By 2025, it is estimated that there will be 1.5 billion people worldwide with systemic arterial hypertension. An elevated blood pressure level is a key risk factor in 49% of coronary artery disease and 69% of stroke, and causes half of all cardiovascular mortality. The observed familial aggregation of blood pressure suggests that the heritability of blood pressure is between 30% and 50%. This led to the anticipation that the genetic architecture might be readily tractable, and caused by a small number of genes with appreciable effects.

The elucidation of genes implicated in Mendelian forms of hypertension demonstrates rare variants with substantial effects are responsible, and often these genes lie within pathways managing sodium homeostasis. More recently with advances in affordable high-throughput genotyping strategies, multiple common genetic variants with modest effects on blood pressure (<1 mm Hg systolic) have been discovered in the population. In aggregate, these common variants explain <3% of the variance of blood pressure. Although these findings may offer new mechanistic insights into the biology of blood pressure, a key question is can these findings translate into patient benefit? It is timely to reflect on recent advances in genomics, and the use of new resources, such as the 1000 Genomes Project and the Encyclopedia of DNA Elements, to annotate likely causal variants, and their relevance to cardiovascular disease. In this review, we discuss the advances in relation to our knowledge of the genetic architecture of blood pressure, and whether gene discoveries might influence cardiovascular risk assessment, help to stratify patient response to medicine, or identify new biological pathways for novel therapeutic targets. (Circ Res. 2013;112:1365-1379.)

Key Words: blood pressure ■ exome ■ genome-wide association studies ■ next generation sequencing ■ pharmacogenomics ■ 1000 genome

Advances in Blood Pressure Genomics

Patricia B. Munroe, Michael R. Barnes, Mark J. Caulfield

Abstract: The elucidation of genes implicated in Mendelian forms of hypertension demonstrates rare variants with substantial effects are responsible, and often these genes lie within pathways managing sodium homeostasis. More recently with advances in affordable high-throughput genotyping strategies, multiple common genetic variants with modest effects on blood pressure (<1 mm Hg systolic) have been discovered in the population. In aggregate, these common variants explain <3% of the variance of blood pressure. Although these findings may offer new mechanistic insights into the biology of blood pressure, a key question is can these findings translate into patient benefit? It is timely to reflect on recent advances in genomics, and the use of new resources, such as the 1000 Genomes Project and the Encyclopedia of DNA Elements, to annotate likely causal variants, and their relevance to cardiovascular disease. In this review, we discuss the advances in relation to our knowledge of the genetic architecture of blood pressure, and whether gene discoveries might influence cardiovascular risk assessment, help to stratify patient response to medicine, or identify new biological pathways for novel therapeutic targets. (Circ Res. 2013;112:1365-1379.)

Key Words: blood pressure ■ exome ■ genome-wide association studies ■ next generation sequencing ■ pharmacogenomics ■ 1000 genome

B y 2025, it is estimated that there will be 1.5 billion people worldwide with systemic arterial hypertension. An elevated blood pressure level is a key risk factor in 49% of coronary artery disease and 69% of stroke, and causes half of all cardiovascular mortality. The observed familial aggregation of blood pressure suggests that the heritability of blood pressure is between 30% and 50%. This led to the anticipation that the genetic architecture might be readily tractable, and caused by a small number of genes with appreciable effects.

The elucidation of genes implicated in Mendelian forms of hypertension demonstrated that we can identify rare variants with high penetrance, but with substantial effects on blood pressure. It is noteworthy that these mutations often lie within pathways managing sodium homeostasis. Alongside these early successes family-based genome-wide linkage studies were undertaken using sparse microsatellite marker sets, but we now know that these approaches were underpowered to detect gene loci linked to essential hypertension or blood pressure.
More recently, advances in high-throughput genome-wide chip-based strategies, alongside an improved understanding of the architecture of the human genome, have enabled direct typing and imputation of ≤2.5 million common single-nucleotide polymorphisms (SNPs) across the human genome. An early pathfinder SNP-based genome-wide association study (GWAS) was the Wellcome Trust Case Control Consortium, which studied 2000 hypertensive cases and 3000 common controls, as part of an experiment which included 6 other diseases.12 This analysis found no significant findings for hypertension, although several loci were discovered for type 1 and 2 diabetes mellitus, coronary artery disease, Crohn disease, and rheumatoid arthritis.

This experimental strategy was extended in a genome-wide scan of genomic structural variation on the basis of the copy number variants. This was a particularly challenging experiment because copy number variants can be difficult to genotype, but again no significant associations were found with hypertension.13 A criticism leveled at the Wellcome Trust Case Control Consortium genome-wide SNP and copy number variant analysis was the use of a common control group, which did not have blood pressure measurements. In a common control group derived from the population, there will be between 25% and 30% with hypertension potentially diluting their value as controls. However, it is now clear from further research that the early SNP-based genome-wide association analyses for hypertension were simply underpowered to detect the effect sizes that loci exert on blood pressure.

There were limited hypertension case–control resources available to do enlarged meta-analyses, which prompted researchers to study blood pressure as a quantitative trait. Significant association was reported with SNPs in STK39 (serine threonine kinase 39) and diastolic blood pressure,14 for a SNP upstream of CDH13 (adhesion glycoprotein T-cadherin) and systolic blood pressure,15 and for a SNP near ATP2B1 (ATPase calcium transporting, plasma membrane 1) with systolic and diastolic blood pressure.16 Only variants near ATP2B1 were replicated in subsequent studies with much larger sample sizes, suggesting that the associations near STK39 and CDH13 may be either false positives or are population-specific effects. The modest findings from individual GWAS for blood pressure traits led to the progression to meta-analyses of GWAS. The analyses of ≥20000 individuals led to the discovery of multiple common variants with modest effects.17–21 To date, these findings explain <3% of blood pressure variance. Although these findings might offer new mechanistic insights into blood pressure biology, a key question is how can these findings translate into patient benefit? It is timely to reflect on recent advances in the genomics of blood pressure, and whether these impact on assessment of cardiovascular risk, or might be used to stratify patient response to therapies, or identify new biological pathways for novel therapeutic targets (Figure 1).

**Advances in Mendelian Disorders Influencing Blood Pressure**

The discovery of loci influencing blood pressure began with Mendelian forms of hypertension and hypotension. These syndromes usually have early onset and possess distinctive phenotypic features (eg, hypo- or hyperkalemia).6,22 In Table 1, we summarized the main clinical features of many of these syndromes and indicate the causal genes and mechanisms. An interesting feature is that for many the mutated gene product affects salt/water homeostasis in the kidney, and several of the pathogenic mutations typically exhibit large effects on blood pressure and in some cases delineate likely therapeutic response.

Recent work on familial hyperkalemic hypertension (also known as Gordon syndrome and pseudohypoaldosteronism type Ia) illustrates that the so-called simple Mendelian traits may really be quite complex. In 2002, linkage analysis of Gordon kindreds identified mutations in the serine threonine kinases (WNK1 and WNK4) as causes of this familial hyperkalemic hypertensive phenotype.23 These kinases operate as molecular switches in a cascade, which regulates sodium reabsorption via the sodium chloride cotransporter (NCC) and, to some extent, the epithelial sodium channel in the collecting ducts. Mutations in WNK1 and WNK4 cause overactivity of the NCC, which explains the association with hyperkalemia, hyperchloremia, and thiazide responsiveness of blood pressure. This has led to the identification of a phosphorylation cascade with multiple points for potential therapeutic targeting for antihypertensives.

More recently, kindreds with familial hyperkalemic hypertension who did not have mutations at the WNK1 and WNK4 loci have been subjected to either SNP-based genome-wide linkage scans or whole-exome next generation sequencing (NGS), focused on coding regions or exons.24,25 This analysis identified the Kelch-like protein 3 (KLHL3) and Cullin as 2 further loci responsible for this hypertensive phenotype. Direct sequencing of affected individuals revealed rare missense mutations in KLHL3 associated with heterogeneous phenotypes, and different modes of inheritance (recessive and dominant), whereas at the Cullin locus, splice-site mutations inherited in a dominant manner were found. In an elegant series of functional experiments, the KLHL3 protein, which is an actin-binding protein, was shown to recruit substrates for Cullin 3–based ubiquitin ligase complexes. Furthermore, the KLHL3 protein is coexpressed along with the NCC and downregulates expression of this distal convoluted tubule ion channel at the cell surface. These studies reveal the power of NGS to uncover previously unknown pathways for sodium homeostasis in the distal nephron that affect blood pressure.
Translation of Findings From Mendelian Traits Into Population Blood Pressure

A natural extension of findings beyond Mendelian blood pressure phenotypes is to test whether common, or rare variants, at these gene loci affect blood pressure, or hypertension in the population. There are some data suggesting that rare haplotypes of WNK1 and common variants at the inwardly rectifying potassium channel ROMK may have modest effects on blood pressure.38–40 In the Framingham Heart Study, the effect of rare variants in the SLC12A3 (NCC), SLC12A1 (NKCC2), and the inwardly rectifying potassium channel, KCNJ1 (ROMK), genes were tested for to see if they influence blood pressure.41 These genes cause rare recessive disorders with substantial effects on blood pressure (≥6–10 mm Hg systolic). It is, therefore, of considerable interest to discover that individuals heterozygous for rare mutations in the same genes that alter renal salt handling are associated with significant reductions in blood pressure in the population. More recently, resequencing of 11 candidate genes (of which 10 are mutated in monogenic forms of hypertension) was performed in 560 unrelated individuals (European and African ancestry) with extreme hypertension.42 Sequencing identified 2535 variants, testing of these rare and common variants individually in relatively large independent sample sets revealed no evidence of association with systolic blood pressure. However, an analysis, including all the variants pooled, was significant. These results are encouraging and suggest that variations at these genes may also be important in the development of systemic hypertension. Follow-up work in larger sample sizes will be required to back up these initial observations.

Large-Scale Genome-Wide Studies Identify Multiple Blood Pressure Loci

Since 2009 a sequential series of large-scale meta-analyses of genome-wide scans for blood pressure in individuals of European ancestry have identified loci affecting systolic, diastolic, pulse, and mean arterial blood pressure.17,18,21 In 2011, the International Consortium for Blood Pressure genome-wide association studies (ICBP-GWAS) published the largest meta-analysis for systolic and diastolic blood pressure in >69,899 European individuals, followed by validation in 132,000 individuals.20 To reduce potential for confounding in this analysis, blood pressure was adjusted for age, age² (this accounts for the middle age plateau of diastolic blood pressure), body mass index, and antihypertensive medication. The approach taken to adjust for the presence of antihypertensive therapy added 15/10 mm Hg to the blood pressure of people on blood pressure-lowering medication. Blood pressure and SNP association were analyzed using inverse variance weighted meta-analysis with an additive model and a P value of <5×10⁻⁸ as genome-wide statistical significance. The impact of common variants on the trait is measured as the effect size expressed independently for systolic and diastolic blood pressure in mm Hg.

The ICBP-GWAS study identified and validated 29 SNPs at 28 loci for systolic and diastolic blood pressure. Among these loci there are some genes which encode proteins within pathways or systems which may have a clear physiological impact on blood pressure, such as genes from the nitric oxide-natriuretic peptide pathway. In >50% of cases, the lead SNP is intragenic but for some loci there was no gene within the region of association that encoded a protein with a biologically plausible effect on blood pressure. These findings also highlight locus pleiotropy across phenotypes exemplified by the association of the SH2B3 locus being common to blood pressure, celiac disease, myocardial infarction, and type 1 diabetes mellitus,20,33,34 and the SLC39A8 locus to blood pressure, body mass index, and high density lipoprotein.

Figure 1. Summary of study designs, analytic strategies, key findings, and their translational potential in the elucidation of the genetic architecture of blood pressure. CV indicates cardiovascular.
### Table 1. Monogenic Forms of Hypertension and Hypotension

<table>
<thead>
<tr>
<th>Condition</th>
<th>Mode of Inheritance</th>
<th>Phenotype</th>
<th>Chromosome(s)</th>
<th>Gene, Biological Effect, and Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hypertension</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pseudohypoaldosteronism type 2; familial hyperkalemic hypertension; Gordons syndrome</td>
<td>Autosomal dominant and autosomal recessive</td>
<td>Hypertension</td>
<td>2q36, 5q31, 12p13, 17q21</td>
<td>Protein kinase, lysine deficient 1 and 4 (WNK 1, WNK 4)&lt;sup&gt;23&lt;/sup&gt; Kelch-like protein 3 (KLHL3) and Cullin 3 (Cullin)&lt;sup&gt;24,25&lt;/sup&gt; Overactivity of sodium chloride cotransporter, via different regulatory mechanisms</td>
</tr>
<tr>
<td>Hypertension associated with PPARγ mutations</td>
<td>Autosomal dominant</td>
<td>Hypertension Insulin resistance, type 2 diabetes mellitus</td>
<td>3p25</td>
<td>Peroxisome proliferator-activated receptor gamma (PPAR-)γ&lt;sup&gt;92&lt;/sup&gt; Malfunctioning nuclear receptor PPARγ</td>
</tr>
<tr>
<td>Hypertension exacerbated by pregnancy</td>
<td>Autosomal dominant</td>
<td>Early hypertension, exacerbated by pregnancy</td>
<td>4q31.23</td>
<td>Mineralocorticoid receptor (NR3C2)&lt;sup&gt;20&lt;/sup&gt;</td>
</tr>
<tr>
<td>Glucocorticoid remediable aldosteronism</td>
<td>Autosomal dominant</td>
<td>Hypertension with other variable features</td>
<td>8q24.3</td>
<td>11β-hydroxylase/aldosterone synthase (CYP11B1/CYP11B2) chimera&lt;sup&gt;24&lt;/sup&gt;</td>
</tr>
<tr>
<td>Congenital adrenal hyperplasia with 11β-hydroxylase deficiency</td>
<td>Autosomal recessive</td>
<td>Hypertension with other variable features</td>
<td>8q</td>
<td>11β-hydroxylase (CYP11B1)↓&lt;sup&gt;109&lt;/sup&gt; ↓11β hydroxylase</td>
</tr>
<tr>
<td>Congenital adrenal hyperplasia with 17α-hydroxylase deficiency</td>
<td>Autosomal recessive</td>
<td>Hypertension absent sexual maturation, ↑K+</td>
<td>10q24.3</td>
<td>17α-hydroxylase (CYP17A1)↑&lt;sup&gt;110&lt;/sup&gt; ↓17α hydroxylase</td>
</tr>
<tr>
<td>Hypertension and brachydactyly</td>
<td>Autosomal dominant</td>
<td>Hypertension brachydactyly</td>
<td>12p</td>
<td>Unknown&lt;sup&gt;112&lt;/sup&gt;</td>
</tr>
<tr>
<td>Apparent mineralocorticoid excess</td>
<td>Autosomal recessive</td>
<td>Hypertension</td>
<td>16q</td>
<td>11β-hydroxysteroid dehydrogenase (HSD11B2).&lt;sup&gt;113&lt;/sup&gt; ↓11β-hydroxysteroid dehydrogenase</td>
</tr>
<tr>
<td>Liddle syndrome</td>
<td>Autosomal dominant</td>
<td>Hypertension</td>
<td>12p13.31, 16p12.2</td>
<td>Sodium channel nonvoltage-gated 1, β and γ subunits (SCCN1A, SCNN1B, SCNN1G),&lt;sup&gt;114&lt;/sup&gt; Abnormal ENaC: reduced receptor clearance</td>
</tr>
<tr>
<td>Hypotension</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bartter syndrome type 3</td>
<td>Autosomal recessive</td>
<td>Birth at term, nephrocalcinosis, ↑K+, increased renin and aldosterone</td>
<td>1p36.13</td>
<td>CLCNKB,&lt;sup&gt;105&lt;/sup&gt; kidney chloride channel B gene. Chloride channel current is compromised in renal tubule.</td>
</tr>
<tr>
<td>Bartter syndrome type 2</td>
<td>Autosomal recessive</td>
<td>Prematurity, nephrocalcinosis, postnatal transient ↑K+ hypotension</td>
<td>11q24.3</td>
<td>KCNJ1.&lt;sup&gt;115&lt;/sup&gt; The ATP-sensitive inwardly-rectifying potassium channel, ROMK. Dysregulation of potassium secretion and NaCl reabsorption in the kidney</td>
</tr>
<tr>
<td>Bartter syndrome type 1</td>
<td>Autosomal recessive</td>
<td>Prematurity, nephrocalcinosis, hypotension</td>
<td>15q21.1</td>
<td>SLC12A1&lt;sup&gt;117&lt;/sup&gt; sodium-potassium-chloride cotransporter-2 (NKCC2). Defect in chloride transport cross the medullary thick ascending limb</td>
</tr>
<tr>
<td>Gitelman syndrome</td>
<td>Autosomal recessive</td>
<td>Salt-losing tubulopathy, ↓Ca&lt;sup&gt;2+&lt;/sup&gt;, ↓Mg&lt;sup&gt;2+&lt;/sup&gt;, Hypotension</td>
<td>16q13</td>
<td>SLC12A3&lt;sup&gt;118&lt;/sup&gt; thiazide-sensitive sodium-chloride (Na-CI) cotransporter (NCC7). Loss of function of NCCT</td>
</tr>
</tbody>
</table>

This table summarizes the key features of monogenic forms of hypertension and hypotension, and causative genes and mechanism. The arrow(s) indicate increased or decreased of level. Plasma Ca<sup>2+</sup> indicates calcium; plasma K+, potassium; and plasma Mg<sup>2+</sup>, magnesium.
Extension of Novel Blood Pressure Loci to Other Ancestries

Alongside gene discoveries in individuals of European ancestry, there have been several reports describing genetic variants affecting blood pressure, in individuals of East Asian, South Asian, and African ancestry. Some studies have suggested that the effect sizes of blood pressure variants may differ between different ethnic groups, (eg, the FGF5 locus). The ICBP-GWAS study performed a comprehensive analysis of whether findings made in people of European ancestry can be extended to individuals of non-European origin. They tested 29 SNPs identified in Europeans for systolic and diastolic blood pressure in a meta-analysis of results from non-European populations. A genetic risk score was associated with blood pressure in all the ethnic groups tested. Significant associations with individual SNPs were also observed in populations of East or South Asian ancestry (or both) after correction for multiple testing. The lack of significant associations for all the blood pressure SNPs tested was most likely because of low statistical power as the sample sizes of the other ethnic groups were all much smaller than the European data set.

The Impact on Cardiovascular Risk of These Gene Variants

The impact of 29 blood pressure variants discovered by ICBP-GWAS on cardiovascular outcomes was tested using a genetic risk score approach. Although the effect sizes of the individual variants are modest the aggregate impact on blood pressure was of similar magnitude to that of a standard antihypertensive therapy. It is important to note that the effects of these variants operate across the entire population blood pressure distribution, and we know from observational data that an increase in systolic blood pressure of 2 mmHg increases stroke risk by 10% and the coronary risk by 7%. It is, therefore, reassuring that the aggregate impact of the 29 variants did affect the risk of stroke, coronary disease, left ventricular hypertrophy, and hypertension but not renal disease. The latter finding may seem surprising given the role of the kidney in hypertension. This might arise, in part, because of the kidney being both a cause of hypertension and a target organ which is damaged by high blood pressure. It is noteworthy that some loci are shared between chronic kidney disease and blood pressure, such as Uromodulin, which encodes the Tamm Horsfall protein, and is related to both phenotypes. Similarly, rare variants at PLCE1 are associated with focal glomerulosclerosis, a form of familial glomerulonephritis, and common variants are at the same locus are associated with blood pressure.

Enhancing Gene Discovery and Locus Fine Mapping for Blood Pressure

The success of GWAS and gene-centric genotyping arrays for discovering disease-causing loci prompted further bespoke genotyping arrays. These included CardioMetabochip (Illumina, San Diego, CA) and ImmunoChip (Illumina, San Diego, CA). Both of these bespoke arrays provide cost-effective approaches for further gene discovery, and have content for fine mapping causal genetic variant(s) and loci. The CardioMetabochip array content (contains probes for ≤200,000 SNPs) was created by several GWAS meta-analysis consortia working on metabolic, cardiovascular, and related traits. The ICBP-GWAS consortium contributed the blood pressure content for this array; this included a set of 8000 LD pruned SNPs; these were derived from the interim results of the ICBP-GWAS discovery analyses for systolic and diastolic blood pressure. For practical purposes, this corresponds to tagging of all SNPs that have \( P \leq 0.01 \), for either systolic or diastolic blood pressure in the ICBP-GWAS discovery analysis. The array also genotypes 21 regions surrounding lead blood pressure SNPs that achieved genome-wide significance in an interim analysis, including some replication data. The SNPs were derived from several sources, the International Haplotype Mapping Project and the August 2009 release of the 1000 Genomes Project, and cosmopolitan tagging for multiethnic application of the chip was used.

The CardioMetabochip array has been genotyped in thousands of individuals to date, across many different ethnic groups, and the results for some metabolic and cardiovascular traits have recently been published. Using genome-wide complex trait analysis-independent signals have been observed at some loci. Fine mapping of disease loci using CardioMetabochip individuals of mainly European descent has not been as successful as hoped. The resolution of loci is improved in only a handful of cases. However, genotyping of the CardioMetabochip array in individuals of African
### Table 2. Summary of Blood Pressure Loci

<table>
<thead>
<tr>
<th>SNP</th>
<th>Locus Nickname</th>
<th>Chr</th>
<th>BP Position (Build 37)</th>
<th>CEU RAF</th>
<th>Potential Blood Pressure Candidates</th>
<th>Novel Drug Network, ENCODE or rsSNP Candidates</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs880315</td>
<td>CASZ1</td>
<td>1</td>
<td>10 796 866</td>
<td>0.65§</td>
<td>CASZ1†</td>
<td>...</td>
<td>39</td>
</tr>
<tr>
<td>rs5068‡</td>
<td>NPPA/NPPB</td>
<td>1</td>
<td>11 905 974</td>
<td>0.06</td>
<td>MTHFR†, CLCN6†, NPPA*, NPPB*</td>
<td>PLOD1*, AGTRAP*</td>
<td>73</td>
</tr>
<tr>
<td>rs17030613</td>
<td>ST7L-CAPZA1</td>
<td>1</td>
<td>113 190 807</td>
<td>0.49§</td>
<td>SLC16A1</td>
<td>CAPZA1†, ST7L†, MOV10†</td>
<td>19</td>
</tr>
<tr>
<td>rs2004776‡</td>
<td>AGT</td>
<td>1</td>
<td>230 848 702</td>
<td>0.24</td>
<td>AGT†</td>
<td>...</td>
<td>109</td>
</tr>
<tr>
<td>rs16849225‡</td>
<td>FIGN-GRB14</td>
<td>2</td>
<td>164 906 820</td>
<td>0.61§</td>
<td>FIGN</td>
<td>...</td>
<td>19</td>
</tr>
<tr>
<td>rs13082711</td>
<td>SLC4A7</td>
<td>3</td>
<td>27 537 909</td>
<td>0.78</td>
<td>SLC4A7†</td>
<td>EOMES*</td>
<td>20</td>
</tr>
<tr>
<td>rs9815354</td>
<td>ULK4</td>
<td>3</td>
<td>41 912 651</td>
<td>0.17</td>
<td>ULK4†, CTNNB1*</td>
<td>...</td>
<td>18</td>
</tr>
<tr>
<td>rs319690</td>
<td>MAP4</td>
<td>3</td>
<td>47 927 484</td>
<td>0.51</td>
<td>MAP4†, SMARCC1</td>
<td>CDC25A*</td>
<td>19, 21</td>
</tr>
<tr>
<td>rs419076</td>
<td>MECOM</td>
<td>3</td>
<td>169 100 886</td>
<td>0.47</td>
<td>MECOM†</td>
<td>...</td>
<td>20</td>
</tr>
<tr>
<td>rs871606</td>
<td>CHIC2</td>
<td>4</td>
<td>54 799 245</td>
<td>0.85</td>
<td>PDEGFA*</td>
<td>...</td>
<td>21</td>
</tr>
<tr>
<td>rs16998073</td>
<td>FGF5</td>
<td>4</td>
<td>81 184 341</td>
<td>0.21</td>
<td>FGF5†</td>
<td>...</td>
<td>17</td>
</tr>
<tr>
<td>rs13107325</td>
<td>SLC39A8</td>
<td>4</td>
<td>103 188 709</td>
<td>0.05</td>
<td>SLC39A8†</td>
<td>...</td>
<td>20</td>
</tr>
<tr>
<td>rs6825911</td>
<td>ENPEP</td>
<td>4</td>
<td>111 381 638</td>
<td>0.51§</td>
<td>ENPEP†, PITX2</td>
<td>...</td>
<td>19</td>
</tr>
<tr>
<td>rs13139571</td>
<td>GUCY1A3-GUCY1B3</td>
<td>4</td>
<td>156 645 513</td>
<td>0.76</td>
<td>GUCY1A3†, GUCY1B3†</td>
<td>...</td>
<td>20</td>
</tr>
<tr>
<td>rs1799945</td>
<td>HFE</td>
<td>6</td>
<td>260 911 179</td>
<td>0.14</td>
<td>HFE†</td>
<td>SLC17A1*, HIST1H4C†, HIST1H2B†</td>
<td>20</td>
</tr>
<tr>
<td>rs805303</td>
<td>BAT2-BAT5</td>
<td>6</td>
<td>31 616 366</td>
<td>0.61</td>
<td>SLC2DH2*, HSPA1L*, HSPA1A*, HSPA1B*</td>
<td>CSNK2B*, C6ORF25*, NEU1*, EHMT2*</td>
<td>20</td>
</tr>
<tr>
<td>rs17477177</td>
<td>PIK3CG</td>
<td>7</td>
<td>106 411 858</td>
<td>0.72</td>
<td>...</td>
<td>...</td>
<td>21</td>
</tr>
<tr>
<td>rs3918226</td>
<td>NOS3</td>
<td>7</td>
<td>150 690 176</td>
<td>0.08</td>
<td>ABP1*, KCNH2*, NOS3†, ACCN3,</td>
<td>...</td>
<td>21</td>
</tr>
<tr>
<td>rs2898290</td>
<td>BLK-GATA4</td>
<td>8</td>
<td>11 433 909</td>
<td>0.53</td>
<td>PINX1*, GATA4*, MTRM9</td>
<td>...</td>
<td>110</td>
</tr>
<tr>
<td>rs2071518</td>
<td>NOV</td>
<td>8</td>
<td>120 435 812</td>
<td>0.17</td>
<td>NOV†, ENPP2</td>
<td>...</td>
<td>21</td>
</tr>
<tr>
<td>rs11014166‡</td>
<td>CACNB2</td>
<td>10</td>
<td>18 708 798</td>
<td>0.66</td>
<td>CACNB2†</td>
<td>...</td>
<td>18</td>
</tr>
<tr>
<td>rs1530440</td>
<td>c10orf107†</td>
<td>10</td>
<td>63 524 591</td>
<td>0.19</td>
<td>C10orf107†</td>
<td>...</td>
<td>17</td>
</tr>
<tr>
<td>rs932764</td>
<td>PLCE1</td>
<td>10</td>
<td>95 895 940</td>
<td>0.44</td>
<td>PLCE1†</td>
<td>...</td>
<td>20</td>
</tr>
<tr>
<td>rs11191548‡</td>
<td>CYP17A1-NT5C2</td>
<td>10</td>
<td>104 846 178</td>
<td>0.91</td>
<td>CYP17A1†, NT5C2†</td>
<td>...</td>
<td>17</td>
</tr>
<tr>
<td>rs1801253‡</td>
<td>ADRB1</td>
<td>10</td>
<td>115 805 056</td>
<td>0.63</td>
<td>ADRB1†</td>
<td>...</td>
<td>109</td>
</tr>
<tr>
<td>rs661346</td>
<td>LSP1-TNNT3</td>
<td>11</td>
<td>190 292†</td>
<td>0.43</td>
<td>CTSD*</td>
<td>LSP1†</td>
<td>47</td>
</tr>
<tr>
<td>rs7129220</td>
<td>ADM</td>
<td>11</td>
<td>10 350 538</td>
<td>0.89</td>
<td>ADM†</td>
<td>...</td>
<td>20</td>
</tr>
<tr>
<td>rs2014408</td>
<td>SOX6</td>
<td>11</td>
<td>18 365 282</td>
<td>0.21</td>
<td>SOX6†</td>
<td>...</td>
<td>47</td>
</tr>
<tr>
<td>rs381815‡</td>
<td>PLEKH7</td>
<td>11</td>
<td>16 902 268</td>
<td>0.30</td>
<td>PLEKH7†</td>
<td>...</td>
<td>18</td>
</tr>
<tr>
<td>rs633185‡</td>
<td>FLJ32810-TMEM133</td>
<td>11</td>
<td>11 538 410</td>
<td>0.28</td>
<td>PGR*, TRPC6*</td>
<td>ARHGAP42†</td>
<td>20</td>
</tr>
<tr>
<td>rs1122084‡</td>
<td>ADAMTS8</td>
<td>11</td>
<td>130 273 230</td>
<td>0.38</td>
<td>ADAMTS8†</td>
<td>...</td>
<td>21</td>
</tr>
<tr>
<td>rs7249754‡</td>
<td>ATP2B1</td>
<td>12</td>
<td>90 060 586</td>
<td>0.37</td>
<td>ATP2B1†</td>
<td>DUSP5†</td>
<td>16</td>
</tr>
<tr>
<td>rs653178‡</td>
<td>SH2B3</td>
<td>12</td>
<td>112 007 756</td>
<td>0.53</td>
<td>ALDH2*, PTPN11*, SH2B3†</td>
<td>ACAD10*, RPL6*, CUX2†</td>
<td>17</td>
</tr>
<tr>
<td>rs2384550‡</td>
<td>TBX3-TBX5</td>
<td>12</td>
<td>115 352 731</td>
<td>0.35</td>
<td>TBX3*, TBX5</td>
<td>...</td>
<td>18</td>
</tr>
<tr>
<td>rs1378942‡</td>
<td>CYP1A1-CSK</td>
<td>15</td>
<td>75 077 367</td>
<td>0.36</td>
<td>CYP1A1*, CYP1A2†, CYP1A1*</td>
<td>LMAN1L*, SCAMP2*, hsa-mir-4513†</td>
<td>17</td>
</tr>
<tr>
<td>rs2521501</td>
<td>FES-FURIN</td>
<td>15</td>
<td>91 437 388</td>
<td>0.31</td>
<td>FES†, FURIN†</td>
<td>PRC1†</td>
<td>20</td>
</tr>
<tr>
<td>rs13333226</td>
<td>UMOD</td>
<td>16</td>
<td>20 365 654</td>
<td>0.81</td>
<td>UMOD†</td>
<td>ACVR2A†, ORC4†</td>
<td>44</td>
</tr>
<tr>
<td>rs17608766</td>
<td>GOSR2</td>
<td>17</td>
<td>45 013 271</td>
<td>0.86</td>
<td>GOSR2†</td>
<td>WNT3*, NSP*, RPRML†</td>
<td>20</td>
</tr>
<tr>
<td>rs16948048</td>
<td>ZNF652</td>
<td>17</td>
<td>47 440 466</td>
<td>0.39</td>
<td>ZNF652†, NGFR</td>
<td>PHB*, IGF2BP1†</td>
<td>17</td>
</tr>
</tbody>
</table>

(Continued)
The Technical Challenge of GWAS Data Interpretation

Although GWAS have revolutionized the study of complex disease genetics, accurate interpretation of the results still represents a considerable technical challenge. This approach exploits genome-wide correlation between markers, known as LD, to detect association. Each association highlights a set of marker variants in LD, which may often extend across many thousands of base pairs of the genome, spanning many genes. Determination of the causative variant among possibly hundreds of variants is very challenging. The magnitude of association alone is not a good guide, where all markers have not been genotyped, particularly where the causal variant and genotyped variants differ greatly in allele frequency. Using in silico methods, such as genotype imputation, can assist in the evaluation of association at each locus, but ultimately all variants linked to an association need to be evaluated for functional impact. Functional impact may be subtle and effects can be very diverse, from direct impact on coding sequences (e.g., nonsynonymous variation, to impact on regulatory regions, often influencing remotely located genes). The Encyclopedia of DNA Elements (ENCODE) project enables such analysis (see below), but this only serves to allow prioritization for laboratory on the basis of the follow-up. It may be too difficult to prove causality unequivocally (although functional studies in cell lines and model organisms can help). Genetic studies of blood pressure have shown that many associations exert a very modest influence on the phenotype. However, this does not mean that genes with modest effects are of limited value as targets or biological mechanisms of disease. Indeed, the Glitazone class of diabetes mellitus drugs targets the gene product of PPARG, a gene with an odds ratio of ≈1.2 for association with type 2 diabetes mellitus.62

The Usage of the 1000 Genomes Project for Cardiovascular Genetics

The goal of the 1000 Genomes Project is to create an integrated gene variant map. From the early phase of the project, we have access to information ≈95% of common variation arising because of SNPs and genomic structural variants, but until now much less was known about rarer variants.53 The project has now released their latest inventory of gene variants, which offers a tremendous resource for both rare and common disease cardiovascular research. This was achieved by integrating data from NGS of 1092 individuals from 14 populations, using a range of depths of whole-genome (6×) and exome-sequencing at 50× or 100×.63 By combining these approaches it has been possible to generate a validated haplotype map of 38 million SNPs, 1.4 million short insertions and deletions (indels), and >14,000 larger deletions. These capture ≥98% of common and uncommon alleles and confirm that individuals of different ancestries carry different profiles of rare and common variants. Furthermore, the 1000 Genomes experiment confirms that rare variants, with a frequency of <1% in the population, show substantial geographic differentiation and that the abundance of these rarer variants varies across biological pathways.63,64 These findings confirm that low-frequency variants are enriched for potentially functional mutations with 85% of nonsynonymous variants and >90% of stop-gain and splice-site disrupting variants below frequencies of 0.5%, compared with 65% of synonymous variants (no amino acid change).

Data from the 1000 Genomes Project are already being widely used to screen variants discovered in exome data from individuals with cardiovascular disorders. The augmented data set that has recently been published enables enhanced imputation strategies which infer missing genotypes, by association with known genotype. This improves the power to

<table>
<thead>
<tr>
<th>SNP</th>
<th>Locus Nickname</th>
<th>Chr</th>
<th>BP Position (Build 37)</th>
<th>CEU RAF</th>
<th>Potential Blood Pressure Candidates</th>
<th>Novel Drug Network, ENCODE or nsSNP Candidates</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs1327235</td>
<td>JAG1</td>
<td>20</td>
<td>10969030</td>
<td>0.46</td>
<td>JAG1</td>
<td></td>
<td>20</td>
</tr>
<tr>
<td>rs6015450</td>
<td>GNAS-EDN3</td>
<td>20</td>
<td>5775117</td>
<td>0.12</td>
<td>GNAS*, ZNF831†, EDN3*</td>
<td></td>
<td>20</td>
</tr>
</tbody>
</table>

This table contains a subset of SNPs associated with blood pressure traits, listed in chromosomal order. bp indicates base pair; CEU, Utah residents with ancestry from northern and western Europe; Chr, chromosome; nsSNP, if index SNP at a locus or a SNP in high linkage disequilibrium (P<0.8) is nonsynonymous SNP; RAF, risk allele frequency; and SNP, single-nucleotide polymorphism.

*Risk allele frequency in E Asians. The locus nickname indicates the genes nearest to the associated genetic variant, as described in the original papers. Previously suggested hypertension candidate genes are indicated on the basis of the literature reports.

†Genes which interact with the antihypertensive drug network.

‡Genes with putative functional LD proxy variants supported by ENCODE or Annovar nsSNP analysis.

§The identity of the first SNP associated at a particular locus.

| bp indicates base pair; CEU, Utah residents with ancestry from northern and western Europe; Chr, chromosome; nsSNP, if index SNP at a locus or a SNP in high linkage disequilibrium (P<0.8) is nonsynonymous SNP; RAF, risk allele frequency; and SNP, single-nucleotide polymorphism. |

| rs6015450  | JAG1           | 20  | 10969030               | 0.46    | JAG1                                |                                             | 20        |
| rs6015450  | GNAS-EDN3      | 20  | 5775117                | 0.12    | GNAS*, ZNF831†, EDN3*               |                                             | 20        |

The Usage of the 1000 Genomes Project for Cardiovascular Genetics

The goal of the 1000 Genomes Project is to create an integrated gene variant map. From the early phase of the project, we have access to information ≈95% of common variation arising because of SNPs and genomic structural variants, but until now much less was known about rarer variants.53 The project has now released their latest inventory of gene variants, which offers a tremendous resource for both rare and common disease cardiovascular research. This was achieved by integrating data from NGS of 1092 individuals from 14 populations, using a range of depths of whole-genome (6×) and exome-sequencing at 50× or 100×.63 By combining these approaches it has been possible to generate a validated haplotype map of 38 million SNPs, 1.4 million short insertions and deletions (indels), and >14,000 larger deletions. These capture ≥98% of common and uncommon alleles and confirm that individuals of different ancestries carry different profiles of rare and common variants. Furthermore, the 1000 Genomes experiment confirms that rare variants, with a frequency of <1% in the population, show substantial geographic differentiation and that the abundance of these rarer variants varies across biological pathways.63,64 These findings confirm that low-frequency variants are enriched for potentially functional mutations with 85% of nonsynonymous variants and >90% of stop-gain and splice-site disrupting variants below frequencies of 0.5%, compared with 65% of synonymous variants (no amino acid change).

Data from the 1000 Genomes Project are already being widely used to screen variants discovered in exome data from individuals with cardiovascular disorders. The augmented data set that has recently been published enables enhanced imputation strategies which infer missing genotypes, by association with known genotype. This improves the power to
detect de novo gene associations and will potentially enable us to undertake more granular fine mapping of existing association signals in CV traits. For common variants, the accuracy of genotype imputation at sites not featured on the original SNP array is between 90% and 95% in non-African and perhaps 90% in people of African ancestry.

As indicated earlier there are limited examples of successful fine mapping with identification of a single causal variant from signals derived from genome-wide association studies, probably because of extensive haplotype structure within regions of association. The latest data from the 1000 Genomes Project show that in Europeans, a single GWAS signal is typically in LD ($r^2 > 0.5$) with an average of 56 other variants. As a result of different ancestral genetic architectures and, therefore, different patterns of LD it is possible that transethnic fine mapping experiments, which have not been used to date, could be especially valuable in locating the causative variant. The upgraded 1000 Genomes data set increases the number of variants in LD with each GWAS signal by $\approx 25\%$, compared with the earlier phase of the project, and by more than double those in LD from the original Haplotype Mapping Project data set. Although the increase in available variation data from genome resequencing projects may make the task of fine mapping causal alleles appear more onerous because of the increased number of variants identified for functional follow-up studies that are required to demonstrate causality. Although there are many reasons for optimism moving forward with the application of methods, such as genotype imputation, it seems reasonable to expect that causal variants will be present among the known variants for common diseases. One could be forgiven for believing that it will simply be a matter of pinpointing the most likely variants and demonstrating their function; however, proving causality is much more complex than it might first appear. However, prospects for functional prioritization using in silico data continue to improve, the impact of variation in genes is already well served by a large range of tools, but until recently intergenic and intronic associations have proved much more difficult to characterize. This situation has changed with the recent publication of the regulatory annotation from the ENCODE project, described below. Now with most of the tools and the data there may be a real opportunity to address the bottleneck that currently exists between replicated associations and causal alleles, be prudent application of high-throughput methods to enable the functional characterization of thousands of variants that have been associated in GWAS studies to date.

**Rare Coding Region Variants and Blood Pressure**

The 1000 Genomes Project and data from NGS efforts have enabled the creation of an exome chip containing $\approx 250\,000$ rare and common variants (Illumina, San Diego, CA and Affymetrix, Santa Clara, CA). This chip offers much greater depth of coverage for rare variants, especially for putative functional exonic variants (eg, nonsynonymous variants, which have to be observed three times in 2 data sets) and nonsense and splice-site variants; [http://genome.sph.umich.edu/wiki/Exome_Chip_Design](http://genome.sph.umich.edu/wiki/Exome_Chip_Design). Genetic variants were selected from exome- and whole-genome sequencing of $>12\,000$ individuals, of different ethnicities (European, African, Chinese, and Hispanics), and disease areas (type 2 diabetes mellitus, cancer, metabolic, and psychiatric disorders). It is anticipated that genotyping of the exome chip will provide a cost-effective way to test the role of rare variants in complex disease, and that results may yield data on the missing heritability observed for many complex traits. From our experience and that of others with genotyping of this chip, there is a high specificity and sensitivity for detecting the minor allele of rare variants. Genotyping of the exome chip is on-going in large sample sizes, for many diseases, including blood pressure. Both single variant and burden testing (assessment of the cumulative effects of multiple variants in a genomic region) are planned. It is hoped that similar to early GWAS studies, there may be some low hanging fruit from the early efforts, but it is generally anticipated that hundreds of thousands of samples will be likely including samples for validation.

**ENCODE: Unraveling the Function of the Genome**

One of the Grand Challenges for genomics is to move findings beyond genes to a broader description of the genetic architecture of common complex diseases. The ENCODE project has recently presented findings of relevance to many cardiovascular researchers, whether they are interested in cell biology, physiology, or genomics. To functionally annotate the genome, the ENCODE team carried out 24 different genome-wide analyses, in >150 different cell lines. In >1600 experiments, they identified DNA methylation sites for epigenetic regulation of gene expression. They also detected regions of DNA with open chromatin accessible by regulatory proteins, defined positions for RNA binding and areas transcribed into RNA, sites of histone modification, and located transcription factor-binding sites. When the human genome project first reported, it seemed inconceivable to many scientists that the extended regions of DNA between the $\approx 20\,000$ human genes, sometimes referred to as gene deserts simply consisted of junk DNA. The real success of the ENCODE project has been to demonstrate the fallacy of junk DNA, instead showing that 80% of DNA shows evidence of function in $\leq 1$ of the cell lines/types studied in the project. Overall, they identified >70000 promoter regions, usually just upstream of genes, and almost 40000 enhancer regions regulating expression of distant genes.

Viewed from the perspective of GWAS, the ramifications of their findings are made clear by the observation that 99% of the genome lies within 17.4 kb of $\leq 1$ of the biochemical events measured by ENCODE. This illustrates a potential weakness in any tendency toward gene-centric interpretation of GWAS data and prompts consideration of the functional potential of almost any region of the genome. As a further bonus for researchers, ENCODE data offer a unique context to any genomic region, by allowing cell- or tissue-specific analysis, by comparison of the activity across 150 cell types. Cardiovascular researchers are well served by data from 14 different vascular cell types, including different types of vascular endothelial cells, fibroblasts, and smooth muscle (Table 3; [http://genome.ucsc.edu/ENCODE/dataMatrix/encodeDataMatrixHuman.html](http://genome.ucsc.edu/ENCODE/dataMatrix/encodeDataMatrixHuman.html)).
Table 3. Cardiovascular Relevant Cell Lines and Cell Types Investigated by the ENCODE Project

<table>
<thead>
<tr>
<th>Human Cell Type</th>
<th>Abbreviation</th>
<th>Tier</th>
<th>Data Sets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human umbilical vein endothelial cells</td>
<td>HUVEC</td>
<td>2</td>
<td>61</td>
</tr>
<tr>
<td>Aortic adventitial fibroblasts</td>
<td>HAaAF</td>
<td>3</td>
<td>9</td>
</tr>
<tr>
<td>Aortic smooth muscle cells</td>
<td>HAoSMC</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>Aortic endothelial cells</td>
<td>HAoEC</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Brain microvascular endothelial cells</td>
<td>HBMVEC</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>Brain vascular pericytes</td>
<td>HBVP</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Brain vascular smooth muscle</td>
<td>HMVSMC</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Cardiac fibroblasts</td>
<td>HCF, HCFia</td>
<td>3</td>
<td>14</td>
</tr>
<tr>
<td>Cardiac myocytes</td>
<td>HCM</td>
<td>3</td>
<td>9</td>
</tr>
<tr>
<td>Primary heart tissue (5 donors)</td>
<td>Heart_OC</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Dermal microvascular endothelial cells</td>
<td>HMVEC</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Pulmonary artery endothelial cells</td>
<td>HPAEC</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Pulmonary artery fibroblasts</td>
<td>HPAF</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>Saphenous vein endothelial cells</td>
<td>HSaVEC</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Adult skeletal muscle myoblasts</td>
<td>HSMMM</td>
<td>3</td>
<td>25</td>
</tr>
<tr>
<td>Embryonic skeletal muscle myoblasts</td>
<td>HSMMM_emb</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Embryonic kidney cells</td>
<td>HEK293</td>
<td>3</td>
<td>13</td>
</tr>
<tr>
<td>Primary kidney tissue (1 donor)</td>
<td>BC_Kidney_01-11002</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Renal cortical epithelial cells</td>
<td>HRCEpIC</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>Renal epithelial cells</td>
<td>HRE</td>
<td>3</td>
<td>10</td>
</tr>
<tr>
<td>Renal glomerular endothelial cells</td>
<td>HRGEc</td>
<td>3</td>
<td>2</td>
</tr>
</tbody>
</table>


By integrating ENCODE and gene-related information genome-wide, a researcher can use a Web-based browser, such as the UCSC genome browser, to functionally annotate a gene region, gene, or lead SNP by layering complex information across regions identified by genome-wide scans for cardiovascular traits.

To illustrate the potential for cardiovascular researchers, we annotated 2 validated blood pressure loci; 1 where there was a strong biologically plausible candidate gene in the associated interval, NPR3 (natriuretic peptide C receptor), and a second locus located on chromosome 1 (MOV10), where there was no obvious candidate gene.20

The MOV10 locus clearly illustrates the potential value that ENCODE annotation can bring to the characterization of GWAS results (Online Figure I). One of the previously reported lead SNPs (rs2932538) is located 500 base pair upstream of MOV10, in the putative promoter region. Review of ENCODE data in the UCSC genome browser shows good evidence of DNA binding in the SNP region in multiple CV-relevant cell lines. On the basis of DNase I activity, the SNP is predicted to lie within a strong enhancer region, this is likely to be cis acting on MOV10, but equally it could be acting in a trans manner on a remote gene. The sentinel SNP is annotated within a CCCTC-binding factor–binding site, identified by ChIP-seq. We note that the CCCTC-binding factor has also been shown to regulate the renin gene21 and the angiotensin-converting enzyme 2 gene,22 CCCTC-binding factor–binding sites are ubiquitous in the genome; however, this regulatory mechanism might be worth further investigation.

By contrast, Online Figure II shows how ENCODE can highlight evidence of cell-type–specific regulatory elements. In this case, the lead SNP (rs1173771) is located 20 kb downstream of NPR3. Histone marks in 7 cell lines show some evidence of regulatory activity (from embryonic and adult skeletal muscle myoblasts) in the region of the lead SNP, but not in the region of other SNPs in LD (r^2>0.8). When DNase I activity is reviewed in a range of cell lines, a strong signal in the lead SNP region is seen only in aortic smooth muscle and skeletal muscle myoblasts. This suggests a muscle-specific regulatory element downstream of NPR3. The data produced by the ENCODE project offer a remarkably fine detailed view of genomic function, importantly in a range of cell types, that are relevant to the biology of specific diseases, including cardiovascular disease. The biological activity suggested by ENCODE data would require verification by further experiments in targeted cell models or possibly in vivo; however, the detail and granularity of the data presented help immensely to focus further experimentation toward highly specific hypotheses, such as the muscle-specific element suggested in the NPR3 locus. To gain some insight into potential functionalities of published blood pressure SNPs, we have used ENCODE and pathway analysis tools to annotate each SNP and its close proxies, the results are presented in Table 2.

Pathway Analysis as a Tool to Explore the Therapeutic Potential of GWAS

Some from the pharmaceutical industry view hypertension as an area, where further innovation is unnecessary and unlikely to be adopted because of cost. However, studies in Western countries have revealed continuing poor blood pressure control and burgeoning epidemic in emergent economies. This is a complex problem, reflecting poor application of current therapies to control blood pressure, unresponsiveness, or variable uptake by patients who are sometimes noncompliant or importantly intolerant to available medicines. In the era of genomic medicine, it should be possible to examine whether variants at gene loci could inform therapeutic targeting or stratification of patient response.

Exploration of the genes discovered to date gives a partial picture of the genetic architecture of blood pressure. Nonetheless results do hint that a Guytonian-type complex circuit for blood pressure regulation can begin to be populated. Furthermore, this circuit will probably include many systems that had showed no prior evidence for a role in blood pressure regulation. Studies of Mendelian forms of hypertension have highlighted some potential therapeutic targets for exploitation (eg, the ROMK and the WNK kinase phosphorylation cascade). These findings give greater prominence to targets, such as the epithelial sodium channel, which causes Liddle syndrome (autosomal dominant, early onset, hypokalemia hypertension). This sodium channel is the site of action for amiloride which we rarely use for hypertension but perhaps we should explore further. From our genetic findings of blood pressure in the population, we can highlight the nitric oxide-natriuretic signaling pathway as being of interest;
These findings along with therapeutic innovations, such as use of inorganic nitrate to lower blood pressure, warrant further investigation (see pathway analysis below).

With the current resources available on common genetic variants, functional data from ENCODE, and pathway analysis tools, we can more thoroughly interrogate the loci discovered to date, in terms of known function and therapeutic potential. These analyses may identify alternative targets that may be more amenable to safe therapeutic targeting, where gene loci encode key regulatory points which may not be safe to target or lead to significant risk of adverse events because of off-target effects.

Pathway and network analysis is a powerful tool to apply GWAS findings, and they can be used in multiple complementary ways. Perhaps the most frequent application of pathway analysis is that focused on the identification of significantly enriched pathways in a genome-wide dataset. Such analyses are able to detect group behavior among gene signals, perhaps linking genome-wide significant associations or expression changes, with perisignificant results in the data set. There is value in such approaches; however, they are prone to bias on the basis of nature of the input data set, (eg, in the case of regions of genetic association, genes colocated on chromosomes often share function, because of gene duplication events). There are also underlying biases in the curation of pathway databases in key areas of biomedical research (eg, cancer research is highly represented in the literature and hence in pathway databases). More often overlooked is the use of pathway analysis for data annotation, drawing on the richly populated body of literature captured within pathway analysis databases, enabling, for example, the identification of indirect interactions between genes. Often information on this nature could not be found by direct literature searching, and hence it can represent a valuable tool for further expanding knowledge around a set of genes. It should be noted, however, that reported interactions may need further validation in the laboratory.

To demonstrate an example of such a process, we have used the main existing classes of antihypertensive medications in regular clinical use, to identify a network of genes on which these medications are known to act (directly or indirectly). We then explored their relationship with genes associated with blood pressure by GWAS, some of which, given the caveat of extended LD discussed earlier may contribute to the genetic architecture of blood pressure. The existing medicines target a priori known physiological systems, and act by distinct mechanisms differing in efficacy and side effect profiles among individual patients according to their underlying disease pathology and genetic background (Online Table 1). Pathway analysis offers a powerful tool to infer direct and indirect interactions (connections) between large numbers of genes, derived from different sources (Figure 2A). To do this, we used a comprehensively curated commercial database, GeneGo MetaCore (Thomson Reuters, United Kingdom), to identify all known gene interactions of the 9 main classes of antihypertensive drugs—our so-called drug interactome. We then defined a list of BP GWAS-associated genes at the 29 loci for blood pressure (on the basis of the LD of genic markers \( r^2>0.5 \) published by the ICBP-GWAS study). Both lists of genes were used to seed a gene interaction network in MetaCore, on the basis of direct or once removed interaction between genes in the network (Figure 2B).

Our analysis identified interaction between 72/301 GWAS-associated genes and 144/224 antihypertensive drug interactors. Four of the 72 GWAS-associated genes, PDGFRα, PGR, ALDH2, and ADRB1, are targeted by known drugs across a range of diseases (PDGFRα, imatinib mesylate, for cancer; PGR, progesterone agonists, various; ALDH2, disulfiram, for alcohol excess; ADRB1, adrenergic receptor antagonists, hypertension). At first glance 2 of the 3 disease/drug groupings may seem unrelated with blood pressure, but the ALDH2 locus is of interest as Kato et al reported strong association between the gene in East Asians and an independent association also seems to be present in our European data. The ALDH2 inhibitor, disulfiram, is used to treat alcoholism, but it is well established that the drug induces hypertension with extended use. In contrast, antihypertensive molecules, such as Amyl Nitrite, are known to be ALDH2 agonists with vasodilator effects.

Twenty-seven of the 72 genes we identified have no prior support in hypertension and can, therefore, be considered as novel hypertension candidates. Six of these novel candidate genes are potentially druggable, determined using the Chembl database: https://www.ebi.ac.uk/chembl/. In Figure 2C, we classify the GWAS-drug interactome network by each class of antihypertensive drug. This shows some of the insight that could be gained into the clinical use of these classes of compounds, given more detailed analysis. The even distribution of genes across each drug mechanism also underscores the direct link between disease causality and therapeutic effect, and suggests that population-level disease susceptibility variants may have a relatively even distribution across antihypertensive therapeutic mechanisms. As more data emerge, it may be possible to stratify therapeutic nonresponders by deleterious genetic load on particular pathways involved in the regulation of blood pressure. Perhaps the most interesting observation relates to the 41 potentially druggable genes that are associated with BP by GWAS, but do not overlap with the known drug interactome. Because LD was used to determine GWAS association, many of these genes will not be causally linked to blood pressure; however, those that are causally linked (perhaps with some help from ENCODE analysis) could represent novel disease mechanisms that are not targeted by current drugs.

By viewing drug mechanism–disease interactions at a network level, there may be potential to improve understanding of the action and side effects of a drug. In addition, we may possibly identify new therapeutic intervention points to create new therapies or synergise with existing therapies.

As we alluded to earlier, our analysis has a key caveat—although the GWAS loci described here have been adjusted for multiple testing and have been validated, it is not possible in most cases to pinpoint the causal gene among several genes in the average locus. Instead we have included all genes with evidence of LD \( r^2>0.5 \) with the associated marker. Although this will introduce many genes into our analysis which are not truly associated, it also serves the purpose of providing further
annotation and support for those genes that interact with anti-hypertensive drugs at a direct level or network level.

This preliminary analysis may offer important pointers toward the development of new molecular taxonomies of hypertension and the different classes of drugs that target this disease. This will be an important step toward more efficacious interventions at first using a stratified approach, with the ultimate goal of personalized medicine.

**Blood Pressure Loci and Cardiovascular Risk Assessment**

A further argument that might be deployed in support of a panel of gene variants could be to add value to risk prediction in the young. These individuals are less well served by current conventional risk assessment tools that integrate multiple risk factors and allow computation of cardiovascular risk over a 10-year period. Using such approaches the young (<40 years) may not reach thresholds of risk that require treatment as their cardiovascular risk is <20% of an event >10 years. However, tools are now emerging that permit the computation of lifetime risk comparable with the approach used by actuarial life-tables, which offers a clearer picture of life-course risk, http://www.qrisk.org/lifetime/. In spite of the impact of blood pressure gene loci within a risk score on cardiovascular outcomes (above) at this time, the modest added value of blood pressure loci to these risk assessments seem unlikely to supersede the value of direct measurement blood pressure, cholesterol, or smoking in the assessment of cardiovascular risk but we cannot entirely rule out a role in the future especially among the young.

**Stratifying Antihypertensive Response by Genotype**

There is a limited pharmacogenomic literature for antihypertensives and very limited validation of any of the findings. There are cardiovascular medications with appreciable variation in response by genotype or risk of adverse effect. In blood pressure, the Mendelian hypertensive syndromes do
demonstrate to some extent the potential for stratified medicine. In the glucocorticoid suppressible hyperaldosteronism syndrome, a chimeric gene combines the regulatory sequence of 11β-hydroxylase with the coding region of the aldosterone gene. This brings production of this sodium-losing hormone under adrenocorticotrophic hormone control, which is suppressed by steroids. In Liddle syndrome, mutations of the epithelial sodium channel enhance sodium retention and loss of potassium. This channel is the target for amiloride which very effectively treats the hypertensive phenotype of Liddle syndrome. At this time, there is no panel of genotypes which we could deploy to reliably stratify patient response in the clinic, but this does not rule out that possibility.

Mendelian Randomization in Cardiovascular Disease
Mendelian randomization allows us to test for a causal effect of a genotype on a trait of interest from observational data even in the presence of confounding factors. It uses Mendel Second Law, which states that genotypes are randomly assigned during meiosis when transmitted from parents to offspring. In this circumstance, the population genotype distribution should be unrelated to the confounders that typically plague observational epidemiology studies. This means that Mendelian randomization can be thought of as a natural randomized controlled trial, where the common genetic variant acts as an instrument for the exposure of interest and must only affect the disease status indirectly via its effect on the exposure of interest.

Mendelian randomization offers an important opportunity to distinguish between gene variants, which identify a causal pathway for disease and offer greatest predictive potential, and may also reveal potential interventional targets compared with gene variants where the association is caused by reverse causation or confounding. Recent application of this approach in cardiovascular traits suggests that association of C-reactive protein and high-density lipoprotein are unlikely to cause coronary heart disease. Interestingly, for the interleukin 6 receptor it suggests an association with coronary disease and has led to discussions around use of IL6 monoclonal in acute coronary syndromes. This approach is less valuable when there is genetic heterogeneity, pleiotropic effects of genotype on multiple traits, or population substructure. The method has yet to be effectively deployed in blood pressure genetics, but this may be useful at identifying the therapeutic potential of pathways identified by gene discovery.

Future Directions and Translational Potential of Genomics for Blood Pressure
A key conundrum for almost all common complex cardiovascular diseases is that our understanding of the genetic architecture of these traits explains only a small proportion of the heritability. This has prompted consideration of where to search for the missing heritability. Earlier, we indicated that for the given effect size of the blood pressure variants detected to date we might expect to double the number of variants with a similar effect for an effective doubling of the sample size. The application of dense gene-centric arrays is increasing gene discovery; this and other approaches will be required to discover a greater proportion of heritability.

With the arrival of exome chips and the use of NGS the search will refocus from common to rare variants. At the present time, the optimal strategy for NGS of either exomes or the whole genome is yet to be established. These approaches offer an unbiased potential to identify rare variants which may have a greater effect size, and in the future move discovery from populations to individual profiles of risk for at least some cardiovascular disorders. It is also possible as we have seen from Mendelian hypertensive phenotypes that rare variants may help to stratify subsets of individuals with specific treatment responses, but the clinical use at an individual level and affordability remains to be proven. At this time, a key issue which needs to be addressed is the challenge of detecting and analysis of rare variants with the fidelity and reproducibility necessary for clinical use with the present technologies. A second challenge is the affordability of NGS when deployed at scale but akin to chip-based genotyping this will reduce in price.

The NGS strategies are based on massive parallel sequencing, leading to repeated reads at a depth of 30- to 50-fold coverage, or even greater, depth to identify rare variants with high degrees of confidence. The optimal strategy is likely to be revealed by only direct and perhaps iterative experimentation. For exome-sequencing focused on coding regions, the strategy is becoming clearer, but the adequacy of capture of exon-based variants is still limited to 80% to 90% sensitivity, and possibly a specificity of 90% to 95% for rarer diseases which leaves open the risk of missing some causal variants. With whole-genome sequencing, the risk of false positives is amplified considerably by the high proportion of ≈75% of alleles in the population genome being singleton or doubleton. The risk inherent in the elimination of singleton and doubleton variants is that important causal variants with large effect sizes could be missed. A further problem highlighted by the gene encoding the sarcomeric protein titin is that even when a series of rare variants which may encode nonsynonymous amino acid change are identified how do you delineate those with clinical relevance. Mendelian randomization will be of limited help here because the allele is very rare and thus too infrequent to triangulate or refute causality in the population except in extremely large population resources.

Somatic Mutations in the Adrenal Causing Hypertension
The potential of exome-focused NGS to identify new causes of hypertension is beautifully demonstrated by recent analysis of patients with endocrine hypertension caused by Conn syndrome. This can arise to aldosterone-secreting adrenocortical hyperplasia or a discrete adenoma. In a very elegant series of publications somatic mutations of KCNJ5 (encoding a potassium channel and ATP1A1 [encoding an Na+/K+ ATPase α subunit] and ATP2B3 [encoding a Ca2+ ATPase]) genes have been demonstrated to affect aldosterone secretion. It is possible that such somatic mutations leading to adrenal adenomas or hyperplasia may account for 5% of hypertension.
NGS in Large-Scale Studies and Adoption Into Health Care

There are studies based on NGS of individuals with blood pressure measurements at the population extremes underway. There is also an opportunity to test the value of whole-genome NGS within UK Biobank. This is a cohort of 500,000 people, aged 40 to 69 years, who have blood pressure measurements, and of which 100,000 are undergoing deeper phenotyping for cardiovascular phenotypes (including cardiac imaging). This resource has been created alongside other Biobanks to offer the scale to examine gene and environment interaction. The sampling strategy for UK Biobank enables the integration of genotypic, metabonomic, and proteomic biomarkers. This multi-omic approach may prove profitable in gene identification or refining patient response to medication by defining intermediary phenotypes, which are closer to the gene of interest than blood pressure. The ICBP-GW AS study attempted to add value by incorporating a limited metabonomic panel alongside blood pressure gene variants but this did not add any new observations.

In a complementary but distinct strategy the UK National Health Service will use NGS to sequence 100,000 individuals with rare diseases (CV and other disorders) and common cancers. This is with the goal of developing new diagnostics and therapeutics indicating that health services are adopting these technologies.

We have described multiple complimentary strategies to detect gene loci for blood pressure and cardiovascular risk, and the new advances in our understanding of the architecture of the human genome to aid further locus discovery. The potential of these strategies will be to facilitate exploration of new pathways or repurpose existing molecules, and contribute to stratification of patient response to antihypertensives. In time with additional discoveries facilitated by NGS strategies, we may be in a position to create a comprehensive genomic Guytonian circuit for blood pressure which may inform therapeutic development and improved control of blood pressure which is the commonest cardiovascular risk factor worldwide.

Acknowledgments

This work forms part of the research themes contributing to the translational research portfolio of the National Institute for Health Cardiovascular Biomedical Research Unit at Barts, Michael Barnes is funded by this award. We thank the Bioinformatics Group at Barts and The London Genome Center.

Sources of Funding

Our research is funded by the Medical Research Council of Great Britain (G9521010D), the British Heart Foundation (PG/02/128), the Wellcome Trust, National Institute for Health Research. Mark Caulfield and Patricia B. Munroe are Wellcome Trust, National Institute for Health Research. Mark Caulfield, and the Wellcome Trust Case Control Consortium. Echogen consortium; AortaGen Consortium; CHARGE Consortium Heart Failure Working Group; KidneyGen consortium; KidneyGen consortium; CardioGen consortium; CardioGram. Genome-wide association study identifies six new loci influencing blood pressure and mean arterial pressure. Nat Genet 2011;43:531–538.


References


12. Wellcome Trust Case Control Consortium W. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature. 2007;447:661–678.


Disclosures

None.


Am J Physiol Heart. Hydrogenase mediates vasodilator responses of glyceryl trinitrate and so-


BMJ. and coronary heart disease: Mendelian randomisation analysis based on

Hum Mol Genet. Pathway and network-based analysis of genome-wide association studies


ICSGfB-l. Mutations in the gene encoding the inwardly-rectifying renal

Bartter's syndrome, hypokalaemic alkalosis with hypercalciuria, is


Science. 2011;331:768–772.


Am J Physiol Heart. Hydrogenase mediates vasodilator responses of glyceryl trinitrate and so-


Am J Physiol Heart. Hydrogenase mediates vasodilator responses of glyceryl trinitrate and so-


Am J Physiol Heart. Hydrogenase mediates vasodilator responses of glyceryl trinitrate and so-


Advances in Blood Pressure Genomics
Patricia B. Munroe, Michael R. Barnes and Mark J. Caulfield

Circ Res. 2013;112:1365-1379
doi: 10.1161/CIRCRESAHA.112.300387

Circulation Research is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2013 American Heart Association, Inc. All rights reserved.
Print ISSN: 0009-7330. Online ISSN: 1524-4571

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circres.ahajournals.org/content/112/10/1365

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Circulation Research can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at: http://www.lww.com/reprints

Subscriptions: Information about subscribing to Circulation Research is online at: http://circres.ahajournals.org//subscriptions/