hg, (ii) DBP ≥90 mm Hg, or (iii) taking BP-lowering medication; otherwise, individuals were classified as non-hypertensive. Descriptive summary statistics are provided for all individuals (**Supplementary Table 1**).

Statistical methods. The statistical approaches used for the discovery and replication of loci are reported in detail below. We also describe methods used for identification of secondary signals; lookups in non-European populations and for monogenic BP genes; functional and experimental methods; and construction of a GRS for analysis with BP traits and cardiovascular outcomes. All P values are from two-sided tests.

Analysis models. For the GWAS, we performed linear regression analyses of the three (untransformed) continuous, medication-adjusted BP traits (SBP, DBP and PP) for all measured and imputed genetic variants in dosage format using SNPTEST software[65] under an additive genetic model. We carried out a similar analysis for the exome content. Quantile–quantile plots are shown in **Supplementary Figure 16**. Each analysis included the following covariates: sex, age, age^2, BMI, the top ten principal components and a binary indicator variable for UK Biobank versus UK BiLEVE to adjust for the different genotyping chips. We also ran an association analysis within UKB for validated BP-associated SNVs and hypertension using logistic regression under an additive model with adjustments as above. There were 76,554 hypertensive cases, and the 64,384 remaining participants were treated as non-hypertensive controls. This sample size is slightly larger than the N = 140,866 used in the main analyses, as participants with only one BP measurement but with reported BP-lowering medication could be included as hypertensive.

Previously reported variants. We compiled a list of all SNVs previously reported to be associated with BP at the time of analysis (**Supplementary Table 13**). This list includes all published SNVs that have been identified and validated from previous GWAS, CardioMetabochip and exome chip projects[7–11]. We augmented this list to include all 34,459 SNVs in LD with these previously reported SNVs, according to a threshold of $r^2 \geq 0.2$. Results for all these variants were extracted for each of the three BP traits, to check previously reported BP associations in the UKB data, according to whether the sentinel SNV or a variant at the locus in LD ($r^2 \geq 0.2$) with it showed evidence of support ($P < 0.01$) for association with at least one of the three BP traits.

Replication strategy. We used three independent external data sets for replication (**Supplementary Note**). First, for the GWAS analysis based on advanced 1000 Genomes imputation enhanced by UK10K data, we considered SNVs with MAF ≥1% and performed a reciprocal replication exchange with the International Consortium of Blood Pressure (ICBP) 1000G meta-analysis (max N = 150,134). The imputation strategy for ICBP 1000G meta-analysis was based on an earlier imputation grid for the 1000 Genomes Project. In addition, we recruited further cohorts with 1000 Genomes data that had not contributed to the ICBP-1000G discovery meta-analysis: ASCOT-UK (N = 3,803), ASCOT-SC (N = 2,462), BRIGHT (N = 1,791), Generation Scotland (GS) (N = 9,749), EGCUT (N = 5,468), Lifelines (N = 13,292) and PREVEND (N = 3,619). This gave a total of N = 190,318 independent replication samples for the GWAS analysis.

Second, because the UK Biobank and UK BiLEVE genotyping chips contain exome content, we sought replication from two BP exome consortia (European exome consortium and the Cohorts for Heart and Aging Research in Genome Epidemiology (CHARGE) BP exome consortium), to allow validation of coding variants and variants with lower frequency. The European exome consortium (N = 161,926) and CHARGE consortium (N = 119,792) gave a total of N = 281,718 independent replication samples for the exome analysis.

Note that the lookups for GWAS and exome discovery were distinct sets of SNVs. Loci were assigned sequentially, prioritizing the primary GWAS discovery first and then considering any remaining loci with non-overlapping exome content for replication in the independent exome replication resources.

Statistical criteria for replication. For the GWAS discovery, there were ~9.8 million SNVs with MAF ≥1% and INFO >0.1. We considered for follow-up any SNVs with $P < 1 \times 10^{-6}$ for any of the three BP traits. For the exome discovery, there were 149,026 exome SNVs (**Supplementary Note**) that were polymorphic with INFO >0.1; for follow-up, we considered all SNVs with MAF ≥0.01% and $P < 1 \times 10^{-5}$. All such SNVs were annotated to loci according to both an LD threshold of $r^2 \geq 0.2$ and a 1-Mb interval (**Supplementary Note**), and signals were classified either as belonging to unvalidated loci or being potential secondary signals at previously reported loci at the time of analysis.

Selection of variants for follow-up. The sentinel (most significant) SNV from each association signal was selected for follow-up, all of which were pairwise independent by LD ($r^2 < 0.2$). For the GWAS discovery, we checked that potential lookup SNVs were covered in the ICBP-1000G replication data (**Supplementary Tables 28** and **29**, and **Supplementary Note**). Of the 235 novel loci containing previously unreported SNVs at the time of analysis with MAF ≥1%, INFO >0.1 and $P < 1 \times 10^{-6}$, 218 were covered, and similarly 100 of the 123 potential secondary SNVs at 51 of the 54 previously reported BP loci at the time of analysis were available for follow-up. For the exome discovery, by following up SNVs with MAF ≥0.01%, INFO >0.1 and $P < 1 \times 10^{-5}$ across the three BP traits, we carried forward for replication sentinel SNVs at 22 unvalidated loci and potential secondary SNVs at 3 previously reported loci at the time of analysis. We produced LocusZoom plots for each of the lookup variants.

Replication meta-analyses. The replication and combined meta-analyses were performed within METAL software[66] using fixed-effects inverse-variance-weighted meta-analysis (**Supplementary Note**). The combined meta-analysis of both the UKB discovery (N = 140,886) and GWAS replication meta-analysis (max N = 190,070) included a total maximum sample size of N = 330,956. For the exome combined meta-analysis, we synthesized data from the UKB discovery exome content (max N = 140,866) with the replication data set from both exome consortia (total max N = 281,718), giving a maximum sample size of N = 422,604.

Validation criteria. In our study, a signal was declared to be validated if it satisfied all three of the following criteria: (i) the sentinel SNV was genome-wide significant ($P < 5 \times 10^{-8}$) in the combined meta-analysis for any of the three BP traits; (ii) the sentinel SNV showed evidence of support ($P < 0.01$) in the replication meta-analysis alone for association with the most significantly associated BP trait from the combined meta-analysis (N.B.: $P < 0.01$

is more stringent than a range of thresholds calculated according to FDR; **Supplementary Note**); and (iii) the sentinel SNV had concordant directions of effect between the UKB discovery and the replication meta-analysis for the most significantly associated BP trait from the combined meta-analysis.

Secondary signals. By conditional analysis within UKB data, we assessed all validated secondary signals from our validated and previously reported loci at the time of analysis for independence from the sentinel or previously reported SNV, respectively (**Supplementary Note**). We declared a secondary signal to be independent of the previously reported SNV if there was less than a 1.5-fold difference between the main association and conditional association P values on a $-\log_{10}$ scale, i.e., if $-\log_{10}(P)/-\log_{10}(P_{cond}) < 1.5$. Note that the lookup criteria already ensured that the secondary variant was not in LD ($r^2 < 0.2$) with the previously reported SNV. If more than one SNV in a region was found to be independent, we undertook further rounds of iterative conditional analysis.

Lookups in non-European ancestries. As a secondary analysis, we looked up 102 and 5 validated SNVs from the GWAS and exome analyses, respectively, in non-European ancestry samples. These comprised analysis of East Asian (N = 31,513) and South Asian (N = 33,115) ancestry data from the iGEN-BP consortium[11] for the GWAS lookups and South Asian (N = 25,937), African-American (N = 21,488) and Hispanic (N = 4,581) ancestry data from the CHARGE BP exome consortium[10] and CHD+ Exome consortium[9] for the exome content lookups (**Supplementary Note**). We carried out a binomial (sign) test based on the number of SNVs with consistent directions of effect between UKB and each of the samples of non-European ancestry.

Monogenic blood pressure gene lookups. The UKB arrays include some rare coding variants for monogenic disorders. We collated a list of all specific mutations within genes known to be associated with monogenic BP disorders[19]. Results from the UKB association analyses for all three BP traits were extracted for any of these SNVs directly covered in the UKB data set (**Supplementary Table 14**). Note that a search of proxies did not augment the list of available variants, so results are reported for the specific variants only.

Functional analyses. To prioritize associated SNVs, we used an integrative bioinformatics approach to collate functional annotation (**Supplementary Table 30**) at both the variant and gene level for each SNV within the BP loci (all SNVs in LD $r^2 \geq 0.8$ with the BP-associated SNVs). At the variant level, we used ANNOVAR[67] to obtain comprehensive functional characterization of variants, including gene location, conservation and amino acid substitution impact based on a range of prediction tools including SIFT and PolyPhen-2. All nonsynonymous variants were predicted to be damaging by two or more methods.

We used the UCSC Genome Browser to review the sequence-specific context of SNVs in relation to function, particularly in the Encyclopedia of DNA Elements (ENCODE) data set[68]. We used the UCSC table browser to annotate SNVs in ENCODE regulatory regions. We evaluated SNVs for impact on putative miRNA target sites in the 3′ UTRs of transcripts by a query of the miRNASNP database[69]. We evaluated all SNVs in LD ($r^2 \geq 0.8$) with our validated sentinel SNVs for evidence of mediation of eQTLs in all 44 tissues using the GTEx database, to identify validated loci that are highly expressed and to highlight specific tissue types that show eQTLs for a large proportion of validated loci. We further sought to identify validated loci with the strongest evidence of eQTL associations in arterial tissue in particular.

At the gene level, we used Ingenuity Pathway Analysis (IPA) software (IPA, Qiagen) to review genes with prior links to BP, on the basis of annotation with the "Blood Pressure" Medline Subject Heading (MESH) term, which is annotated to 684 genes. We also used IPA to identify genes that interact with BP MESH-annotated genes and evaluated genes for evidence of small molecule druggability on the basis of queries of the Chembl and Drug Gene Interaction database.

We then performed overall enrichment testing across all loci. First, we used DEPICT[70] (Data-driven Expression-Prioritized Integration for Complex Traits) to identify tissues and cells in which the BP loci were highly expressed. DEPICT uses a large number of microarrays (~37,000) to identify cells and tissues where the genes are highly expressed and uses precomputed GWAS phenotypes to adjust for co-founding sources. DEPICT provides a P value of enrichment and an FDR-adjusted P value for each tissue and cell type tested.

Furthermore, to investigate regulatory regions, we employed a two-tiered approach to investigate cell-type-specific enrichment within DNase I–hypersensitive sites using FORGE, which tests for enrichment of SNVs within DNase I–hypersensitive sites in 123 cell types from the Epigenomics Roadmap Project and ENCODE[71] (**Supplementary Note**). Validated sentinel SNVs from our study were analyzed along with previously reported SNVs at the time of analysis and secondary signals (with $P < 1 \times 10^{-4}$) to evaluate the overall tissue-specific enrichment of BP-associated variants. In a second analysis, we used FORGE (with no LD filter) to investigate directly our curated candidate regulatory SNVs for overlap with cell-type-specific DNase I–hypersensitive signals.

GenomeRunner[72] was used to search for enrichment of validated and previously reported sentinel SNVs with histone modification mark genomic features (**Supplementary Note**). Relevant cardiovascular tissue expression was investigated using Fantom5 reference transcript expression data (http://fantom.gsc.riken.jp/5) (**Supplementary Note**).

We used IPA to identify biological pathways and transcriptional upstream regulators enriched for genes within the BP loci. The transcriptional upstream regulator analysis aimed to identify transcription factors, compounds, drugs, kinases and other molecules for which the target is one of the BP-associated genes under investigation.

We queried SNVs against PhenoScanner[16] to investigate trait pleiotropy, extracting all association results with nominal significance at $P < 0.05$ for full reporting (**Supplementary Table 16**), and then extracted genome-wide significant results to highlight the validated loci with the strongest evidence of association with other traits (**Supplementary Fig. 5a**). We also used the Genomic Regions Enrichment of Annotations Tool (GREAT) to study gene set enrichment of mouse phenotype and disease ontology terms within our validated and previously reported loci at the time of analysis, using default SNV-to-gene mapping settings[73].

We carried out metabolomics analysis using two sets of data. First, we used ^{1}H NMR lipidomics data on plasma from a subset of 2,000 participants of the Airwave Health Monitoring Study[74,75] (**Supplementary Note**). For each validated BP-associated SNV, we ran association tests with the lipidomics data using linear regression analyses, adjusted for age and sex. We computed significance thresholds using a permutation-derived family-wise error rate (5%) to account for the high correlation structure of these data (ENT = 35) (ref. 76). We also tested each validated SNV against published genome-wide versus metabolome-wide associations in plasma and urine using publicly available data from the Metabolomics GWAS Server to identify metabolites that have been associated with variants of interest at $P < 3.0 \times 10^{-4}$ (Bonferroni-corrected P value for validated signals)[22,23].

Experimental methods. We prioritized genes for laboratory testing on the basis of evidence of SNV function (including coding variants, eQTLs and Hi-C interactions), biological support for relevance to BP (from literature review) and mouse transgenic phenotype. We performed genotyping and qRT–PCR for the selected sentinel variants of interest using human VSMCs and endothelial cells and tested for expression levels (**Supplementary Table 31** and **Supplementary Note**). All three SNVs were tested using an additive model.

Genetic risk scores. GRSs were constructed using data from the independent Airwave study[74] to assess the combined effect of the BP-associated variants on BP and risk of hypertension (**Supplementary Note**), while avoiding bias from 'winner's curse'. We created weighted GRSs for all pairwise-independent, LD-filtered ($r^2 < 0.2$) previously reported variants at the time of analysis and our validated variants (sentinel and secondary SNVs) combined, using available SNVs (**Supplementary Table 22**). For the previously reported variants at the time of analysis, we weighted BP-increasing alleles by the β coefficients from the UKB analysis. For our validated variants, the β coefficients of the replication meta-analysis were used as independent, unbiased weights.

For analyses of the variance explained within the independent Airwave cohort, we used three trait-specific GRSs (SBP, DBP and PP). Each GRS

included all variants, but weights were trait specific, using the β coefficients from the analysis of each of the three different BP traits; for example, the SBP GRS was weighted by the β coefficients from the SBP GWAS. To calculate the percent of variance for each BP trait explained by its corresponding trait-specific GRS, not accounted for by known factors, we generated the residuals from the regression model of each trait against covariates of age, age^2, sex and BMI. We then fit a second linear model for the trait residuals with all the variants in the GRS plus the top ten principal components.

For risk score analyses, we calculated a single BP GRS, as the average of the SBP and DBP GRSs. We standardized the average GRS to have mean of 0 and s.d. of 1. We assessed the association of the continuous average GRS variable with each BP trait by simple linear regression. We also ran logistic regression to examine the association of the average GRS with risk of hypertension. We performed each analysis both with and without adjustment for sex. We tested for interaction between age (<50 years and ≥50 years) and the effect of the GRS on BP. We then compared BP levels and risk of hypertension for individuals in the top and bottom 20% of the GRS distribution at ≥50 years of age using linear and logistic regression, respectively.

We also assessed the association of the average BP GRS with cardiovascular outcomes in the UKB data. We included all pairwise-independent previously reported BP variants at the time of analysis and our validated variants. We used logistic regression with binary outcome variables for coronary heart disease, stroke and cardiovascular disease (**Supplementary Note**) and the GRS as an explanatory variable (with and without sex adjustment).

Data availability. The data generated during the current study are available from the UK Biobank data repository (http://biota.osc.ox.ac.uk/), which can be accessed by researchers upon application. These data include the derived GWAS analysis summary data from our UK Biobank discovery data for all three BP traits. The genetic and phenotypic UK Biobank data are also available upon application to the UK Biobank (https://www.ukbiobank.ac.uk/). All replication data generated during this study are included in the published article. For example, association results of lookup variants from our replication analyses and the subsequent combined meta-analyses are contained within the supplementary tables provided.

63. Wain, L.V. *et al.* Novel insights into the genetics of smoking behaviour, lung function, and chronic obstructive pulmonary disease (UK BiLEVE): a genetic association study in UK Biobank. *Lancet Respir. Med.* **3**, 769–781 (2015).
64. Tobin, M.D., Sheehan, N.A., Scurrah, K.J. & Burton, P.R. Adjusting for treatment effects in studies of quantitative traits: antihypertensive therapy and systolic blood pressure. *Stat. Med.* **24**, 2911–2935 (2005).
65. Marchini, J., Howie, B., Myers, S., McVean, G. & Donnelly, P. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat. Genet.* **39**, 906–913 (2007).
66. Willer, C.J., Li, Y. & Abecasis, G.R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190–2191 (2010).
67. Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* **38**, e164 (2010).
68. Barnes, M.R. Exploring the landscape of the genome. *Methods Mol. Biol.* **628**, 21–38 (2010).
69. Gong, J. *et al.* Genome-wide identification of SNPs in microRNA genes and the SNP effects on microRNA target binding and biogenesis. *Hum. Mutat.* **33**, 254–263 (2012).
70. Pers, T.H. *et al.* Biological interpretation of genome-wide association studies using predicted gene functions. *Nat. Commun.* **6**, 5890 (2015).
71. Dunham, I., Kulesha, E., Iotchkova, V., Morganella, S. & Birney, E. FORGE: a tool to discover cell specific enrichments of GWAS associated SNPs in regulatory regions. *F1000Res.* **4**, 18 (2015).
72. Dozmorov, M.G., Cara, L.R., Giles, C.B. & Wren, J.D. GenomeRunner: automating genome exploration. *Bioinformatics* **28**, 419–420 (2012).
73. McLean, C.Y. *et al.* GREAT improves functional interpretation of *cis*-regulatory regions. *Nat. Biotechnol.* **28**, 495–501 (2010).
74. Elliott, P. *et al.* The Airwave Health Monitoring Study of police officers and staff in Great Britain: rationale, design and methods. *Environ. Res.* **134**, 280–285 (2014).
75. Petersen, M. *et al.* Quantification of lipoprotein subclasses by proton nuclear magnetic resonance–based partial least-squares regression models. *Clin. Chem.* **51**, 1457–1461 (2005).
76. Chadeau-Hyam, M. *et al.* Metabolic profiling and the metabolome-wide association study: significance level for biomarker identification. *J. Proteome Res.* **9**, 4620–4627 (2010).

Corrigendum: Genome-wide association analysis identifies novel blood pressure loci and offers biological insights into cardiovascular risk

Helen R Warren, Evangelos Evangelou, Claudia P Cabrera, He Gao, Meixia Ren, Borbala Mifsud, Ioanna Ntalla, Praveen Surendran, Chunyu Liu, James P Cook, Aldi T Kraja, Fotios Drenos, Marie Loh, Niek Verweij, Jonathan Marten, Ibrahim Karaman, Marcelo P Segura Lepe, Paul F O'Reilly, Joanne Knight, Harold Snieder, Norihiro Kato, Jiang He, E Shyong Tai, M Abdullah Said, David Porteous, Maris Alver, Neil Poulter, Martin Farrall, Ron T Gansevoort, Sandosh Padmanabhan, Reedik Mägi, Alice Stanton, John Connell, Stephan J L Bakker, Andres Metspalu, Denis C Shields, Simon Thom, Morris Brown, Peter Sever, Tõnu Esko, Caroline Hayward, Pim van der Harst, Danish Saleheen, Rajiv Chowdhury, John C Chambers, Daniel I Chasman, Aravinda Chakravarti, Christopher Newton-Cheh, Cecilia M Lindgren, Daniel Levy, Jaspal S Kooner, Bernard Keavney, Maciej Tomaszewski, Nilesh J Samani, Joanna M M Howson, Martin D Tobin, Patricia B Munroe, Georg B Ehret, Louise V Wain, The International Consortium of Blood Pressure (ICBP) 1000G Analyses, The CHD Exome+ Consortium, The ExomeBP Consortium, The T2D-GENES Consortium, The GoT2DGenes Consortium, The Cohorts for Heart and Ageing Research in Genome Epidemiology (CHARGE) BP Exome Consortium, The International Genomics of Blood Pressure (iGEN-BP) Consortium, Michael R Barnes, Ioanna Tzoulaki, Mark J Caulfield, Paul Elliott & for The UK Biobank CardioMetabolic Consortium BP working group
***Nat. Genet.*; doi:10.1038/ng.3768; corrected online 20 February 2017**

In the version of this article the name of Chiara Batini was misspelled as Chiara Battini in the list of collaborators affiliated with International Consortium of Blood Pressure (ICBP) 1000G Analyses.