Trans-ancestry meta-analyses identify novel rare and common variants associated with blood pressure and hypertension

Praveen Surendran [1,167], Fotios Drenos [2,3,167], Robin Young [1,167], Helen Warren [4,5,167], James P Cook [6,7,167], Alisa K Manning [8,9,10,167], Niels Grarup [11,167], Xueling Sim [12,13,14,167], Daniel R Barnes [1], Kate Witkowska [4,5], James R Staley [1], Vinicius Tragante [15], Taru Tukiainen [8,9,16], Hanieh Yaghootkar [17], Nicholas Masca [18,19], Daniel F Freitag [1], Teresa Ferreira [20], Olga Giannakopoulou [21], Andrew Tinker [21,5], Magdalena Harakalova [15], Evelin Mihailov [22], Chunyu Liu [23], Aldi T Kraja [24,25], Sune Fallgaard Nielsen [26], Asif Rasheed [27], Maria Samuel [27], Wei Zhao [28], Lori L Bonnycastle [29], Anne U Jackson [13,12], Narisu Narisu [29], Amy J Swift [29], Lorraine Southam [30,20], Jonathan Marten [31], Jeroen R Huyghe [13,12], Alena Stančáková [32], Cristiano Fava [33,34], Therese Ohlsson [33], Angela Matchan [30], Kathleen E Stirrups [21,35], Jette Bork-Jensen [11], Anette P Gjesing [11], Jukka Kontto [36], Markus Perola [36,37,22], Susan Shaw-Hawkins [4], Aki S Havulinna [36], He Zhang [38], Louise A Donnelly [39], Christopher J Groves [40], N William Rayner [40,20,30], Matt J Neville [40,41], Neil R Robertson [20,40], Andrianos M Yiorkas [42,43], Karl-Heinz Herzig [44,45], Eero Kajantie [36,46,47], Weihua Zhang [48,49], Sara M Willems [50], Lars Lannfelt [51], Giovanni Malerba [52], Nicole Soranzo [53,35,54], Elisabetta Trabetti [52], Niek Verweij [55,9,56], Evangelos Evangelou [48,57], Alireza Moayyeri [48,58], Anne-Claire Vergnaud [48], Christopher P Nelson [18,19], Alaitz Poveda [59,60], Tibor V Varga [59], Muriel Caslake [61], Anton JM de Craen [62,63], Stella Trompet [62,64], Jian'an Luan [50], Robert A Scott [50], Sarah E Harris [65,66], David CM Liewald [65,67], Riccardo Marioni [65,66,68], Cristina Menni [69], Aliki-Eleni Farmaki [70], Göran Hallmans [71], Frida Renström [59,71], Jennifer E Huffman [31,23], Maija Hassinen [72], Stephen Burgess [1], Ramachandran S Vasan [23,73,74], Janine F Felix [75], CHARGE-Heart Failure Consortium [76], Maria Uria-Nickelsen [77], Anders Malarstig [78], Dermot F Reily [79], Maarten Hoek [80], Thomas Vogt [80,81], Honghuang Lin [23,82], Wolfgang Lieb [83], EchoGen Consortium [76], Matthew Traylor [84], Hugh F Markus [84], METASTROKE Consortium [76], Heather M Highland [85], Anne E Justice [85], Eirini Marouli [21], GIANT Consortium [76], Jaana Lindström [36], Matti Uusitupa [86,87], Pirjo Komulainen [72], Timo A Lakka [72,88,89], Rainer Rauramaa [72,89], Ozren Polasek [90,91], Igor Rudan [90], Olov Rolandsson [92], Paul W Franks [59,92,93], George Dedoussis [70], Timothy D Spector [69], EPIC-InterAct Consortium [76], Pekka Jousilahti [36], Satu Männistö [36], Ian J Deary [65,67], John M Starr [65,94], Claudia Langenberg [50], Nick J Wareham [50], Morris J Brown [4], Anna F Dominiczak [95], John M Connell [39], J Wouter Jukema [64,96], Naveed Sattar [95], Ian Ford [61], Chris J Packard [61], Tõnu Esko [22,97,8,9], Reedik Mägi [22], Andres Metspalu [22,98], Rudolf A de Boer [99], Peter van der Meer [99], Pim van der Harst [99,100,101], Lifelines Cohort Study [76], Giovanni Gambaro [102], Erik Ingelsson [103,104], Lars Lind [103], Paul IW de Bakker [105,106], Mattijs E Numans [107,106], Ivan Brandslund [108,109], Cramer Christensen [110], Eva RB Petersen [111], Eeva Korpi-Hyövälti [112], Heikki Oksa [113], John C Chambers [48,49,114], Jaspal S Kooner [49,115,114], Alexandra IF Blakemore [42,43], Steve Franks [116], Marjo-Riitta Jarvelin [117,118,119,120], Lise L Husemoen [121], Allan Linneberg [121,122,123], Tea Skaaby [121], Betina Thuesen [121], Fredrik Karpe [40,41], Jaakko Tuomilehto [36,124,125,126], Alex SF Doney [39], Andrew D Morris [127], Colin NA Palmer [39], Oddgeir Lingaas Holmen [128,129], Kristian Hveem [128,130], Cristen J Willer [38,131,132], Tiinamaija Tuomi [133,134], Leif Groop [135,134], AnneMari Käräjämäki [136,137], Aarno Palotie [16,9,134,138], Samuli Ripatti [134,139,30], Veikko Salomaa [36], Dewan S Alam [140], Abdulla al Shafi Majumder [141], Emanuele Di Angelantonio [1,54], Rajiv Chowdhury [1], Mark I McCarthy [40,41,20], Neil Poulter [142], Alice V Stanton [143], Peter Sever [142], Philippe Amouyel [144,145,146,147], Dominique Arveiler [148], Stefan Blankenberg [149,150], Jean Ferrières [151], Frank Kee [152], Kari Kuulasmaa [36], Martina Müller-Nurasyid [153,154,155], Giovanni Veronesi [156], Jarmo Virtamo [36], Panos Deloukas [21,157], Wellcome Trust Case Control Consortium [76], Paul Elliott [117], Understanding Society Scientific Group [76], Eleftheria Zeggini [30], Sekar Kathiresan [56,158,159,9], Olle Melander [33], Johanna Kuusisto [32], Markku Laakso [32], Sandosh Padmanabhan [95], David Porteous [66], Caroline Hayward [31], Generation Scotland [160], Francis S Collins [29], Karen L Mohlke [161], Torben Hansen [11], Oluf Pedersen [11], Michael Boehnke [13,12], Heather M Stringham [13,12], EPIC-CVD Consortium [76], Philippe Frossard [27], Christopher Newton-Cheh [56,158], CHARGE+ Exome Chip Blood Pressure Consortium [76], Martin D Tobin [6,

47   Børge Grønne Nordestgaard [26], T2D-GENES Consortium [76], GoT2DGenes Consortium [76], ExomeBP

48   Consortium [76], CHD Exome+ Consortium [76], Mark J Caulfield [4,5], Anubha Mahajan [20], Andrew P Morris

49   [20,7], Maciej Tomaszewski [18,19,162], Nilesh J Samani [18,19], Danish Saleheen [28,27,1,167], Folkert W Asselbergs

50   [15,101,163,167], Cecilia M Lindgren [164,9,20,167], John Danesh [1,165,54,167], Louise V Wain [6,167], Adam S Butterworth

51   [1,166,167], Joanna MM Howson [1,167,168], Patricia B Munroe [4,5,167,168]

52

53   1. Cardiovascular Epidemiology Unit, Department of Public Health and Primary Care, University of

54   Cambridge, Cambridge, UK

55   2. Medical Research Council Integrative Epidemiology Unit, School of Social and Community Medicine,

56   University of Bristol, Oakfield House, Oakfield Grove, Bristol, UK

57   3. Centre for Cardiovascular Genetics, Institute of Cardiovascular Science, Rayne Building University

58   College London, London, UK

59   4. Clinical Pharmacology, William Harvey Research Institute, Queen Mary University of London, London,

60   UK

61   5. National Institute for Health Research Barts Cardiovascular Biomedical Research Unit, Queen Mary

62   University of London, London, UK

63   6. Department of Health Sciences, University of Leicester, Leicester, UK

64   7. Department of Biostatistics, University of Liverpool, Liverpool, UK

65   8. Department of Genetics, Harvard Medical School, Boston, Massachusetts, USA

66   9. Program in Medical and Population Genetics, Broad Institute, 7 Cambridge Center, Cambridge,

67   Massachusetts, USA

68   10. Department of Molecular Biology, Massachusetts General Hospital, Boston, Massachusetts, USA

69   11. The Novo Nordisk Foundation Center for Basic Metabolic Research, Faculty of Health and Medical

70   Sciences, University of Copenhagen, Copenhagen, Denmark

71   12. Center for Statistical Genetics, University of Michigan, Ann Arbor, Michigan, USA

72   13. Department of Biostatistics, University of Michigan, Ann Arbor, Michigan, USA

73   14. Saw Swee Hock School of Public Health, National University of Singapore, National University Health

74   System, Singapore

75   15. Department of Cardiology, University Medical Center Utrecht, Utrecht, The Netherlands

76   16. Analytic and Translational Genetics Unit, Department of Medicine, Massachusetts General Hospital,

77   Boston, Massachusetts, USA

78   17. Genetics of Complex Traits, Institute of Biomedical and Clinical Science, University of Exeter Medical

79   School, Exeter, UK

80   18. Department of Cardiovascular Sciences, University of Leicester, Leicester, UK

81   19. National Institute for Health Research Leicester Biomedical Research Unit in Cardiovascular Disease,

82   Leicester, UK

83   20. Wellcome Trust Centre for Human Genetics, Nuffield Department of Medicine, University of Oxford,

84   Oxford, UK

85   21. Heart Centre, William Harvey Research Institute, Barts and The London School of Medicine and

86   Dentistry, Queen Mary University of London, London, UK

87   22. Estonian Genome Center, University of Tartu, Tartu, Estonia

88   23. National Heart, Lung, and Blood Institute's and Boston University's Framingham Heart Study,

89   Framingham, Massachusetts, USA

90   24. Division of Statistical Genomics, Center for Genome Sciences and Systems Biology, Washington

91   University School of Medicine, St. Louis, Missouri, USA

92   25. Department of Genetics, Washington University School of Medicine, St. Louis, Missouri, USA

93   26. Department of Clinical Biochemistry Herlev Hospital, Copenhagen University Hospital, Herlev,

94   Denmark

95   27. Centre for Non-Communicable Diseases, Karachi, Pakistan

96   28. Department of Biostatistics and Epidemiology, Perelman School of Medicine, University of

97   Pennsylvania, Philadelphia, Pennsylvania, USA

98   29. Medical Genomics and Metabolic Genetics Branch, National Human Genome Research Institute, NIH,

99   Bethesda, Maryland, USA

100   30. Wellcome Trust Sanger Institute, Genome Campus, Hinxton, UK

101   31. Medical Research Council Human Genetics Unit, Medical Research Council Institute of Genetics and

102   Molecular Medicine, University of Edinburgh, Edinburgh, UK

103   32. Department of Medicine, University of Eastern Finland and Kuopio University Hospital, Kuopio,

104   Finland

105   33. University of Lund, Department of Clinical Sciences, Malmö, Sweden

106   34. University of Verona, Department of Medicine, Verona, Italy

107   35. Department of Haematology, University of Cambridge, Cambridge, UK

108   36. Department of Health, National Institute for Health and Welfare, Helsinki, Finland

109   37. Institute of Molecular Medicine FIMM, University of Helsinki, Finland

110   38. Department of Internal Medicine, Division of Cardiovascular Medicine, University of Michigan, Ann

111   Arbor, Michigan, USA

112   39. Medical Research Institute, University of Dundee, Ninewells Hospital and Medical School, Dundee, UK

113   40. Oxford Centre for Diabetes, Endocrinology, and Metabolism, Radcliffe Department of Medicine,

114   University of Oxford, Oxford, UK

115   41. National Institute for Health Research Oxford Biomedical Research Centre, Oxford University Hospital

116   Trusts, Oxford, UK

117   42. Section of Investigative Medicine, Imperial College London, London, UK

118   43. Department of Life Sciences, Brunel University London, London, UK

119   44. Institute of Biomedicine, Biocenter Oulu, University of Oulu, Oulu, Finland

120   45. Department of Gastroenterology and Metabolism, Poznan University of Medical Sciences, Poznan,

121   Poland

122   46. Hospital for Children and Adolescents, Helsinki University Central Hospital and University of Helsinki,

123   Helsinki, Finland

124   47. Department of Obstetrics and Gynaecology, Oulu University Hospital and University of Oulu, Oulu,

125   Finland

126   48. Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London,

127   London, UK

128   49. Department of Cardiology, Ealing Hospital, Middlesex, UK

129   50. Medical Research Council Epidemiology Unit, University of Cambridge School of Clinical Medicine,

130   Box 285 Institute of Metabolic Science, Cambridge Biomedical Campus, Cambridge, UK

131   51. Department of Public Health and Caring Sciences, Uppsala University, Uppsala, Sweden

132   52. Section of Biology and Genetics, Department of Neurosciences, Biomedicine and Movement Sciences,

133   University of Verona, Verona, Italy

134   53. Human Genetics, Wellcome Trust Sanger Institute, Hinxton, UK

135   54. The National Institute for Health Research Blood and Transplant Research Unit in Donor Health and

136   Genomics, University of Cambridge, Cambridge, UK

137   55. University Medical Center Groningen, University of Groningen, Department of Cardiology, The

138   Netherlands

139   56. Center for Human Genetic Research, Massachusetts General Hospital, Boston, Massachusetts, USA

140   57. Department of Hygiene and Epidemiology, University of Ioannina Medical School, Ioannina, Greece

141   58. Farr Institute of Health Informatics Research, Institute of Health Informatics, University College

London, London, UK

59. Department of Clinical Sciences, Genetic and Molecular Epidemiology Unit, Lund University, Malmö, Sweden

60. Department of Genetics, Physical Anthropology and Animal Physiology, Faculty of Science and Technology, University of the Basque Country (UPV/EHU), Bilbao, Spain

61. University of Glasgow, Glasgow, UK

62. Department of Gerontology and Geriatrics, Leiden University Medical Center, Leiden, The Netherlands

63. Mr. De Craen suddenly passed away January 2016

64. Department of Cardiology, Leiden University Medical Center, Leiden, The Netherlands

65. Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, Edinburgh, UK

66. Centre for Genomic and Experimental Medicine, Medical Research Council Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, UK

67. Department of Psychology, University of Edinburgh, Edinburgh, UK

68. Queensland Brain Institute, The University of Queensland, Brisbane, Queensland, Australia

69. Department of Twin Research and Genetic Epidemiology, King's College London, UK

70. Department of Nutrition and Dietetics, School of Health Science and Education, Harokopio University, Athens, Greece

71. Department of Biobank Research, Umeå University, Umeå, Sweden

72. Kuopio Research Institute of Exercise Medicine, Kuopio, Finland

73. Section of Cardiology, Department of Medicine, Boston University Schools of Medicine and Public Health, Boston, Massachusetts, USA

74. Sections of Preventive Medicine and Epidemiology, Department of Medicine, Boston University Schools of Medicine and Public Health, Boston, Massachusetts, USA

75. Department of Epidemiology, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands

76. A full list of members and affiliations appears in the Supplementary Note

77. Development Management and Planning, Pfizer Worldwide Research and Development

78. Pfizer Worldwide Research and Development, Stockholm, Sweden

79. Merck Research Laboratories, Genetics and Pharmacogenomics, Boston, Massachusetts, USA.

80. Merck Research Laboratories, Cardiometabolic Disease, Kenilworth, New Jersey, USA

81. CHDI Management/CHDI Foundation, Princeton, New Jersey, USA

82. Section of Computational Biomedicine, Department of Medicine, Boston University School of Medicine, Boston, Massachusetts, USA

83. Institute of Epidemiology and Biobank Popgen, Kiel University, Kiel, Germany

84. Neurology Unit, University of Cambridge, Cambridge Biomedical Campus, Cambridge, UK

85. University of North Carolina at Chapel Hill, Department of Epidemiology, Chapel Hill, North Carolina, USA

86. Department of Public Health and Clinical Nutrition, University of Eastern Finland, Finland

87. Research Unit, Kuopio University Hospital, Kuopio, Finland

88. Institute of Biomedicine/Physiology, University of Eastern Finland, Kuopio Campus, Finland

89. Department of Clinical Physiology and Nuclear Medicine, Kuopio University Hospital, Kuopio, Finland

90. Centre for Global Health Research, Usher Institute for Population Health Sciences and Informatics, University of Edinburgh, Edinburgh, UK

91. Faculty of Medicine, University of Split, Croatia

92. Department of Public Health and Clinical Medicine, Umeå University, Umeå, Sweden

93. Department of Nutrition, Harvard School of Public Health, Boston, Massachusetts, USA

94. Alzheimer Scotland Research Centre, University of Edinburgh, Edinburgh, UK

95. Institute of Cardiovascular and Medical Sciences, College of Medical, Veterinary and Life Sciences,

University of Glasgow, Glasgow, UK

96. The Interuniversity Cardiology Institute of the Netherlands, Utrecht, The Netherlands

97. Division of Endocrinology, Boston Children's Hospital, Boston, Massachusetts, USA

98. Institute of Molecular and Cell Biology, Tartu, Estonia

99. Department of Cardiology, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands

100. Department of Genetics, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands

101. Durrer Center for Cardiogenetic Research, ICIN-Netherlands Heart Institute, Utrecht, The Netherlands

102. Division of Nephrology, Department of Internal Medicine and Medical Specialties, Columbus - Gemelli University Hospital, Catholic University, Rome, Italy

103. Department of Medical Sciences, Molecular Epidemiology and Science for Life Laboratory, Uppsala University, Uppsala, Sweden

104. Department of Medicine, Division of Cardiovascular Medicine, Stanford University School of Medicine, Stanford, California, USA

105. Department of Medical Genetics, Center for Molecular Medicine, University Medical Center Utrecht, Utrecht, The Netherlands

106. Department of Epidemiology, Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, The Netherlands

107. Department of Public Health and Primary Care, Leiden University Medical Center, Leiden, The Netherlands

108. Department of Clinical Biochemistry, Lillebaelt Hospital, Vejle, Denmark

109. Institute of Regional Health Research, University of Southern Denmark, Odense, Denmark

110. Medical Department, Lillebaelt Hospital, Vejle, Denmark

111. Department of Clinical Immunology and Biochemistry, Lillebaelt Hospital, Vejle, Denmark

112. South Ostrobothnia Central Hospital, Seinäjoki, Finland

113. Tampere University Hospital, Tampere, Finland

114. Imperial College Healthcare NHS Trust, London, UK

115. National Heart and Lung Institute, Imperial College London, London, UK

116. Institute of Reproductive and Developmental Biology, Imperial College London, London, UK

117. Department of Epidemiology and Biostatistics, Medical Research Council Public Health England Centre for Environment and Health, School of Public Health, Faculty of Medicine, Imperial College London, St. Mary's Campus, London, UK

118. Centre for Life Course Epidemiology, Faculty of Medicine, University of Oulu, Oulu, Finland

119. Biocenter Oulu, University of Oulu, Oulu, Finland

120. Unit of Primary Care, Oulu University Hospital, Oulu, Finland

121. Research Centre for Prevention and Health, Capital Region of Denmark, Copenhagen, Denmark

122. Department of Clinical Experimental Research, Glostrup University Hospital, Glostrup, Denmark

123. Department of Clinical Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark

124. Dasman Diabetes Institute, Dasman, Kuwait

125. Centre for Vascular Prevention, Danube-University Krems, Krems, Austria

126. King Abdulaziz University, Jeddah, Saudi Arabia

127. School of Molecular, Genetic and Population Health Sciences, University of Edinburgh, Medical School, Teviot Place, Edinburgh, UK

128. HUNT Research Centre, Department of Public Health and General Practice, Norwegian University of Science and Technology, Levanger, Norway

129. St. Olav Hospital, Trondheim University Hospital, Trondheim, Norway

130. Department of Medicine, Levanger Hospital, Nord- Trøndelag Health Trust, Levanger, Norway

131. Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, Michigan, USA

132. Department of Human Genetics, University of Michigan, Ann Arbor, Michigan, USA

133. Folkhälsan Research Centre, Helsinki, Finland; Department of Endocrinology, Helsinki University Central Hospital, Helsinki, Finland

134. Institute for Molecular Medicine Finland University of Helsinki, Helsinki, Finland

135. Department of Clinical Sciences, Diabetes and Endocrinology, Lund University Diabetes Centre, Malmö, Sweden.

136. Department of Primary Health Care, Vaasa Central Hospital, Vaasa, Finland

137. Diabetes Center, Vaasa Health Care Center, Vaasa, Finland

138. Psychiatric and Neurodevelopmental Genetics Unit, Department of Psychiatry, Massachusetts General Hospital, Boston, Massachusetts, USA

139. Department of Public Health, University of Helsinki, Finland

140. ICDDR, B; Mohakhali, Dhaka, Bangladesh

141. National Institute of Cardiovascular Diseases, Sher-e-Bangla Nagar, Dhaka, Bangladesh

142. International Centre for Circulatory Health, Imperial College London, UK

143. Molecular and Cellular Therapeutics, Royal College of Surgeons in Ireland, Dublin, Ireland

144. University of Lille, UMR1167, Risk Factors and Molecular Determinants of aging-related diseases, Lille, France

145. Inserm, Lille, France

146. Centre Hospitalier Universitaire Lille, Public Health, Lille, France

147. Institut Pasteur de Lille, Lille, France

148. Department of Epidemiology and Public Health, EA 3430, University of Strasbourg, Strasbourg, France

149. Department of General and Interventional Cardiology, University Heart Center Hamburg, Germany

150. University Medical Center Hamburg-Eppendorf, Hamburg, Germany

151. Department of Epidemiology, UMR 1027- INSERM, Toulouse University-CHU Toulouse, Toulouse, France

152. Director, UKCRC Centre of Excellence for Public Health, Queens University, Belfast, Northern Ireland

153. Institute of Genetic Epidemiology, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg, Germany

154. Department of Medicine I, University Hospital Grosshadern, Ludwig-Maximilians-Universität, Munich, Germany

155. DZHK (German Centre for Cardiovascular Research), partner site Munich Heart Alliance, Munich, Germany

156. Research Center in Epidemiology and Preventive Medicine, Department of Clinical and Experimental Medicine, University of Insubria, Varese, Italy

157. Princess Al-Jawhara Al-Brahim Centre of Excellence in Research of Hereditary Disorders (PACER-HD), King Abdulaziz University, Jeddah, Saudi Arabia

158. Cardiovascular Research Center, Massachusetts General Hospital, Boston, Massachusetts, USA

159. Department of Medicine, Harvard Medical School, Boston, Massachusetts, USA

160. A Collaboration between the University Medical Schools and NHS, Aberdeen, Dundee, Edinburgh and Glasgow, UK

161. Department of Genetics, University of North Carolina, Chapel Hill, North Carolina, USA

162. Institute of Cardiovascular Sciences, University of Manchester, Manchester, UK

163. Faculty of Population Health Sciences, Institute of Cardiovascular Science, University College London, London, UK

286   164. The Big Data Institute at the Li Ka Shing Centre for Health Information and Discovery, University of
287   Oxford, Oxford, UK
288   165. Wellcome Trust Sanger Institute, Hinxton, UK
289   166. The National Institute for Health Research Blood and Transplant Research

290   167. Contributed equally to this work
291   168. Jointly supervised the work
292
293   Corresponding authors: Joanna M M Howson, jmmh2@medschl.cam.ac.uk and Patricia B Munroe,
294   p.b.munroe@qmul.ac.uk
295

296

297

Abstract

High blood pressure is a major risk factor for cardiovascular disease and premature death. However, there is limited knowledge on specific causal genes and pathways. To better understand the genetics of blood pressure, we genotyped 242,296 rare, low-frequency and common genetic variants in up to ~192,000 individuals, and used ~155,063 samples for independent replication. We identified 31 novel blood pressure or hypertension associated genetic regions in the general population, including three rare missense variants in *RBM47*, *COL21A1* and *RRAS* with larger effects (>1.5mmHg/allele) than common variants. Multiple rare, nonsense and missense variant associations were found in *A2ML1* and a low-frequency nonsense variant in *ENPEP* was identified. Our data extend the spectrum of allelic variation underlying blood pressure traits and hypertension, provide new insights into the pathophysiology of hypertension and indicate new targets for clinical intervention.

Introduction

High blood pressure (BP) or hypertension is a highly prevalent chronic disorder. It is estimated to be responsible for a larger proportion of global disease burden and premature mortality than any other disease risk factor[1]. Elevated systolic and/or diastolic BP increases the risk of several cardiovascular disorders including stroke, coronary heart disease (CHD), heart failure, peripheral arterial disease and abdominal aortic aneurysms[2]. BP is a complex, heritable, polygenic phenotype for which genome-wide association studies (GWAS) have identified over 67 genetic regions associated with BP and/or hypertension to date[3-11]. These variants are common (minor allele frequency, MAF≥0.05), mostly map to intronic or intergenic regions, with the causal alleles and genes not readily identified due to linkage disequilibrium (LD)[4,5], and explain only ~2% of trait variance[12]. Low-frequency (0.01<MAF<0.05) and rare (MAF≤0.01) single nucleotide variants (SNVs), predominantly unexplored by GWAS may have larger phenotypic effects than common SNVs[13], and may help to explain the missing heritability, and identify causative genes as demonstrated previously[14].

To identify novel coding variants and loci influencing BP traits and hypertension we performed the largest meta-analysis to date that included a total of ~350,000 individuals, directly genotyped with the Exome chip. The Exome chip contains ~240,000 mostly rare and low-frequency variants (Methods). A single-variant discovery analysis was performed, and candidate SNVs were taken forward for validation using independent replication samples. Gene-based tests were used to identify BP associated genes harboring multiple rare variant associations. We next assessed whether the newly identified BP associated SNVs were associated with expression levels of nearby genes, and tested these variants in aggregate for a causal association of BP with other cardiovascular traits and risk factors. Our findings highlight the contribution of rare variants in the aetiology of blood pressure in the general population, and provide new insights into the pathophysiology of hypertension.

Results

Discovery of single variant BP associations

336   We genotyped 192,763 individuals from 51 studies, and assessed association of 242,296 SNVs with

337   diastolic BP (DBP), systolic BP (SBP), pulse pressure (PP) and hypertension (HTN; Supplementary Tables

338   1, 2 and 3; Methods). An overview of the SNV discovery study design is given in Figure 1. A fixed effects

339   meta-analysis for each trait was performed using study-level association summary statistics from i) samples

340   of European (EUR) ancestry (up to 165,276 individuals), and ii) a trans-ethnic meta-analysis of the EUR and

341   additional South Asian (SAS) ancestry samples (EUR_SAS; up to 192,763 individuals). Two analyses of

342   DBP, SBP and PP were performed, one in which the trait was inverse normal transformed and a second in

343   which the raw phenotype was analysed. Both sets of results were consistent (Methods), therefore to

344   minimise sensitivity to deviations from normality in the analysis of rare variants, the results from the

345   analyses of the transformed traits were used for discovery. Strong correlations between the BP traits were

346   observed across studies (Methods), hence no adjustment of significance thresholds for independent trait

347   testing was applied.

348   The discovery meta-analyses identified 50 genomic regions with genome-wide significant (GWS) evidence

349   of association with at least one of the four BP traits tested ($P<5\mathrm{x}10^{-8}$; Supplementary Table 4). There were

350   45 regions associated in the EUR_SAS samples, of which 13 were novel (Figure 2). An additional five

351   regions were GWS in the EUR only meta-analyses of which two were novel (Supplementary Figure 1). In

352   total, 16 genomic regions were identified that were GWS for at least one BP trait that have not been

353   previously reported.

354   **Replication of single variant BP associations**

355   Next we sought support for our findings, in an independent replication dataset comprising of 18 studies, 16

356   of which were from the Cohorts for Heart and Aging Research in Genomic Epidemiology+ (CHARGE+)

357   exome chip blood pressure consortium (Figure 1; Liu *et al.* Nature Genetics, *submitted*). Variants were

358   selected for replication first using the larger (transformed) EUR_SAS data, with additional variants from the

359   (transformed) EUR data also selected. SNVs were selected if they mapped outside of known BP genomic

360   regions and had MAF≥0.05 and $P<1\mathrm{x}10^{-5}$ or MAF<0.05 and $P<1\mathrm{x}10^{-4}$ with at least one BP trait, *i.e.*

361   choosing a lower significance threshold for the selection of rare variants (full details of the selection criteria

362   are provided in the Methods). In total 81 candidate SNVs were selected for replication (Supplementary

Table 5). Eighty variants were selected from EUR_SAS (transformed) results and one SNV at the *ZNF101*

locus from the EUR (transformed) analyses. The results for EUR_SAS and EUR were consistent

(association statistics were correlated, ρ=0.9 across ancestries for each of the traits). Of the 81 variants, 30

SNVs were selected for association with DBP as the primary trait, 26 for SBP, 19 for PP and 6 for HTN,

with the primary trait defined as the BP trait with the smallest association *P*-value in the EUR-SAS

discovery analyses.

Meta-analyses were performed on results from analyses of untransformed DBP, SBP, PP and HTN (as only

results of untransformed traits were available from CHARGE+) in (i) up to 125,713 individuals of EUR

descent, and (ii) up to 155,063 individuals of multiple ethnicities (4,632 of Hispanic descent, 22,077 of

African American descent, 2,641 SAS samples with the remainder EUR; Figure 1). Given that a large

proportion of the ancestries in the trans-ethnic meta-analyses were not included in our discovery samples,

we used the EUR meta-analyses as the main data set for replication, but we also report any additional

associations identified within the larger trans-ethnic dataset.

Novel BP-SNV associations were identified based on two criteria (Figure 1; Methods). Firstly, replication of

the primary BP trait-SNV association was sought at a Bonferroni adjusted *P*-value threshold in the

replication data ($P \leq 6.17 \times 10^{-4}$, assuming α=0.05 for 81 SNVs tested and same direction of effect; Methods)

without the need for GWS. Secondly, meta-analyses of discovery and replication results across all four

(untransformed) BP traits were performed to assess the overall level of support across all samples for the 81

candidate SNVs; those BP-SNV associations that were GWS (with statistical support in the replication

studies; *P*<0.05 and the same direction of effect) were also declared as novel.

Seventeen SNV-BP associations formally replicated with concordant direction of effect at a Bonferroni

adjusted significance level for the primary trait. Fourteen were in the EUR meta-analyses, and amongst these

was a rare non-synonymous (ns) SNV mapping to *COL21A1* (Table 1, Supplementary Table 6). Three

associations were in the trans-ethnic meta-analyses, these included two rare nsSNVs in *RBM47* and *RRAS*

(Table 1, Supplementary Table 7; Methods).

389

390 In addition to the 17 SNV-BP trait associations that formally replicated, we identified 13 further SNV-

391 associations that were GWS in the combined (discovery and replication) meta-analyses. Ten of these were

392 GWS in the combined EUR analyses, (Table 2; Supplementary Tables 6 and 8a), and three were GWS in the

393 combined trans-ethnic meta-analyses (Table 2; Supplementary Tables 7 and 8b).

394

395 This gives a total of 30 novel SNV-BP associations (15 SNV-DBP, 9 SNV-SBP and 6 SNV-PP; Tables 1

396 and 2; Supplementary Figures 2 and 3). Five of the SNVs were GWS with more than one BP trait (Figure 3:

397 Tables 1 and 2; Supplementary Table 8). Four loci (*CERS5*, *TBX2*, *RGL3* and *OBFC1*) had GWS

398 associations with HTN in addition to GWS associations with DBP and SBP. The *PRKAG1* locus had GWS

399 associations with both SBP and PP.

400

401 Conditional analyses were performed to identify secondary signals of association within the novel BP loci.

402 The RAREMETALWORKER (RMW) package (Methods)[15] allows conditional analyses to be performed

403 using summary level data. Hence, analyses of the transformed primary traits and HTN were re-run in RMW

404 across the discovery studies (Figure 4). The results of the RMW single variant tests were consistent with the

405 initial discovery analyses (Supplementary Information). Given the RMW analyses were based on our

406 discovery samples, the larger EUR-SAS data was used as the main analysis to increase power, but we also

407 report any additional associations with evidence in EUR.

408

409 We identified secondary independent signals of association in four loci, *PREX1, PRKAG1* and *RRP1B*

410 within the EUR_SAS analyses and *COL21A1* in the EUR analyses ($P_{conditional}<1x10^{-4}$, Bonferroni adjusted

411 for ~500 variants within each region; Methods; Supplementary Tables 9 and 10). Three independent

412 association signals were identified in the *MYH6* locus in the EUR_SAS analyses (Supplementary Table 11).

413

414 **Gene-based BP associations**

415 To improve statistical power to detect associations in genes harbouring rare variants, analytical methods that

416 combine effects of variants across a gene into a single test have been devised and are implemented in the

417 RMW package[15]. We applied the gene-based sequence kernel association test (SKAT)[16] and Burden tests[17]

418 to the RMW dataset (MAF<0.05 or MAF<0.01; Figure 4; Methods). One previously unidentified BP gene

419 (*A2ML1*) was associated with HTN ($P= 7.73 \times 10^{-7}$) in the EUR_SAS studies and also in EUR studies

420 (Supplementary Table 12; Bonferroni-corrected threshold of significance $P<2.8 \times 10^{-6}$, after adjusting for

421 17,996 genes tested, Methods). The gene showed residual association with the primary BP trait after

422 conditioning on the most associated SNV in the gene ($P_{conditional}=5.00 \times 10^{-4}$; Supplementary Table 12),

423 suggesting that the association is due to multiple rare variants in the gene. One nonsense variant

424 (rs199651558, p.Arg893*, MAF=3.5×10^{-4}) was observed, and there were multiple missense variants (Figure

425 5). *A2ML1* encodes alpha-2-macroglobulin-like 1 protein, and is a member of the alpha macroglobulin

426 superfamily, which comprises protease inhibitors targeting a wide range of substrates. Mutations in this gene

427 are associated with a disorder clinically related to Noonan syndrome, a developmental disorder which

428 involves cardiac abnormalities[18]. We sought replication in the CHARGE+ studies for this gene, however

429 there was no evidence of association with HTN ($P= 0.45$). Given the very low frequencies of the variants

430 involved, however, studies in which the variants are polymorphic will be required to replicate the

431 association with HTN. The DBH gene was found to be associated with DBP using the SKAT test

432 ($P=2.88 \times 10^{-6}$). However, this was not due to multiple rare variants as the association was driven by

433 rs77273740 (Supplementary Table 5) and the SNV was not validated in the replication samples.

434

435 **Rare and common variant associations in established BP loci**

436 Of the 67 established BP loci, 35 loci were on the Exome chip (N=43 SNVs or close proxies $r^2>0.7$). All 43

437 SNVs had at least nominal evidence of association with BP in our discovery samples ($P<0.01$;

438 Supplementary Table 13). We also assessed if any of the established BP loci contained coding variants that

439 are associated with BP traits and in LD ($r^2>0.2$) with the known BP variants on the Exome chip

440 (Supplementary Table 13), using the 1000G phase 3 release for LD annotation. Focusing on SNVs that were

441 GWS for any BP trait from our transformed discovery data for either ancestry, there were 25 coding

442 variants, of which 6 were predicted to be damaging at loci labelled *CDC25A, SLC39A8, HFE, ULK4, ST7L-*

443 *CAPZA1-MOV10* and *CYP1A1-ULK3*. Three of these are published variants at loci labelled *SLC39A8, HFE*

444 and *ST7-CAPZA1-MOV10.* At *CYP1A1-ULK3*, the coding variant was in moderate LD with the reported

445   variant, but was less significantly associated with DBP in our EUR_SAS dataset (P=2.24x10^{-8} compared to

446   P=1.68x10^{-15} for the published variant). At the *ULK4* locus the predicted damaging coding variant had

447   similar association as the published coding variant (predicted to be benign), and prior work has already

448   indicated several associated nsSNVs in strong LD in *ULK4* [19]. The nsSNV within the *CDC25A* locus

449   (rs11718350 in *SPINK8*) had similar association with DBP as the intergenic published SNV in our

450   EUR_SAS dataset (P=2.00x10^{-8} compared to P=2.27x10^{-8} for the published variant). Overall at least 5 of

451   the known loci are consistent with having a coding causal variant.

452   Gene-based SKAT tests of all genes that map within 1 Mb of a previously reported SNV association

453   (Supplementary Table 14), indicated no genes with multiple rare or low-frequency variant associations.

454   Single variant conditional analyses showed that rs33966350, a rare nonsense variant in *ENPEP* (MAF=0.01)

455   was associated with SBP ($P_{conditional}$=1.61x10^{-5}) in the EUR_SAS samples (Supplementary Tables 14 and 15;

456   Methods) independently of the known SNV (rs6825911). *ENPEP* encodes aminopeptidase A (APA) an

457   enzyme of the renin-angiotensin-aldosterone system (RAAS) that converts angiotensin II (AngII) to AngIII.

458   There were no other established loci with convincing low-frequency or rare SNV associations in the

459   EUR_SAS samples. However, *HOXC4*, had evidence of a second independent signal with a rare missense

460   SNV in EUR samples (rs78731604; MAF=0.005, $P_{conditional}$= 5.76x10^{-5}; Supplementary Table 15). The

461   secondary signal in the *HOXC4* region, mapped to *CALCOCO1*, ~300kb from the known SNV. The gene

462   association (MAF≤0.01, P=2.37x10^{-5}) was below the required significance threshold and attributable to

463   rs78731604, which is not predicted to have detrimental effects on protein structure. Therefore, replication of

464   this association is required. Three loci (*ST7L-CAPZA1-MOV10, FIGN-GRB14,* and *TBX5-TBX3*) had

465   evidence of a second independent signal in the region in EUR_SAS samples with a common variant

466   ($P_{conditional}$<1x10^{-4}; Supplementary Table 15) that has not been previously reported.

467   Having identified 30 novel loci associated with BP traits, as well as additional new independent SNVs at

468   four novel loci and five known loci, we calculated the percent of the trait variance explained (Methods).

469   This was 2.08%/2.11%/1.15% for SBP/DBP/PP for the 43 previously reported BP-SNVs covered in our

470   dataset, increasing to 3.38%/3.41%/2.08% respectively with the inclusion of the 30 lead SNVs from novel

471   loci, plus new independent SNV-BP associations identified from novel and known loci.

472

Effect of BP SNVs on cardiovascular traits & risk factors

474 Amongst our novel BP-SNV associations, some have previously been reported to be associated with other

475 cardiovascular traits and risk factors (Supplementary Table 16); these include coronary heart disease (CHD:

476 *PHACTR1*, *ABO*)[20,21], QT interval (*RNF207*)[22], heart rate (*MYH6*)[23], and cholesterol levels (2q36.3, *ABO*,

477 *ZNF101*)[24].

478 To test the impact of BP variants on cardiovascular endpoints and risk factors we created three weighted

479 genetic risk scores (GRS) according to SBP/DBP/PP based on the newly identified and previously published

480 BP variants (up to N=125; Methods). The GRS models were used to test the causal effect of BP on the

481 following traits: ischemic stroke (including the subtypes, cardiometabolic, large and small vessel [25]), CHD,

482 heart failure,[26] left ventricular mass[27], left ventricular wall thickness[27], high-density lipoprotein cholesterol

483 (HDL-c), low-density lipoprotein (LDL-c), triglycerides, total cholesterol, body mass index (BMI), waist-

484 hip ratio adjusted BMI, height and estimated glomerular filtration rate (eGFR) (Methods). As expected, BP

485 was positively associated with increased CHD risk (OR [95% CI]=1.39[1.22-1.59] per 10mmHg increase in

486 SBP, $P=6.07\times10^{-7}$; 1.62[1.28-2.05] per 10mmHg increase in DBP, $P=5.99\text{x}10^{-5}$; 1.70[1.34-2.16] per

487 10mmHg increase in PP, $P=1.20\text{x}10^{-5}$; Table 3), and increased risk of ischemic stroke (OR [95%

488 CI]=1.93[1.47-2.55] per 10mmHg increase in DBP, $P=2.81\times10^{-6}$; 1.57[1.35-1.84] per 10mmHg increase in

489 SBP, $P=1.16\times10^{-8}$; 2.12[1.58-2.84] per 10mmHg increase in PP, $P=5.35\text{x}10^{-7}$). The positive association with

490 ischemic stroke was primarily due to large vessel stroke (Table 3). DBP and SBP were also positively

491 associated with left ventricular mass (9.57 [3.98-15.17] gram increase per 10mmHg increase in DBP,

492 $P=8.02\text{x}10^{-4}$ and 5.13 [1.77-8.48] gram increase per 10mmHg increase in SBP, $P=0.0027$) and left

493 ventricular wall thickness (0.10 [0.06-0.13] cm increase per 10mmHg increase in DBP, $P=1.88\text{x}10^{-8}$ and

494 0.05 [0.03-0.07] cm increase per 10mmHg increase in SBP, $P=5.52\text{x}10^{-6}$, Table 3). There was no convincing

495 evidence to support the BP associated variants having an effect on lipid levels ($P>0.1$), BMI ($P>0.005$),

496 waist hip ratio adjusted BMI ($P>0.1$), height ($P>0.06$), eGFR ($P>0.02$) or heart failure ($P>0.04$). The causal

497 associations with CHD, stroke, and left ventricular measures augment the results from a previous association

498 analysis using 29 BP variants[28]. Our data strongly support the previous observations of no causal

499 relationship between BP and eGFR. Lack of evidence of a BP effect with heart failure may only be due to

500 lack of power, as the association was in the expected direction.

501

502

Possible functional variants at BP loci and candidate genes

504 Twenty-six of our newly discovered BP associated SNVs had MAF≥0.05 and therefore due to extensive LD

505 with other SNVs not genotyped on the Exome array, identifying the causal genes requires additional

506 information. If a SNV is associated with increased or decreased expression of a particular gene, *i.e.* it is an

507 expression quantitative trait locus (eQTL) this suggests the gene on which the SNV acts could be in the

508 causal pathway. To help identify potential candidate causal genes in the novel BP loci (Supplementary Table

509 9), information from publicly available eQTL databases was investigated (MuTHER for LCL, adipose and

510 skin and GTEx for nine tissues including the heart and tibial artery; Methods).

511 The DBP increasing allele of the nsSNV, rs7302981-A, was associated with increased expression of *CERS5*

512 in: LCLs ($P_{MuTHER}=3.13 \times 10^{-72}$) skin ($P_{MuTHER}=2.40 \times 10^{-58}$) adipose ($P_{MuTHER}=2.87 \times 10^{-54}$) and nerve

513 ($P_{GTEx}=4.5 \times 10^{-12}$) (Supplementary Figure 4). Additional testing (Methods) provided no evidence against

514 colocalisation of the eQTL and DBP association signals, implicating *CERS5* as a candidate causal gene for

515 this DBP locus. *CERS5* (LAG1 homolog, ceramide synthase 5) is involved in the synthesis of ceramide, a

516 lipid molecule involved in several cellular signaling pathways. *Cers5* knockdown has been shown to reduce

517 cardiomyocyte hypertrophy in mouse models[29]. However, it is unclear whether the blood pressure raising

518 effects at this locus are the cause or result of any potential effects on cardiac hypertrophy. Future studies

519 investigating this locus in relation to parameters of cardiac hypertrophy and function (*e.g.* ventricular wall

520 thickness) should help address this question.

521 The DBP raising allele of the nsSNV (rs867186-A) was associated with increased expression of *PROCR* in

522 adipose tissue ($P_{MuTHER}=3.24 \times 10^{-15}$) and skin ($P_{MuTHER}=1.01 \times 10^{-11}$) (Supplementary Figure 4). There was no

523 evidence against colocalisation of the eQTL and DBP association thus supporting *PROCR* as a candidate

524 causal gene. *PROCR* encodes the Endothelial Protein C receptor, a serine protease involved in the blood

525 coagulation pathway, and rs867186 has previously been associated with coagulation and haematological

526 factors.[30,31] The PP decreasing allele of, rs10407022-T, which is predicted to have detrimental effects on

527 protein structure (Methods) was associated with increased expression of *AMH* in muscle ($P_{GTEx}=9.95\times10^{-15}$),

528 thyroid ($P_{GTEx}=8.54\times10^{-7}$), nerve ($P_{GTEx}=7.15\times10^{-8}$), tibial artery ($P_{GTEx}=6.46\times10^{-9}$), adipose

529 ($P_{GTEx}=4.69\times10^{-7}$), and skin ($P_{GTEx}=5.88\times10^{-8}$) (Supplementary Figure 4). There was no evidence against

530 colocalisation of the eQTL and PP association, which supports *AMH* as a candidate causal gene for PP. Low

531 AMH levels have been previously associated with hypertensive status in women with the protein acting as a

532 marker of ovarian reserve[32]. The intergenic SBP raising allele of rs4728142-A was associated with reduced

533 expression of *IRF5* in skin ($P_{MuTHER}=5.24\times10^{-31}$) and LCLs ($P_{MuTHER}=1.39\times10^{-34}$), whole blood

534 ($P_{GTEx}=3.12\times10^{-7}$) and tibial artery ($P_{GTEx}=1.71\times10^{-7}$).

535

536 Three novel rare nsSNVs were identified that map to *RBM47*, *RRAS* (both associated with SBP) and

537 *COL21A1* (associated with PP). They had larger effect sizes than common variant associations (>1.5mmHg

538 per allele; Supplementary Figure 5) and were predicted to have detrimental effects on protein structure

539 (Supplementary Table 16; Methods). In *RBM47*, rs35529250 (p.Gly538Arg) is located in a highly conserved

540 region of the gene and was most strongly associated with SBP (MAF=0.008; +1.59 mmHg per T allele;

541 $P=5.90\times10^{-9}$). *RBM47* encodes the RNA binding motif protein 47 and is responsible for post-transcriptional

542 regulation of RNA, through its direct and selective binding with the molecule.[33] In *RRAS*, rs61760904

543 (p.Asp133Asn) was most strongly associated with SBP (MAF=0.007; +1.51 mmHg per T allele; $P=8.45\times10^{-8}$

544). *RRAS* encodes a small GTPase belonging to the Ras subfamily of proteins H-RAS, N-RAS, and K-RAS

545 and has been implicated in actin cytoskeleton remodelling, and controlling cell proliferation, migration and

546 cycle processes[34]. The nsSNV in *COL21A1* (rs200999181, p.Gly665Val) was most strongly associated with

547 PP (MAF=0.001; +3.14 mmHg per A allele; $P=1.93\times10^{-9}$). *COL21A1* encodes the collagen alpha-1 chain

548 precursor of type XXI collagen, a member of the FACIT (fibril-associated collagens with an interrupted

549 triple helix) family of proteins[35]. The gene is detected in many tissues, including the heart and aorta. Based

550 on our results, these three genes represent good candidates for functional follow-up. However, due to the

551 incomplete coverage of all SNVs across the region on the Exome chip, it is possible that other non-

552 genotyped SNVs may better explain some of these associations. We therefore checked for variants in LD

553   (r^2>0.3) with these three rare nsSNVs in the UK10K + 1000G dataset[36] to ascertain if there are other

554   candidate SNVs at these loci (Supplementary Table 17). There were no SNVs within 1Mb of the *RBM47*

555   locus in LD with the BP associated SNV. At the *COL21A*1 locus there were only SNVs in moderate LD, and

556   these were annotated as intronic, intergenic or in the 5'UTR. At the *RRAS* locus, there were two SNVs in

557   strong LD with the BP associated SNV, which both mapped to introns of *SCAF1* and are not predicted to be

558   damaging. All SNVs in LD at both loci were rare as expected (Supplementary Table 17) supporting a role

559   for rare variants. Hence, the rare BP associated nsSNVs at *RBM47, COL21A1* an*d RRAS* remain the best

560   causal candidates.

561

562   **Pathway and network analyses**

563   To identify connected gene sets and pathways implicated by the BP associated genes we used Meta-Analysis

564   Gene-set Enrichment of variant Associations (MAGENTA)[37] and GeneGo MetaCore (Thomson Reuters,

565   UK). MAGENTA tests for over-representation of BP associated genes in pre-annotated pathways (gene

566   sets) (Methods and Supplementary Table 18a). GeneGo Metacore identifies potential gene networks. The

567   MAGENTA analysis was used for hypothesis generation and results were compared with the GeneGo

568   Metacore outputs to cross-validate findings.

569   Using MAGENTA there was an enrichment (P<0.01 and FDR<5% in either the EUR_SAS or the EUR

570   participants) of six gene sets with DBP, three gene sets with HTN and two gene sets for SBP

571   (Supplementary Table 18b). The RNA polymerase I promoter clearance (chromatin modification) pathway

572   showed the most evidence of enrichment with genes associated with DBP ($P_{Reactome}$=8.4x10^{-5}, FDR=2.48%).

573   NOTCH signalling was the most associated pathway with SBP ($P_{Reactome}$ = 3.00x10^{-4}, FDR = 5%) driven by

574   associations at the *FURIN* gene. The inorganic cation anion solute carrier (SLC) transporter pathway had

575   the most evidence of enrichment by HTN associated genes ($P_{Reactome}$=8.00x10^{-6}, FDR=2.13%).

576   Using GeneGo MetaCore, five network processes were enriched (FDR<5%; Methods; Supplementary

577   Tables 19 and 20). These included several networks with genes known to influence vascular tone and BP:

578   inflammation signalling, P=1.14x10^{-4} and blood vessel development P=2.34x10^{-4}. The transcription and

579   chromatin modification network (P=2.85x10^{-4}) was also enriched, a pathway that was also highlighted in the

580   MAGENTA analysis, with overlap of the same histone genes (*HIST1H4C*, *HIST1H2AC*, *HIST1H2BC*,

581   *HIST1H1T*) and has also been recently reported in an integrative network analysis of published BP loci and

582   whole blood expression profiling[38]. Two cardiac development pathways were enriched: the oxidative stress-

583   driven (ROS/NADPH) (P=4.12x10^{-4}) and the Wnt/β-catenin/integrin-driven (P=0.0010). Both these cardiac

584   development pathways include the *MYH6*, *MYH7*, and *TBX2* genes, revealing a potential overlap with

585   cardiomyopathies and hypertension, and suggesting some similarity in the underlying biological

586   mechanisms.

587

Discussion

589   By conducting the largest ever genetic study of BP, we identified further novel common variants with small

590   effects on BP traits, similar to what has been observed for obesity and height[39,40]. More importantly, our

591   study identified some of the first rare coding variants of strong effect (>1.5mmHg) that are robustly

592   associated with BP traits in the general population, complementing and extending the previous discovery

593   and characterisation of variants underlying rare Mendelian disorders of blood pressure regulation [41]. Using

594   SNV associations in 17 genes reported to be associated with monogenic disorders of blood pressure

595   (Methods) we found no convincing evidence of enrichment ($P_{enrichment}$=0.044). This suggests that BP control

596   in the general population may occur through different pathways to monogenic disorders of BP re-enforcing

597   the importance of our study findings. The identification of 30 novel BP loci plus further new independent

598   secondary signals within four novel and five known loci (Methods) has augmented the trait variance

599   explained by 1.3%, 1.2% and 0.93% for SBP, DBP and PP respectively within our data-set. This suggests

600   that with substantially larger sample sizes, for example through UK BioBank[42], we expect to identify 1000s

601   more loci associated with BP traits, and replicate more of our discovery SNV associations that are not yet

602   validated in the current report.

603   The discovery of rare missense variants has implicated several interesting candidate genes, which are often

604   difficult to identify from common variant GWAS, and should therefore lead to more rapidly actionable

605   biology. *A2ML1*, *COL21A1*, *RRAS* and *RBM47* all warrant further follow-up studies to define the role of

606 these genes in regulation of BP traits, as well as functional studies to understand their mechanisms of action.

607 *COL21A1* and *RRAS* warrant particular interest since both are involved in blood vessel remodelling, a

608 pathway of known aetiological relevance to hypertension.

609 We observed a rare nonsense SBP associated variant in *ENPEP* (rs33966350; p.Trp317*): this overlaps a

610 highly conserved region of both the gene and protein and is predicted to result in either a truncated protein

611 with reduced catalytic function or is subject to nonsense mediated RNA decay. ENPEP converts angiotensin

612 II (AngII) to Ang-III. AngII activates the angiotensin 1 (AT1) receptor resulting in vasoconstriction, while

613 AngIII activates the angiotensin 2 (AT2) receptor that promotes vasodilation and protects against

614 hypertension.[43] The predicted truncated protein may lead to predominant AngII signaling in the body, and

615 increases in BP. This new observation could potentially inform therapeutic strategies. Of note, angiotensin-

616 converting-enzyme (ACE) inhibitors are commonly used in the treatment of hypertension. However, patients

617 who suffer from adverse reactions to ACE inhibitors, such as dry cough and skin rash, would benefit from

618 alternative drugs that target RAAS. Murine studies have shown that in the brain, AngIII is the preferred AT1

619 agonist that promotes vasoconstriction and increases blood pressure, as opposed to AngII in the peripheral

620 system. These results have motivated the development of brain specific APA inhibitors to treat

621 hypertension[44]. Our results confirm APAs, such as ENPEP, as a valid target to modify blood pressure, but

622 suggest that long-term systemic reduction in APA activity may lead to an increase in blood pressure. Future

623 studies are needed to examine the effects of the p.Trp317* variant on the RAAS system, specifically in the

624 brain and peripheral vasculature, in order to test the benefits of the proposed therapeutic strategy in humans.

625 In addition to highlighting new genes in pathways of established relevance to BP and hypertension, and

626 identifying new pathways, we have also identified multiple signals at new loci. For example, there are three

627 distinct signals at the locus containing the *MYH6*/*MYH7* genes, and we note that *TBX2* maps to one of the

628 novel regions. These genes are related to cardiac development and/or cardiomyopathies, and provide an

629 insight into the shared inheritance of multiple complex traits. Unravelling the causal networks within these

630 polygenic pathways may provide opportunities for novel therapies to treat or prevent both hypertension and

631 cardiomyopathies.

632

URLs

Exome chip design information: http://genome.sph.umich.edu/wiki/Exome_Chip_Design

RareMetalWorker information: http://genome.sph.umich.edu/wiki/RAREMETALWORKER

Summary SNV association results: http://www.phenoscanner.medschl.cam.ac.uk

Databases used for variant annotation: http://www.ncbi.nlm.nih.gov/SNP/

http://www.ensembl.org/info/docs/tools/index.html and http://evs.gs.washington.edu/EVS/

UCSC reference file used for annotation of variants with gene and exon information:

http://hgdownload.soe.ucsc.edu/goldenPath/hg19/database/refFlat.txt.gz

Databases used for pathway analysis: MAGENTA (https://www.broadinstitute.org/mpg/magenta/) and

THOMSON REUTERS MetaCore™ Single Experiment Analysis workflow tool

(http://thomsonreuters.com/en/products-services/pharma-life-sciences/pharmaceutical-

research/metacore.html).

Naeem Khan, Sadiq Khan, Ayaz Ali, Madad Ali, Saeed Ahmed, Muhammad Waqar Khan, Muhammad Razaq Khan, Abdul Ghafoor, Mir Alam, Riazuddin, Muhammad Irshad Javed, Abdul Ghaffar, Tanveer Baig Mirza, Muhammad Shahid, Jabir Furqan, Muhammad Iqbal Abbasi, Tanveer Abbas, Rana Zulfiqar, Muhammad Wajid, Irfan Ali, Muhammad Ikhlaq, Danish Sheikh and Muhammad Imran.

2. ExomeBP Consortium

3. GoT2D and T2D-GENES consortia

936

983 **Author contributions**

984 Supervision and management of the project: JMHH and PBM. The following authors contributed to the

985 drafting of the manuscript: JMMH, PBM, PSu, HW, ASB, FD, JPC, DRB, KW, MT, FWA, LVW, NJS, JD

986 AKM, HY, CMM, NG, XS, TaT, DFF, MHs, OG, TF, VT. All authors critically reviewed and approved the

987 final version of the manuscript. **Statistical analysis review:** JMMH, PSu, FD, HW, JPC, RY, NM, PBM,

988 LVW, HY, TF, EMi, ADM, AM, AM, EE, ASB, FWA, MJC, CF, TF, SEH, ASH, JEH, JL, GM, JM, NM,

989 APM, APo, NJS, RAS, LS, KE, MT, VT, TVV, NV, KW, AMY, WZg, NG, CML, AKM, XS, TT. **Central**

990 **Data QC**: JMMH, ASB, PSu, RY, FD, HW, JPC, TF, LVW, PBM, EMi, NM, CML, NG, XS, AKM.

991 **Central Data analysis:** JMMH, PSu, FD, HW, JPC, NG, CML, AKM, XS. **Pathway analysis and**

Conflict of interests

030 **References**

031 1. Lim, S.S. *et al.* A comparative risk assessment of burden of disease and injury attributable to 67 risk
032 factors and risk factor clusters in 21 regions, 1990-2010: a systematic analysis for the Global Burden
033 of Disease Study 2010. *Lancet* **380**, 2224-60 (2012).
034 2. Rapsomaniki, E. *et al.* Blood pressure and incidence of twelve cardiovascular diseases: lifetime risks,
035 healthy life-years lost, and age-specific associations in 1.25 million people. *Lancet* **383**, 1899-911
036 (2014).
037 3. Munroe, P.B., Barnes, M.R. & Caulfield, M.J. Advances in blood pressure genomics. *Circ Res* **112**,
038 1365-79 (2013).
039 4. Ehret, G.B. *et al.* Genetic variants in novel pathways influence blood pressure and cardiovascular
040 disease risk. *Nature* **478**, 103-109 (2011).
041 5. Wain, L.V. *et al.* Genome-wide association study identifies six new loci influencing pulse pressure
042 and mean arterial pressure. *Nature Genetics* (2011).
043 6. Johnson, T. *et al.* Blood pressure loci identified with a gene-centric array. *Am J Hum Genet* **89**, 688-
044 700 (2011).
045 7. Tomaszewski, M. *et al.* Genetic architecture of ambulatory blood pressure in the general population:
046 insights from cardiovascular gene-centric array. *Hypertension* **56**, 1069-76 (2010).
047 8. Tragante, V. *et al.* Gene-centric meta-analysis in 87,736 individuals of European ancestry identifies
048 multiple blood-pressure-related loci. *Am J Hum Genet* **94**, 349-60 (2014).
049 9. Ganesh, S.K. *et al.* Loci influencing blood pressure identified using a cardiovascular gene-centric
050 array. *Hum Mol Genet* (2013).
051 10. Simino, J. *et al.* Gene-age interactions in blood pressure regulation: a large-scale investigation with
052 the CHARGE, Global BPgen, and ICBP Consortia. *Am J Hum Genet* **95**, 24-38 (2014).
053 11. Zhu, X. *et al.* Meta-analysis of Correlated Traits via Summary Statistics from GWASs with an
054 Application in Hypertension. *Am J Hum Genet* **96**, 21-36 (2015).

055   12.   Salfati, E., Morrison, A.C., Boerwinkle, E. & Chakravarti, A. Direct Estimates of the Genomic
056   Contributions to Blood Pressure Heritability within a Population-Based Cohort (ARIC). *PLoS One*
057   **10**, e0133031 (2015).

058   13.   Schork, N.J., Murray, S.S., Frazer, K.A. & Topol, E.J. Common vs. rare allele hypotheses for
059   complex diseases. *Curr Opin Genet Dev* **19**, 212-9 (2009).

060   14.   Nejentsev, S., Walker, N., Riches, D., Egholm, M. & Todd, J.A. Rare variants of IFIH1, a gene
061   implicated in antiviral responses, protect against type 1 diabetes. *Science* **324**, 387-9 (2009).

062   15.   Liu, D.J. *et al.* Meta-analysis of gene-level tests for rare variant association. *Nat Genet* **46**, 200-4
063   (2014).

064   16.   Wu, M.C. *et al.* Rare-variant association testing for sequencing data with the sequence kernel
065   association test. *Am J Hum Genet* **89**, 82-93 (2011).

066   17.   Li, B. & Leal, S.M. Methods for detecting associations with rare variants for common diseases:
067   application to analysis of sequence data. *Am J Hum Genet* **83**, 311-21 (2008).

068   18.   Vissers, L.E. *et al.* Heterozygous germline mutations in A2ML1 are associated with a disorder
069   clinically related to Noonan syndrome. *Eur J Hum Genet* **23**, 317-24 (2015).

070   19.   International Consortium for Blood Pressure Genome-Wide Association, S. *et al.* Genetic variants in
071   novel pathways influence blood pressure and cardiovascular disease risk. *Nature* **478**, 103-9 (2011).

072   20.   Coronary Artery Disease Genetics Consortium C4D. A genome-wide association study in Europeans
073   and South Asians identifies five new loci for coronary artery disease. *Nat Genet* **43**, 339-44 (2011).

074   21.   Schunkert, H. *et al.* Large-scale association analysis identifies 13 new susceptibility loci for coronary
075   artery disease. *Nat Genet* **43**, 333-8 (2011).

076   22.   Arking, D.E. *et al.* Genetic association study of QT interval highlights role for calcium signaling
077   pathways in myocardial repolarization. *Nat Genet* **46**, 826-36 (2014).

078   23.   den Hoed, M. *et al.* Identification of heart rate-associated loci and their effects on cardiac conduction
079   and rhythm disorders. *Nat Genet* **45**, 621-31 (2013).

080   24.   Global Lipids Genetics Consortium *et al.* Discovery and refinement of loci associated with lipid
081   levels. *Nat Genet* **45**, 1274-83 (2013).

082   25.   Traylor, M. *et al.* Genetic risk factors for ischaemic stroke and its subtypes (the METASTROKE
083   collaboration): a meta-analysis of genome-wide association studies. *Lancet Neurol* **11**, 951-62
084   (2012).

085   26.   Smith, N.L. *et al.* Association of genome-wide variation with the risk of incident heart failure in
086   adults of European and African ancestry: a prospective meta-analysis from the cohorts for heart and
087   aging research in genomic epidemiology (CHARGE) consortium. *Circ Cardiovasc Genet* **3**, 256-66
088   (2010).

089   27.   Vasan, R.S. *et al.* Genetic variants associated with cardiac structure and function: a meta-analysis
090   and replication of genome-wide association data. *JAMA* **302**, 168-78 (2009).

091   28.   Ehret, G.B. *et al.* Genetic variants in novel pathways influence blood pressure and cardiovascular
092   disease risk. *Nature* **478**, 103-109 (2011).

093   29.   Russo, S.B. *et al.* Ceramide synthase 5 mediates lipid-induced autophagy and hypertrophy in
094   cardiomyocytes. *J Clin Invest* **122**, 3919-30 (2012).

095   30.   Oudot-Mellakh, T. *et al.* Genome wide association study for plasma levels of natural anticoagulant
096   inhibitors and protein C anticoagulant pathway: the MARTHA project. *Br J Haematol* **157**, 230-9
097   (2012).

098   31.   Smith, N.L. *et al.* Novel associations of multiple genetic loci with plasma levels of factor VII, factor
099   VIII, and von Willebrand factor: The CHARGE (Cohorts for Heart and Aging Research in Genome
100   Epidemiology) Consortium. *Circulation* **121**, 1382-92 (2010).

101   32.   Bleil, M.E., Gregorich, S.E., McConnell, D., Rosen, M.P. & Cedars, M.I. Does accelerated
102   reproductive aging underlie premenopausal risk for cardiovascular disease? *Menopause* **20**, 1139-46
103   (2013).

104   33.   Guan, R. *et al.* rbm47, a novel RNA binding protein, regulates zebrafish head development. *Dev Dyn*
105   **242**, 1395-404 (2013).

106   34.   Wozniak, M.A., Kwong, L., Chodniewicz, D., Klemke, R.L. & Keely, P.J. R-Ras controls membrane
107   protrusion and cell migration through the spatial regulation of Rac and Rho. *Mol Biol Cell* **16**, 84-96
108   (2005).

109   35.   Tuckwell, D. Identification and analysis of collagen alpha 1(XXI), a novel member of the FACIT
110   collagen family. *Matrix Biol* **21**, 63-6 (2002).
111   36.   Huang, J. *et al.* Improved imputation of low-frequency and rare variants using the UK10K haplotype
112   reference panel. *Nat Commun* **6**, 8111 (2015).
113   37.   Segre, A.V. *et al.* Common inherited variation in mitochondrial genes is not enriched for associations
114   with type 2 diabetes or related glycemic traits. *PLoS Genet* **6**(2010).
115   38.   Huan, T. *et al.* Integrative network analysis reveals molecular mechanisms of blood pressure
116   regulation. *Mol Syst Biol* **11**, 799 (2015).
117   39.   Locke, A.E. *et al.* Genetic studies of body mass index yield new insights for obesity biology. *Nature*
118   **518**, 197-206 (2015).
119   40.   Wood, A.R. *et al.* Defining the role of common variation in the genomic and biological architecture
120   of adult human height. *Nat Genet* **46**, 1173-86 (2014).
121   41.   Park, H.W. *et al.* Serine-threonine kinase with-no-lysine 4 (WNK4) controls blood pressure via
122   transient receptor potential canonical 3 (TRPC3) in the vasculature. *Proc Natl Acad Sci U S A* **108**,
123   10750-5 (2011).
124   42.   Sudlow, C. *et al.* UK biobank: an open access resource for identifying the causes of a wide range of
125   complex diseases of middle and old age. *PLoS Med* **12**, e1001779 (2015).
126   43.   Te Riet, L., van Esch, J.H., Roks, A.J., van den Meiracker, A.H. & Danser, A.H. Hypertension:
127   Renin-Angiotensin-aldosterone system alterations. *Circ Res* **116**, 960-75 (2015).
128   44.   Gao, J. *et al.* A new strategy for treating hypertension by blocking the activity of the brain renin-
129   angiotensin system with aminopeptidase A inhibitors. *Clin Sci (Lond)* **127**, 135-48 (2014).

130

131

132

133

134

135

136

Figure Legends

Figure 1 Study design and work flow diagram of single variant discovery analyses. EUR=European, SAS=South

Asian, HIS=Hispanic, AA=African American, HTN=hypertension, BP=blood pressure, SBP=systolic blood pressure,

DBP= diastolic blood pressure, PP=pulse pressure, N=sample size, MAF=minor allele frequency, *P*=*P*-value

significance threshold, SNV=single-nucleotide variant, GWS=genome-wide significance *Further details of the selection

criteria are provided in the methods.

Figure 2 Discovery SNV-BP associations. Results are provided for (a) transformed SBP (b) transformed DBP (c)

transformed PP and (d) HTN in the European and South Asian (EUR_SAS) discovery samples. The y-axis represents –

$\log_{10}P$ for association. Red triangles represent variants that map to one of the 81 regions selected for replication, blue

triangles represent SNVs that map to previously published BP regions, and grey triangles represent all remaining SNVs.

SNVs are ordered according to chromosome (black lines on the outside of the plot) and physical position. Genes that

SNVs map to are given in the outer blocks.

Figure 3 Overlap of the 30 novel loci associations across SBP, DBP, PP and HTN. The Venn diagram shows which

of the 30 newly identified BP loci are associated with multiple BP traits. Only SNV-BP trait associations that were

genome-wide significant ($P < 5\times10^{-8}$) in the combined discovery and replication meta-analyses are listed for any given

BP trait, within the corresponding ancestry dataset that the given locus was validated for (see Tables 1 and 2). The

association of *RRAS* variant with SBP was replicated in the independent samples, but did not achieve GWS in the

combined discovery and replication meta-analysis and is therefore only included for SBP. HTN=hypertension,

SBP=systolic blood pressure, DBP= diastolic blood pressure, PP=pulse pressure.

Figure 4 Study design for conditional analyses and rare variant gene-based discovery analyses.

RMW=RareMetalWorker, EUR=European, SAS = South Asian, HTN=hypertension, BP=blood pressure, SBP=systolic

blood pressure, DBP= diastolic blood pressure, PP=pulse pressure. N=sample size, MAF=minor allele frequency, *P*=*P*-

value significance threshold, P_{cond}=conditional *P*-value significance threshold

Figure 5 Locus plot for *A2ML1* and secondary amino acid structure of the gene product. (a) Locus plot for *A2ML1*

associated with HTN identified through gene based tests. The variants' positions along the gene (x axis, based on

163   human genome build 37) and the $-\log_{10}(P$-value of association) (y axis) are indicated. The variants are colour coded:

164   nonsense (black), missense, predicted damaging (blue), and missense (orange). The schematic above the x-axis

165   represents the intron / exon (black vertical bars) structure, the untranslated regions are shown as grey vertical bars.

166   (b) The white box denotes the full-length amino acid sequence for each of the two gene products. Black numbers

167   denote amino acid residue positions of note. Coloured boxes depict putative functional domains (see below). Coloured

168   vertical lines indicate the amino acid substitutions corresponding to the variants depicted in the locus plots above using

169   the same colour coding. Bold, italic indicates the SNV association with smallest P-value.

170   Dark grey – signal peptide sequence. Brown – regions of intramolecular disulfide bonds. For simplicity only those

171   regions coinciding with variants described were indicated. Black – bait region described to interact with proteases.

172   Purple – thiol ester sequence region aiding in interaction with proteases. Light grey – alpha helical regions thought to

173   mediate A2ML1 interaction with LRP1, facilitating receptor-mediated endocytosis.

174

175

176

Table 1 Novel blood pressure trait associated loci. Variants with formal replication

	Variant information			Discovery		Replication			Combined		
Locus	rsID	Chr:Pos (EA:EAF)	Trait	P_t	P_u	N	β	P	N	β	P
EUR											
RNF207	rs709209	1:6.28 (A:0.655)	PP	4.57×10^{-6}	1.60×10^{-6}	122,780	0.17	5.83×10^{-4}	284,683	0.20	9.62×10^{-9}
C5orf56	rs12521868	5:131.78 (T:0.373)	DBP	1.59×10^{-6}	3.03×10^{-7}	122,795	-0.18	2.29×10^{-5}	282,023	-0.19	6.12×10^{-11}
PHACTR1	rs9349379	6:12.90 (A:0.566)	SBP	2.11×10^{-8}	1.78×10^{-7}	122,809	0.24	4.06×10^{-4}	284,673	0.29	8.84×10^{-10}
COL21A1	rs200999181†	6:55.94 (A:0.002)	PP	3.08×10^{-8}	2.46×10^{-7}	121,487	2.70	1.90×10^{-4}	242,486	3.25	6.27×10^{-10}
ABO	rs687621	9:136.14 (A:0.615)	DBP	8.80×10^{-8}	2.55×10^{-7}	122,798	0.16	1.96×10^{-4}	276,014	0.19	5.45×10^{-10}
ADO	rs10995311	10:64.56 (C:0.567)	DBP	1.86×10^{-6}	1.14×10^{-6}	122,798	0.23	8.47×10^{-8}	266,456	0.21	1.12×10^{-12}
LMO1	rs110419	11:8.25 (A:0.48)	DBP	9.41×10^{-6}	2.22×10^{-5}	122,798	0.16	1.81×10^{-4}	279,935	0.16	3.04×10^{-8}
OR5B12	rs11229457	11:58.21 (T:0.236)	SBP	1.58×10^{-6}	4.62×10^{-5}	122,809	-0.32	7.53×10^{-5}	284,680	-0.31	2.70×10^{-8}
CERS5	rs7302981	12:50.54 (A:0.361)	DBP	1.35×10^{-13}	4.60×10^{-11}	122,798	0.24	2.64×10^{-8}	284,718	0.25	1.38×10^{-17}
MYH6	rs452036	14:23.87 (A:0.327)	PP	4.59×10^{-11}	2.80×10^{-13}	122,780	-0.21	1.81×10^{-5}	284,672	-0.28	2.96×10^{-16}
DPEP1	rs1126464	16:89.70 (C:0.256)	DBP	1.19×10^{-9}	4.35×10^{-11}	118,677	0.24	1.68×10^{-6}	261,564	0.28	1.02×10^{-15}
TBX2	rs8068318†	17:59.48 (T:0.698)	DBP	7.46×10^{-13}	5.71×10^{-10}	122,798	0.26	3.23×10^{-8}	281,978	0.26	1.95×10^{-16}
RGL3	rs167479	19:11.53 (T:0.486)	DBP	2.22×10^{-23}	1.97×10^{-22}	122,797	-0.29	3.01×10^{-11}	283,332	-0.33	1.99×10^{-31}
PREX1	rs6095241	20:47.31 (A:0.452)	DBP	5.65×10^{-6}	2.29×10^{-5}	122,798	-0.18	2.56×10^{-5}	281,322	-0.17	4.75×10^{-9}
ALL ancestry											
RBM47	rs35529250†	4:40.43 (T:0.01)	SBP	6.56×10^{-7}	6.15×10^{-6}	148,878	-1.43	5.02×10^{-4}	306,352	-1.55	2.42×10^{-8}
OBFC1	rs4387287	10:105.68 (A:0.157)	SBP	2.23×10^{-8}	1.32×10^{-7}	147,791	0.28	3.37×10^{-4}	320,494	0.36	9.12×10^{-10}
RRAS	rs61760904†	19:50.14 (T:0.008)	SBP	1.96×10^{-6}	1.90×10^{-5}	148,878	1.38	5.70×10^{-4}	322,664	1.50	8.45×10^{-8}

178 SNV-BP associations are reported for the newly identified BP loci that replicated at $P < 6.2 \times 10^{-4}$ (Bonferroni correction for the 81 variants selected for replication for a primary
179 blood pressure trait; Methods). Loci are categorised into EUR and ALL ancestry based on the meta-analysis used to replicate the variants for the primary BP trait shown in columns
180 labelled 'Trait'. In the columns that contains the discovery meta-analyses results, P_t represents the P-value for association of the variant with the transformed primary BP trait in the
181 EUR_SAS discovery meta-analyses (which was also used to select the variant for replication) and P_u represents the P-value for association with the untransformed primary BP trait
182 in the ancestry in which the variant replicated. N, β and P, which denote the number of samples, estimated allelic effect and P-value for association with the primary BP trait, are
183 provided for the untransformed primary BP trait in the replication data and also from the combined (discovery and replication) meta-analyses. NB: ALL ancestry corresponds to all
184 ancestries in the combined (discovery + replication) meta-analyses
185 Locus – Gene or region containing the SNV, rsID - dbSNP rsID. Chr:Pos (EA:EAF) – Chromosome:NCBI Build 37 position in Mb (effect allele:effect allele frequency), Trait – primary
186 blood pressure trait for which the variant was and also replicated, β - effect estimate, N:sample size, EUR - European.
187 † indicates it is a non-synonymous SNV (nsSNV) or is in linkage disequilibrium with a nsSNV ($r^2 > 0.8$) that is predicted to be damaging

188 Table 2 Novel blood pressure trait associated loci. Variants with GWS evidence of association in combined
189 meta-analyses

190

	Variant information			Discovery		Replication			Combined		
Locus	rsID	Chr:Pos (EA:EAF)	Trait	P_t	P_U	N	β	P	N	β	P
EUR											
2q36.3	rs2972146	2:227.10 (T:0.652)	DBP§ (HTN)	1.51×10^{-9}	2.47×10^{-7}	122,798	0.13	2.20×10^{-3}	275,610	0.17	8.40×10^{-9}
ZBTB38	rs16851397	3:141.13 (A:0.953)	DBP§ (SBP)	6.87×10^{-6}	3.20×10^{-5}	122,798	-0.38	1.20×10^{-4}	284,717	-0.38	3.01×10^{-8}
PRDM6	rs1008058	5:122.44 (A:0.135)	SBP	5.09×10^{-7}	1.01×10^{-8}	43,109	0.46	3.61×10^{-3}	176,362	0.55	2.99×10^{-10}
GPR20	rs34591516	8:142.37 (T:0.055)	SBP§ (DBP)	1.54×10^{-6}	1.01×10^{-7}	122,807	0.51	4.20×10^{-4}	282,009	0.64	6.10×10^{-10}
HOXB7	rs7406910	17:46.69 (T:0.118)	SBP	6.07×10^{-10}	2.74×10^{-9}	122,809	-0.20	4.89×10^{-2}	284,690	-0.46	3.80×10^{-8}
AMH	rs10407022†	19:2.25 (T:0.82)	PP	1.63×10^{-7}	1.73×10^{-7}	118,656	-0.19	1.62×10^{-3}	252,525	-0.26	5.94×10^{-9}
ZNF101	rs2304130	19:19.79 (A:0.914)	DBP	1.66×10^{-8}	1.92×10^{-8}	122,798	-0.17	1.71×10^{-2}	284,705	-0.29	1.53×10^{-8}
PROCR	rs867186	20:33.76 (A:0.873)	DBP	1.44×10^{-6}	4.15×10^{-7}	122,798	0.21	2.48×10^{-3}	284,722	0.26	1.19×10^{-8}
RRP1B	rs9306160	21:45.11 (T:0.374)	DBP§ (SBP)	1.04×10^{-8}	1.90×10^{-6}	100,489	-0.16	4.30×10^{-4}	249,817	-0.18	6.80×10^{-9}
TNRC6B	rs470113	22:40.73 (A:0.804)	PP	1.48×10^{-10}	1.31×10^{-9}	122,780	-0.14	1.37×10^{-2}	284,683	-0.25	1.67×10^{-9}
ALL ancestry											
7q32.1	rs4728142	7:128.57 (A:0.433)	SBP	8.10×10^{-6}	4.21×10^{-6}	150,542	-0.21	8.62×10^{-4}	338,338	-0.24	3.45×10^{-8}
PRKAG1	rs1126930†	12:49.40 (C:0.036)	PP	2.12×10^{-6}	4.62×10^{-7}	151,481	0.36	3.74×10^{-3}	314,894	0.50	3.34×10^{-8}
SBNO1	rs1060105	12:123.81 (T:0.209)	DBP	6.66×10^{-7}	1.09×10^{-6}	150,532	-0.15	2.67×10^{-3}	336,413	-0.18	3.07×10^{-8}

191
192 SNV-BP associations are reported for the newly identified BP loci that showed genome-wide significant association ($P < 5 \times 10^{-8}$) in the combined discovery and replication meta-
193 analyses. In the columns that contain results from the discovery meta-analyses, P_t represents the P-value for association of the variant with the transformed *primary* BP trait in the
194 EUR_SAS discovery meta-analyses (used to select the variant for replication) and P_u represents the P-value for association with the untransformed BP trait in the ancestry in which
195 the variant was validated. Loci are categorised into EUR and ALL ancestry based on the ancestry in which the variant showed association with a blood pressure trait at $P < 5 \times 10^{-8}$.
196 N, β and P, which denote the number of samples, estimated allelic effect and P-value for association with the validated BP trait, are provided for the untransformed BP trait in the
197 replication data and also from the combined (discovery and replication) meta-analyses. NB: ALL ancestry corresponds to all ancestries in the combined (discovery + replication)
198 meta-analyses.
199
200 Locus – Gene or region containing the SNV, rsID - dbSNP rsID. Chr:Pos (EA:EAF) – Chromosome:NCBI Build 37 position in Mb (effect allele:effect allele frequency), Trait - blood
201 pressure trait for which association is reported, EUR - European.
202 § At four loci (2q36.3, *ZBTB38, GPR20* and *RRP1B*) the primary trait used to select the variants for replication is given in parentheses because the variant associations were
203 validated in the combined meta-analysis for the listed secondary trait. For these variants, P_t denotes the P-value for association with the primary trait, the other P-values provided
204 are for the secondary trait.
205 † indicates it is a non-synonymous SNV (nsSNV) or is linkage disequilibrium with a nsSNV ($r^2 > 0.8$) that is predicted to be damaging

206 **Table 3 Results of the genetic risk score analyses across CVD traits and risk factors.**

Outcome	Units	N	DBP (per 10mmHg increase)		SBP (per 10mmHg increase)		PP (per 10mmHg increase)	
			Effect [95% CI]	P	Effect [95% CI]	P	Effect [95% CI]	P
CHD	OR	82,056	1.62 [1.28, 2.05]	5.99×10^{-5}	1.39 [1.22, 1.59]	6.07×10^{-7}	1.70 [1.34, 2.16]	1.20×10^{-5}
Ischemic stroke	OR	25,799	1.93 [1.47, 2.55]	2.81×10^{-6}	1.57 [1.35, 1.84]	1.16×10^{-8}	2.12 [1.58, 2.84]	5.35×10^{-7}
Cardioembolic stroke	OR	16,113	1.43 [0.86, 2.39]	0.1683	1.33 [0.99, 1.80]	0.0584	1.73 [1.00, 3.02]	0.0518
Large vessel stroke	OR	13,903	2.26 [1.25, 4.08]	0.0068	1.85 [1.32, 2.59]	3.61×10^{-4}	3.05 [1.64, 5.68]	4.37×10^{-4}
Small vessel stroke	OR	15,617	1.96 [1.13, 3.41]	0.0168	1.56 [1.13, 2.16]	0.0064	1.98 [1.09, 3.61]	0.0248
Heart failure	OR	13,282	1.48 [1.02, 2.17]	0.0409	1.25 [1.00, 1.57]	0.0512	1.33 [0.88, 2.02]	0.1757
Left ventricular mass	g	11,273	9.57 [3.98,15.17]	8.02×10^{-4}	5.13 [1.77, 8.48]	0.0027	5.97 [-0.38,12.31]	0.0653
Left ventricular wall thickness	cm	11,311	0.10 [0.06, 0.13]	1.88×10^{-8}	0.05 [0.03, 0.07]	5.52×10^{-6}	0.05 [0.01, 0.09]	0.0187
HDL	mg/dl	80,395	0.25 [-1.00, 1.51]	0.6930	0.21 [-0.50, 0.92]	0.5622	0.47 [-0.79, 1.73]	0.4668
LDL	mg/dl	77,021	-1.57 [-5.20, 2.06]	0.3972	0.07 [-2.03, 2.16]	0.9498	1.87 [-1.86, 5.59]	0.3255
Total cholesterol	mg/dl	80,455	-1.34 [-5.90, 3.22]	0.5639	0.70 [-1.93, 3.32]	0.6029	3.68 [-0.97, 8.33]	0.1209
Triglycerides	mg/dl	77,779	0.02 [-0.03, 0.08]	0.3859	0.02 [-0.01, 0.05]	0.2697	0.03 [-0.03, 0.08]	0.3025
BMI	INVT	526,508	-0.10 [-0.18,-0.01]	0.0342	-0.07 [-0.13,-0.02]	0.0058	-0.12 [-0.23,-0.02]	0.0165
WHRadjBMI	INVT	344,369	0.03 [-0.04, 0.11]	0.4025	0.03 [-0.02, 0.08]	0.2170	0.06 [-0.03, 0.15]	0.1885
Height	INVT	458,927	0.02 [-0.15, 0.18]	0.8592	-0.04 [-0.15, 0.06]	0.4170	-0.18 [-0.37, 0.01]	0.0683
eGFR	INVT	51,039	-0.02 [-0.15, 0.11]	0.7810	-0.03 [-0.10, 0.04]	0.4080	-0.07 [-0.20, 0.06]	0.2741

207 CHD, coronary heart disease; HDL, high density lipoprotein; LDL, low density lipoprotein; eGFR, estimated glomerular filtration rate; DBP, diastolic blood pressure; SBP systolic blood pressure; PP, pulse pressure; OR, odds ratio; g, grams; INVT, inverse normally

208 transformed (hence no units); N, sample size; P, P-value of association of BP with the trait listed; CI, confidence interval. Results are considered significant if P < 0.0038, which corresponds to a Bonferroni correction for 13 phenotypes tested.

209

210

211 **Online Methods**

212 **Overview of discovery studies**

213 The cohorts contributing to the discovery meta-analyses comprise studies from three consortia (CHD

214 Exome+, ExomeBP, and GoT2D/T2D-GENES) with a total number of 192,763 unique samples. All

215 participants provided written informed consent and the studies were approved by their local Research Ethics

216 Committees and/or Institutional Review Boards.

217 The CHD Exome+ consortium comprised 77,385 samples: eight studies (49,898 samples) of European

218 (EUR) ancestry, two studies (27,487 samples) of South Asian (SAS) ancestry (Supplementary Table 1). The

219 ExomeBP consortium included 25 studies (75,620 samples) of EUR ancestry (Supplementary Table 1). The

220 GoT2D consortium comprised 14 studies (39,758 samples) of Northern EUR ancestry from Denmark,

221 Finland, and Sweden (Supplementary Table 1). The participating studies and their characteristics including

222 BP phenotypes are detailed in Supplementary Tables 1 and 2. Note, any studies contributing to multiple

223 consortia were only included once in all meta-analyses.

224 **Phenotypes**

225 Four blood pressure (BP) traits were analysed: systolic blood pressure (SBP), diastolic blood pressure

226 (DBP), pulse pressure (PP) and hypertension (HTN). For individuals known to be taking BP lowering

227 medication, 15/10 mmHg was added to the raw SBP/DBP values, respectively, to obtain medication-

228 adjusted SBP/DBP values[45]. PP was defined as SBP minus DBP, post-adjustment. For HTN, individuals

229 were classified as hypertensive cases if they satisfied at least one of: (i) SBP≥140 mmHg, (ii) DBP≥90

230 mmHg, (iii) taking anti-hypertensive or BP lowering medication. All other individuals were included as

231 controls. The four BP traits were correlated (SBP:DBP correlations were between 0.6 and 0.8, and SBP:PP

232 correlations were ~0.8). However, they measure partly distinct physiological features including, cardiac

233 output, vascular resistance, and arterial stiffness, all measures for determining a cardiovascular risk profile.

234 Therefore the genetic architecture of the individual phenotypes are of interest, and a multi-phenotype

235 mapping approach was not adopted.

236

Genotyping

All samples were genotyped using one of the Illumina HumanExome Beadchip arrays (Supplementary Table 3). An Exome chip quality control Standard Operating Procedure (SOP) developed by Anubha Mahajan, Neil Robertson and Will Rayner at the Wellcome Trust Centre for Human Genetics, University of Oxford was used by most studies for genotype calling and QC[46] (Supplementary Table 3). All genotypes were aligned to the plus strand of the human genome reference sequence (Build37) prior to any analyses and any unresolved mappings were removed. Genotype cluster plots were reviewed for all the novel rare variants (both lead and secondary signals) and for rare variants that contributed to the gene-based testing.

Meta-analyses

Meta-analyses were performed using METAL[47], for both discovery and replication analyses, using inverse variance weighted fixed effect meta-analysis for the continuous traits (SBP, DBP and PP) and sample size weighted meta-analysis for the binary trait (HTN).

Discovery SNV analyses

Analyses of both untransformed and inverse normal transformed SBP, DBP and PP were conducted within each contributing study. The analyses of the transformed traits were performed in order to minimise sensitivity to deviations from normality in the analysis of rare variants and for discovery of new SNV-BP associations. The residuals from the null model obtained after regressing the medication-adjusted trait on the covariates (age, age^2, sex, BMI, and disease status for CHD) within a linear regression model, were ranked and inverse normalised. These normalised residuals were used to test trait-SNV associations. All SNVs that passed QC were analysed for association, without any further filtering by MAF, but a minor allele count of 10 was used for the analysis of HTN. An additive allelic effects model was assumed.

Two meta-analyses were performed for each trait, one with EUR and SAS ancestries combined (EUR_SAS) and another for EUR ancestry alone. Contributing studies used principal components (PCs) to adjust for population stratification. Consequently minimal inflation in the association test statistics, λ, was observed (

262   λ =1.07 for SBP, 1.10 for DBP, 1.04 for PP and <1 for HTN in the transformed discovery meta-analysis in

263   EUR_SAS; λ = 1.06 for SBP, 1.09 for DBP, 1.05 for PP and <1 for HTN in the transformed discovery

264   meta-analysis in EUR; Supplementary Figure 6). The meta-analyses were performed independently in two

265   centres and results were found to be concordant between centres. Given the studies contributing to the

266   discovery analyses were ascertained on CHD or T2D, we tested potential systematic bias in calculated effect

267   estimates amongst these studies. No evidence of bias in the overall effect estimates was obtained.

268   The results for the transformed traits were taken forward and used to select candidate SNVs for replication.

269   Results (*P*-values) from the transformed and untransformed analyses were strongly correlated (r^2>0.9).

270   **Replication SNV analyses**

271   SNVs associated with any of the transformed traits (SBP, DBP, PP) or HTN were annotated using the

272   Illumina SNV annotation file, humanexome-12v1_a_gene_annotation.txt, independently across two centres.

273   Given the difference in power to detect common versus low frequency and rare variant associations, two

274   different significance thresholds were chosen for SNV selection. For SNVs with MAF≥0.05, $P{\le}1{\times}10^{-5}$ was

275   selected, while, $P{\le}1{\times}10^{-4}$ was used for SNVs with MAF < 0.05. By choosing a significance threshold of

276     $P{<}1{\times}10^{-4}$ we maximized the opportunity to follow-up rare variants (making the assumption that any true

277   signals at this threshold could replicate at Bonferroni adjusted significance, $P{\le}6.17{\times}10^{-4}$, assuming α=0.05

278   for 81 SNVs). All previously published BP associated SNVs and any variants in LD with them (r^2>0.2),

279   were removed from the list of associated SNVs as we aimed to replicate new findings only. SNVs for which

280   only one study contributed to the association result or showed evidence of heterogeneity (P_{het}<0.0001) were

281   removed from the list as they were likely to be an artefact. Where SNVs were associated with multiple traits,

282   to minimise the number of tests performed, only the trait with the smallest *P*-value was selected as the

283   primary trait in which replication was sought. Where multiple SNVs fitted these selection criteria for a

284   single region, only the SNV with the smallest *P*-value was selected. In total, 81 SNVs were selected for

285   validation in independent samples. These 81 SNVs had concordant association results for both transformed

286   and non-transformed traits. Eighty SNVs were selected from EUR_SAS results (with consistent support in

287   EUR), and one SNV from EUR results only. In the next step, we looked up the 81 SNV-BP associations

288   using data from a separate consortium, the CHARGE+ exome chip blood pressure consortium (who had

289 analysed untransformed SBP, DBP, PP and HTN), and UHP and Lolipop (ExomeBP consortium;

290 Supplementary Tables 2 and 3). The analysed residuals from CHARGE+ were approximately normally

291 distributed in their largest studies (Supplementary Figure 7).

292 Two meta-analyses of the replication datasets were performed: one of EUR samples, and a second of EUR,

293 African American, Hispanics and SAS ancestries ("ALL"). Replication was confirmed if P (1-tailed) <

294 $0.05/81=6.17x10^{-4}$ and the effect (beta) was in the direction observed in discovery meta-analyses for the

295 selected trait. A combined meta-analysis was performed of discovery (untransformed results as only

296 untransformed data was available from CHARGE+ exome chip blood pressure consortium) and replication

297 results across the four traits to assess the overall support for each locus. For the combined meta-analyses, a

298 GWS threshold of, $P\leq5x10^{-8}$, was used to declare a SNV as novel rather than a less stringent experiment

299 wide threshold, as GWS is used to declare significance in GWAS and we wish to minimise the possibility of

300 false positive associations. (Note that GWS is equivalent to an exome-wide threshold of $P\leq2x10^{-7}$ adjusted

301 for four traits).

302

303 Note: all validated BP-associated variants were associated at $P<10^{-5}$ in the discovery dataset (for the primary

304 trait). Hence, we could have used the same inclusion criteria for both common and rare SNVs. Therefore the

305 optimal threshold to choose for future experiments may need further consideration.

Conditional analyses and gene-based tests

307 The RAREMETALWORKER (RMW) tool[15] (version 4.13.3) that does not require individual level data to

308 perform conditional analyses and gene-based tests was used for conditional analyses. All studies that

309 contributed to the SNV discovery analyses were re-contacted and asked to run RMW. Only FENLAND,

310 GoDARTS, HELIC-MANOLIS, UKHLS and EPIC-InterAct were unable to run RMW, while two new

311 studies were included, INCIPE and NFBC1966 (Supplementary Table 1 and 2). In total, 43 studies (147,402

312 samples) were included in the EUR analyses and 45 studies (173,329 samples) in the EUR_SAS analyses

313 (Supplementary Tables 2 and 3). Comparison of discovery and RMW study level results were made

314 (Supplementary Information).

315 For each novel locus, the genomic coordinates and size of the region were defined according to

316 recombination rates (Supplementary Table 9) around the lead variant. For known loci, a 1 Mb window was

317 used (Supplementary Table 14). Conditional analyses were performed across each region, in both EUR and

318 EUR_SAS samples, for the transformed phenotype corresponding to the validated BP trait for novel loci and

319 the published BP trait for known loci.

320 Gene based tests were performed in both the EUR and EUR_SAS datasets using the Sequence Kernel

321 Association Test (SKAT)[16] method implemented in RMW as it allows for the SNVs to have different

322 directions and magnitudes of effect. Burden tests were also performed but are not presented as only SKAT

323 provided significant results. The variants in the gene-based tests using SKAT were weighted using the

324 default settings, *i.e.* a beta distribution density function to up-weight rare variants, Beta(MAFj,1,25) where

325 MAFj represents the pooled MAF for variant j across all studies. Analyses were restricted to coding SNVs

326 with MAF<5% and <1%. Genes were deemed to be associated if $P < 2.8 \times 10^{-6}$ (Bonferroni adjusted for

327 17,996 genes). To confirm the gene associations were not attributable to a solitary SNV, a gene-based test

328 conditional on the most associated SNV was performed ($P_{conditional} < 0.001$). The QC of all SNVs

329 contributing to the gene based tests including the number of samples and studies were checked prior to

330 claiming association. We sought replication of associated genes in the CHARGE+ exome chip blood

331 pressure consortium.

332

333 **Pathway analyses with MAGENTA**

334 We tested seven databases in MAGENTA[37] (BioCarta, Kyoto Encyclopedia of Genes and Genomes,

335 Ingenuity, Panther, Panther Biological Processes, Panther Molecular Functions and Reactome) for

336 overrepresentation of the SNV discovery results from both EUR and EUR_SAS ancestries. Each of the four

337 BP phenotypes were tested. Pathways exhibiting $P<0.01$ and FDR<5% were considered statistically

338 significant.

339 **GeneGo MetaCore Network analyses**

340   A set of BP genes based on previously published studies and our current results (locus defined as $r^2>0.4$ and

341   500kb on either side of the lead SNV; Supplementary Table 19) were tested for enrichment using the

342   THOMSON REUTERS MetaCore[TM] Single Experiment Analysis workflow tool. The data were mapped

343   onto selected MetaCore ontology databases: pathway maps, process networks, GO processes and diseases /

344   biomarkers, for which functional information is derived from experimental literature. Outputs were sorted

345   based on *P*- and FDR-values. A gene set was considered enriched for a particular process if $P<0.05$ and

346   FDR<5%.

347
348   **Genetic Risk Score**
349

350   To assess the effect of BP on CHD, ischemic stroke (and subtypes: large vessel, small vessel and

351   cardioembolic stroke) left ventricular mass, left ventricular wall thickness, heart failure, HDL-c, LDL-c,

352   total cholesterol, triglycerides and eGFR, we performed a weighted generalized linear regression of the

353   genetic associations with each outcome variable on the genetic associations with BP.

354   When genetic variants are uncorrelated, the estimates from such a weighted linear regression analysis using

355   summarized data, and a genetic risk score analysis using individual-level data, are equal[48]. We refer to the

356   analysis as a genetic risk score (also known as a polygenic risk score) analysis as this is likely to be more

357   familiar to applied readers. As some of the genetic variants in our analysis are correlated, a generalized

358   weighted linear regression model is fitted that accounts for the correlations between variants, as follows:

359   If β_X are the genetic associations (beta-coefficients) with the risk factor (here, BP) and β_Y are the genetic

360   associations with the outcome, then the causal estimate from a weighted generalized linear regression is $(\beta_X^T$

361     $\Omega^{-1}\beta_X)^{-1} \beta_X^T\Omega^{-1}\beta_Y$, with standard error,

362     $$\hat{\sigma}\sqrt{(\beta_X{}^T\Omega^{-1}\beta_X)^{-1}},$$

363   where ^{T} is a matrix transpose, $\hat{\sigma}$ is the estimate of the residual standard error from the regression model, and

364   the weighting matrix Ω has terms

365     $$\Omega_{j_1 j_2} = \sigma_{Y j_1}\sigma_{Y j_2}\rho_{j_1 j_2}$$

366   ,where σ_{Yj} is the standard error of the genetic association with the outcome for the *j*th SNV, and ρ_{j1j2} is the

367   correlation between the j_1th and j_2th SNVs. The presence of the estimated residual standard error allows for

368 heterogeneity between the causal estimates from the individual SNVs as overdispersion in the regression

369 model (in the case of underdispersion, the residual standard error estimate is set to unity). This is equivalent

370 to combining the causal estimates from each SNV using a multiplicative random-effects model[49].

371

372 For each of SBP, DBP and PP, the score was created using both the novel and known BP SNVs or a close

373 proxy ($r^2 > 0.8$). Both the sentinel SNV association and any secondary SNV associations that remained after

374 adjusting for the sentinel SNV were included in the genetic risk score. For the 30 validated novel SNV-BP

375 associations, βs were taken from the independent replication analyses (Table 1 and 2) to weight the SNV in

376 the genetic risk score. For the secondary SNVs from the seven novel loci and five known loci, βs were taken

377 from the discovery analyses (Supplementary Tables 10 and 15). For the 82 known SNVs, 43 were either

378 genotyped or had proxies on the Exome chip and the βs were taken from discovery results (Supplementary

379 Table 13), the remaining βs were taken from published effect estimates. This strategy for selecting betas for

380 use in the GRS was taken to minimize the influence of winner's curse. The associations between the BP

381 variants with CHD, HDL-c, LDL-c, total cholesterol, log(triglycerides) and log(eGFR) were obtained using

382 the CHD Exome+ Consortium studies, the associations with BMI, waist-hip ratio adjusted BMI and height

383 from the GIANT consortium (unpublished data), ischemic stroke from METASTROKE[25], and left

384 ventricular mass, left ventricular wall thickness and heart failure from EchoGen[27] and CHARGE-HF[26]. A

385 causal interpretation of the association of GRS with the outcome as the effect of BP on the outcome assumes

386 that the effects of genetic variants on the outcome are mediated via blood pressure and not via alternate

387 causal pathways, for example via LV thickness. There are also limitations of the Mendelian randomization

388 approach in distinguishing between the causal effects of different measures of blood pressure, due to the

389 paucity of genetic variants associated with only one measure of blood pressure.

390

391 **eQTL analyses**

392 The MuTHER dataset contains gene expression data from 850 UK twins for 23,596 probes and 2,029,988

393 (HapMap 2 imputed) SNVs. All cis–associated SNVs with FDR<1%, within each of the 30 novel regions

394 (IMPUTE info score >0.8) were extracted from the MuTHER project dataset for, LCL (n=777), adipose

395 (n=776) and skin (n=667) [50]. The pilot phase of the GTEx Project (dbGaP Accession phs000424.v3.p1)

396 provides expression data from up to 156 individuals for 52,576 genes and 6,820,472 genotyped SNVs

397 (imputed to 1000 Genomes project, MAF≥5%)[51]. The eQTL analysis was focused on subcutaneous adipose

398 tissue (n=94), tibial artery (n=112), heart (left ventricle) (n=83), lung (n=119), skeletal muscle (n=138),

399 tibial nerve (n=88), skin (sun exposed, lower leg) (n=96), thyroid (n=105) and whole blood (n=156) which

400 have >80 samples and genes expressed at least 0.1 RPKM in 10 or more individuals in a given tissue. All

401 transcripts with a transcription start site (TSS) within one of the 30 new BP loci and for which there was a

402 cis-associated SNV (IMPUTE info score >0.4) within 1Mb of the TSS at FDR<5%, were identified. Kidney

403 was not evaluated because the sample size was too small (n=8). From each resource, we report eQTL

404 signals, which reach the resource-specific thresholds for significance described above, for SNVs that are in

405 LD (r^2>0.8) with our sentinel SNV.

406 For identified eQTLs, we tested whether they colocalised with the BP associated SNV[52]. Colocalisation

407 analyses were considered to be significant if the posterior probability of colocalisation was greater than 0.95.

408 **Annotation of variants**

409 *In silico* prediction of the functional effect of associated variants was based on the annotation from dbSNP,

410 the Ensembl Variant Effect Predictor tool and the Exome Variant Server, NHLBI GO Exome Sequencing

411 Project (ESP), Seattle, WA.

412 **Trait variance explained**

413 The percentage trait variance explained for SBP, DBP, PP was assessed with 5,861 individuals with

414 complete information for all phenotypes and covariates from the population-based cohort, 1958BC.

415 Two genetic models were investigated: one containing the 43 previously known BP associated SNVs

416 covered on the Exome chip; the other additionally including the 30 novel lead SNVs and 9 conditionally

417 independent SNVs from both novel and known loci. These nine conditionally independent SNVs were

418 taken from the EUR results, as 1958BC is EUR. They included four from novel loci (*PREX1, COL21A1,*

419 *PRKAG1* and *MYH6* (there was only 1 in EUR); Supplementary Table 10) and five from known loci (*ST7L-*

420 *CAPZA1-MOV10, FIGN-GRB14, ENPEP, TBX5-TBX3* and *HOXC4*; Supplementary Table 15).

The residual trait was obtained by adjusting each of the BP traits in a regression model with sex and BMI variables (not age or age^2 as all 1958BC individuals were aged 44 years). The residual trait was regressed on all SNVs within the corresponding model and adjusted for the first ten PCs. The R^2 calculated from this regression model was used as the percentage trait variance explained.

Monogenic Enrichment analyses

To determine if sub-significant signals of association were present in a set of genes associated with monogenic forms of disease, we performed an enrichment analysis of the discovery single variant meta-analyses association results for all four traits, both for EUR and EUR_SAS datasets.

The monogenic gene set included: *WNK1*, *WNK4*, *KLHL3*, *CUL3*, *PPARG*, *NR3C2*, *CYP11B1*, *CYP11B2*, *CYP17A1*, *HSD11B2*, *SCNN1A*, *SCNN1B*, *SCNN1G*, *CLCNKB*, *KCNJ1*, *SLC12A1*, *SLC12A3*[3]. The association results of coding SNVs in these genes were extracted and the number of tests with $P<0.001$ observed. In order to determine how often such an observation would be observed by chance, we constructed 1,000 matched gene sets. The matching criteria for each monogenic gene was the intersection of all genes in the same exon length quintile and all genes in the same coding variant count decile. Within the matched sets, the number of variants with $P<0.001$ was observed. The empirical *P*-value was calculated as the fraction of matched sets with an equal or larger number of variants less than 0.001.

References

45. 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**, 2911-35 (2005).
46. Mahajan, A. *et al.* Identification and Functional Characterization of G6PC2 Coding Variants Influencing Glycemic Traits Define an Effector Transcript at the G6PC2-ABCB11 Locus. *PLoS Genet* **11**, e1004876 (2015).
47. Willer, C.J., Li, Y. & Abecasis, G.R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190-1 (2010).
48. Burgess, S. & Thompson, S.G. Multivariable Mendelian randomization: the use of pleiotropic genetic variants to estimate causal effects. *Am J Epidemiol* **181**, 251-60 (2015).
49. Thompson, S.G. & Sharp, S.J. Explaining heterogeneity in meta-analysis: a comparison of methods. *Stat Med* **18**, 2693-708 (1999).

452   50.   Nica, A.C. *et al.* The architecture of gene regulatory variation across multiple human tissues: the
453   MuTHER study. *PLoS Genet* **7**, e1002003 (2011).
454   51.   GTEx, Consortium,. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis:
455   multitissue gene regulation in humans. *Science* **348**, 648-60 (2015).
456   52.   Giambartolomei, C. *et al.* Bayesian test for colocalisation between pairs of genetic association
457   studies using summary statistics. *PLoS Genet* **10**, e1004383 (2014).

458