

1 **Genetic analysis of over one million people identifies 535 new loci associated with blood**
2 **pressure traits.**

3

4 Short title: blood pressure GWAS in one million people

5 Evangelos Evangelou^{1,2*}, Helen R Warren^{3,4*}, David Mosen-Ansorena^{1*}, Borbala Mifsud^{3*}, Raha
6 Pazoki^{1*}, He Gao^{1,5*}, Georgios Ntritsos^{2*}, Niki Dimou^{2*}, Claudia P Cabrera^{3,4}, Ibrahim Karaman¹,
7 Fu Liang Ng³, Marina Evangelou^{1,6}, Katarzyna Witkowska³, Evan Tzanis³, Jacklyn N Hellwege⁷,
8 Ayush Giri⁸, Digna R Velez Edwards⁸, Yan V Sun^{9,10}, Kelly Cho^{11,12}, J. Michael Gaziano^{11,12}, Peter
9 WF Wilson¹³, Philip S Tsao¹⁴, Csaba P Kovesdy¹⁵, Tonu Esko^{16,17}, Reedik Mägi¹⁶, Lili Milani¹⁶,
10 Peter Almgren¹⁸, Thibaud Boutin¹⁹, Stéphanie Debette^{20,21}, Jun Ding²², Franco Giulianini²³,
11 Elizabeth G Holliday²⁴, Anne U Jackson²⁵, Ruifang Li-Gao²⁶, Wei-Yu Lin²⁷, Jian'an Luan²⁸, Massimo
12 Mangino^{29,30}, Christopher Oldmeadow²⁴, Bram Peter Prins³¹, Yong Qian²², Muralidharan
13 Sargurupremraj²¹, Nabi Shah^{32,33}, Praveen Surendran²⁷, Sébastien Thériault^{34,35}, Niek
14 Verweij^{17,36,37}, Sara M Willems²⁸, Jing-Hua Zhao²⁸, Philippe Amouyel³⁸, John Connell³⁹, Renée de
15 Mutsert²⁶, Alex SF Doney³², Martin Farrall^{40,41}, Cristina Menni²⁹, Andrew D Morris⁴², Raymond
16 Noordam⁴³, Guillaume Paré³⁴, Neil R Poulter⁴⁴, Denis C Shields⁴⁵, Alice Stanton⁴⁶, Simon Thom⁴⁷,
17 Gonçalo Abecasis⁴⁸, Najaf Amin⁴⁹, Dan E Arking⁵⁰, Kristin L Ayers^{51,52}, Caterina M Barbieri⁵³,
18 Chiara Batini⁵⁴, Joshua C Bis⁵⁵, Tineka Blake⁵⁴, Murielle Bochud⁵⁶, Michael Boehnke²⁵, Eric
19 Boerwinkle⁵⁷, Dorret I Boomsma⁵⁸, Erwin P Bottinger⁵⁹, Peter S Braund^{60,61}, Marco Brumat⁶²,
20 Archie Campbell^{63,64}, Harry Campbell⁶⁵, Aravinda Chakravarti⁵⁰, John C Chambers^{1,5,66-68}, Ganesh
21 Chauhan⁶⁹, Marina Ciullo^{70,71}, Massimiliano Cocca⁷², Francis Collins⁷³, Heather J Cordell⁵¹, Gail
22 Davies^{74,75}, Martin H de Borst⁷⁶, Eco J de Geus⁵⁸, Ian J Deary^{74,75}, Joris Deelen⁷⁷, Fabiola Del Greco
23 M⁷⁸, Cumhur Yusuf Demirkale⁷⁹, Marcus Dörr^{80,81}, Georg B Ehret^{50,82}, Roberto Elosua^{83,84}, Stefan
24 Enroth⁸⁵, A Mesut Erzurumluoglu⁵⁴, Teresa Ferreira^{86,87}, Mattias Frånberg⁸⁸⁻⁹⁰, Oscar H Franco⁹¹,
25 Ilaria Gandin⁶², Paolo Gasparini^{62,72}, Vilmantas Giedraitis⁹², Christian Gieger⁹³⁻⁹⁵, Giorgia
26 Girotto^{62,72}, Anuj Goel^{40,41}, Alan J Gow^{74,96}, Vilmundur Gudnason^{97,98}, Xiuqing Guo⁹⁹, Ulf
27 Gyllenstein⁸⁵, Anders Hamsten^{88,89}, Tamara B Harris¹⁰⁰, Sarah E Harris^{63,74}, Catharina A
28 Hartman¹⁰¹, Aki S Havulinna^{102,103}, Andrew A Hicks⁷⁸, Edith Hofer^{104,105}, Albert Hofman^{91,106},
29 Jouke-Jan Hottenga⁵⁸, Jennifer E Huffman^{19,107,108}, Shih-Jen Hwang^{107,108}, Erik Ingelsson^{109,110},
30 Alan James^{111,112}, Rick Jansen¹¹³, Marjo-Riitta Jarvelin^{1,5,114-116}, Roby Joehanes^{107,117}, Åsa
31 Johansson⁸⁵, Andrew D Johnson^{107,118}, Peter K Joshi⁶⁵, Pekka Jousilahti¹⁰², J Wouter Jukema¹¹⁹,
32 Antti Jula¹⁰², Mika Kähönen^{120,121}, Sekar Kathiresan^{17,36,122}, Bernard D Keavney^{123,124}, Kay-Tee
33 Khaw¹²⁵, Paul Knekt¹⁰², Joanne Knight¹²⁶, Ivana Kolcic¹²⁷, Jaspal S Kooner^{5,67,68,128}, Seppo
34 Koskinen¹⁰², Kati Kristiansson¹⁰², Zoltan Kutalik^{56,129}, Maris Laan¹³⁰, Marty Larson¹⁰⁷, Lenore J
35 Launer¹⁰⁰, Benjamin Lehne¹, Terho Lehtimäki^{131,132}, David CM Liewald^{74,75}, Li Lin⁸², Lars Lind¹³³,
36 Cecilia M Lindgren^{40,87,134}, YongMei Liu¹³⁵, Ruth JF Loos^{28,59,136}, Lorna M Lopez^{74,137,138}, Yingchang
37 Lu⁵⁹, Leo-Pekka Lyytikäinen^{131,132}, Anubha Mahajan⁴⁰, Chrysovalanto Mamasoula¹³⁹, Jaume
38 Marrugat⁸³, Jonathan Marten¹⁹, Yuri Milaneschi¹⁴⁰, Anna Morgan⁶², Andrew P Morris^{40,141},
39 Alanna C Morrison¹⁴², Peter J Munson⁷⁹, Mike A Nalls^{143,144}, Priyanka Nandakumar⁵⁰, Christopher
40 P Nelson^{60,61}, Teemu Niiranen^{102,145}, Ilja M Nolte¹⁴⁶, Teresa Nutile⁷⁰, Albertine J Oldehinkel¹⁴⁷,
41 Ben A Oostra⁴⁹, Paul F O'Reilly¹⁴⁸, Elin Org¹⁶, Sandosh Padmanabhan^{64,149}, Walter Palmas¹⁵⁰,
42 Aarno Palotie^{103,151,152}, Alison Pattie⁷⁵, Brenda WJH Penninx¹⁴⁰, Markus Perola^{102,103,153}, Annette
43 Peters^{94,95,154}, Ozren Polasek^{127,155}, Peter P Pramstaller^{78,156,157}, Quang Tri Nguyen⁷⁹, Olli T
44 Raitakari^{158,159}, Meixia Ren¹⁶⁰, Rainer Rettig¹⁶¹, Kenneth Rice¹⁶², Paul M Ridker^{23,163}, Janina S
45 Ried⁹⁴, Harriette Riese¹⁴⁷, Samuli Ripatti^{103,164}, Antonietta Robino⁷², Lynda M Rose²³, Jerome I
46 Rotter⁹⁹, Igor Rudan¹⁶⁵, Daniela Ruggiero^{70,71}, Yasaman Saba¹⁶⁶, Cinzia F Sala⁵³, Veikko
47 Salomaa¹⁰², Nilesh J Samani^{60,61}, Antti-Pekka Sarin¹⁰³, Reinhold Schmidt¹⁰⁴, Helena Schmidt¹⁶⁶,

48 Nick Shrine⁵⁴, David Siscovick¹⁶⁷, Albert V Smith^{97,98}, Harold Snieder¹⁴⁶, Siim Söber¹³⁰, Rossella
49 Sorice⁷⁰, John M Starr^{74,168}, David J Stott¹⁶⁹, David P Strachan¹⁷⁰, Rona J Strawbridge^{88,89}, Johan
50 Sundström¹³³, Morris A Swertz¹⁷¹, Kent D Taylor⁹⁹, Alexander Teumer^{81,172}, Martin D Tobin⁵⁴,
51 Maciej Tomaszewski^{123,124}, Daniela Toniolo⁵³, Michela Traglia⁵³, Stella Trompet^{119,173}, Jaakko
52 Tuomilehto¹⁷⁴⁻¹⁷⁷, Christophe Tzourio²¹, André G Uitterlinden^{91,178}, Ahmad Vaez^{146,179}, Peter J van
53 der Most¹⁴⁶, Cornelia M van Duijn⁴⁹, Anne-Claire Vergnaud¹, Germaine C Verwoert⁹¹, Veronique
54 Vitart¹⁹, Uwe Völker^{81,180}, Peter Vollenweider¹⁸¹, Dragana Vuckovic^{62,182}, Hugh Watkins^{40,41},
55 Sarah H Wild¹⁸³, Gonneke Willemsen⁵⁸, James F Wilson^{19,65}, Alan F Wright¹⁹, Jie Yao⁹⁹, Tatijana
56 Zemunik¹⁸⁴, Weihua Zhang^{1,67}, John R Attia²⁴, Adam S Butterworth^{27,185}, Daniel I Chasman^{23,163},
57 David Conen^{186,187}, Francesco Cucca^{188,189}, John Danesh^{27,185}, Caroline Hayward¹⁹, Joanna MM
58 Howson²⁷, Markku Laakso¹⁹⁰, Edward G Lakatta¹⁹¹, Claudia Langenberg²⁸, Olle Melander¹⁸,
59 Dennis O Mook-Kanamori^{26,192}, Colin NA Palmer³², Lorenz Risch¹⁹³⁻¹⁹⁵, Robert A Scott²⁸, Rodney J
60 Scott²⁴, Peter Sever¹²⁸, Tim D Spector²⁹, Pim van der Harst¹⁹⁶, Nicholas J Wareham²⁸, Eleftheria
61 Zeggini³¹, Daniel Levy^{107,118}, Patricia B Munroe^{3,4}, Christopher Newton-Cheh^{134,197,198}, Morris J
62 Brown^{3,4}, Andres Metspalu¹⁶, Adriana M Hung¹⁹⁹, Christopher J O'Donnell²⁰⁰, Todd L Edwards⁷
63 on behalf of the Million Veteran Program, Bruce M. Psaty^{201,202}, Ioanna Tzoulaki^{1,2,5*}, Michael R
64 Barnes^{3,4*}, Louise V Wain^{54*}, Paul Elliott^{1,5,203-205*‡}, Mark J Caulfield^{3,4*‡}

65

66 * Equal contribution

67 ‡ Corresponding authors

68

- 69 1. Department of Epidemiology and Biostatistics, Imperial College London, London,
70 UK.
- 71 2. Department of Hygiene and Epidemiology, University of Ioannina Medical
72 School, Ioannina, Greece.
- 73 3. William Harvey Research Institute, Barts and The London School of Medicine and
74 Dentistry, Queen Mary University of London, London, UK.
- 75 4. National Institute for Health Research, Barts Cardiovascular Biomedical
76 Research Center, Queen Mary University of London, London, UK.
- 77 5. MRC-PHE Centre for Environment and Health, Imperial College London, London,
78 UK.
- 79 6. Department of Mathematics, Imperial College London, London, UK
- 80 7. Division of Epidemiology, Department of Medicine, Institute for Medicine and
81 Public Health, Vanderbilt Genetics Institute, Vanderbilt University Medical
82 Center, Tennessee Valley Healthcare System (626)/Vanderbilt University,
83 Nashville, TN, USA.
- 84 8. Vanderbilt Genetics Institute, Vanderbilt Epidemiology Center, Department of
85 Obstetrics and Gynecology, Vanderbilt University Medical Center; Tennessee
86 Valley Health Systems VA, Nashville, TN, USA.
- 87 9. Department of Epidemiology, Emory University Rollins School of Public Health,
88 Atlanta, GA, USA.
- 89 10. Department of Biomedical Informatics, Emory University School of Medicine,
90 Atlanta, GA, USA.
- 91 11. Massachusetts Veterans Epidemiology Research and Information Center
92 (MAVERIC), VA Boston Healthcare System, Boston, USA.

- 93 12. Division of Aging, Department of Medicine, Brigham and Women's Hospital,
94 Boston, MA, Department of Medicine, Harvard Medical School, Boston, MA, USA.
- 95 13. Atlanta VAMC and Emory Clinical Cardiovascular Research Institute, Atlanta, GA,
96 USA.
- 97 14. VA Palo Alto Health Care System; Division of Cardiovascular Medicine, Stanford
98 University School of Medicine, CA, USA.
- 99 15. Nephrology Section, Memphis VA Medical Center and University of Tennessee
100 Health Science Center, Memphis, TN, USA.
- 101 16. Estonian Genome Center, University of Tartu, Tartu, Estonia.
- 102 17. Program in Medical and Population Genetics, Broad Institute of Harvard and
103 MIT, Cambridge, MA, USA.
- 104 18. Department Clinical Sciences, Malmö, Lund University, Malmö, Sweden.
- 105 19. MRC Human Genetics Unit, MRC Institute of Genetics and Molecular Medicine,
106 University of Edinburgh, Western General Hospital, Edinburgh, Scotland, UK
- 107 20. Department of Neurology, Bordeaux University Hospital, Bordeaux, France.
- 108 21. Univ. Bordeaux, Inserm, Bordeaux Population Health Research Center, CHU
109 Bordeaux, Bordeaux, France.
- 110 22. Laboratory of Genetics and Genomics, NIA/NIH , Baltimore, MD, USA.
- 111 23. Division of Preventive Medicine, Brigham and Women's Hospital, Boston, MA,
112 USA.
- 113 24. Hunter Medical Research Institute and Faculty of Health, University of Newcastle,
114 New Lambton Heights, New South Wales, Australia.
- 115 25. Department of Biostatistics and Center for Statistical Genetics, University of
116 Michigan, Ann Arbor, MI, USA.
- 117 26. Department of Clinical Epidemiology, Leiden University Medical Center, Leiden,
118 the Netherlands.
- 119 27. MRC/BHF Cardiovascular Epidemiology Unit, Department of Public Health and
120 Primary Care, University of Cambridge, Cambridge, UK.
- 121 28. MRC Epidemiology Unit, University of Cambridge School of Clinical Medicine,
122 Cambridge, UK.
- 123 29. Department of Twin Research and Genetic Epidemiology, Kings College London,
124 London, UK.
- 125 30. NIHR Biomedical Research Centre at Guy's and St Thomas' Foundation Trust,
126 London, UK.
- 127 31. Wellcome Trust Sanger Institute, Hinxton, UK.
- 128 32. Division of Molecular and Clinical Medicine, School of Medicine, University of
129 Dundee, UK.
- 130 33. Department of Pharmacy, COMSATS Institute of Information Technology,
131 Abbottabad, Pakistan.
- 132 34. Department of Pathology and Molecular Medicine, McMaster University,
133 Hamilton, Canada.
- 134 35. Institut universitaire de cardiologie et de pneumologie de Québec-Université
135 Laval, , Quebec City, Canada.
- 136 36. Cardiovascular Research Center and Center for Human Genetic Research,
137 Massachusetts General Hospital, Boston, Massachusetts, MA, USA.
- 138 37. University of Groningen, University Medical Center Groningen, Department of
139 Cardiology, Groningen, The Netherlands.

- 140 38. University of Lille, Inserm, Centre Hosp. Univ Lille, Institut Pasteur de Lille,
141 UMR1167 - RID-AGE - Risk factors and molecular determinants of aging-related
142 diseases, Epidemiology and Public Health Department, Lille, France.
- 143 39. University of Dundee, Ninewells Hospital & Medical School, Dundee, , UK.
- 144 40. Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK.
- 145 41. Division of Cardiovascular Medicine, Radcliffe Department of Medicine,
146 University of Oxford, Oxford, UK.
- 147 42. Usher Institute of Population Health Sciences and Informatics, University of
148 Edinburgh, UK.
- 149 43. Department of Internal Medicine, Section Gerontology and Geriatrics, Leiden
150 University Medical Center, Leiden, The Netherlands.
- 151 44. Imperial Clinical Trials Unit, Stadium House, 68 Wood Lane, London, UK.
- 152 45. School of Medicine, University College Dublin, Ireland.
- 153 46. Molecular and Cellular Therapeutics, Royal College of Surgeons in Ireland,
154 Dublin, Ireland.
- 155 47. International Centre for Circulatory Health, Imperial College London, London,
156 UK.
- 157 48. Center for Statistical Genetics, Dept. of Biostatistics, SPH II, Washington Heights,
158 Ann Arbor, MI, USA.
- 159 49. Genetic Epidemiology Unit, Department of Epidemiology, Erasmus MC,
160 Rotterdam, the Netherlands.
- 161 50. Center for Complex Disease Genomics, McKusick-Nathans Institute of Genetic
162 Medicine, Johns Hopkins University School of Medicine, Baltimore, MD, USA.
- 163 51. Institute of Genetic Medicine, Newcastle University, Newcastle upon Tyne, UK.
- 164 52. Sema4, a Mount Sinai venture, Stamford, CT, USA.
- 165 53. Division of Genetics and Cell Biology, San Raffaele Scientific Institute, Milano,
166 Italy.
- 167 54. Department of Health Sciences, University of Leicester, Leicester, UK.
- 168 55. Cardiovascular Health Research Unit, Department of Medicine, University of
169 Washington, Seattle, WA, USA.
- 170 56. Institute of Social and Preventive Medicine, University Hospital of Lausanne,
171 Lausanne, Switzerland.
- 172 57. Human Genetics Center, School of Public Health, The University of Texas Health
173 Science Center at Houston and Human Genome Sequencing Center, Baylor
174 College of Medicine, One Baylor Plaza, Houston, TX, USA.
- 175 58. Department of Biological Psychology, Vrije Universiteit Amsterdam, EMGO+
176 institute, VU University medical center, Amsterdam, the Netherlands.
- 177 59. The Charles Bronfman Institute for Personalized Medicine, Icahn School of
178 Medicine at Mount Sinai, NY, USA.
- 179 60. Department of Cardiovascular Sciences, University of Leicester, Leicester, UK.
- 180 61. NIHR Leicester Biomedical Research Centre, Glenfield Hospital, Groby Road,
181 Leicester, UK.
- 182 62. Department of Medical, Surgical and Health Sciences, University of Trieste, ,
183 Trieste, Italy.
- 184 63. Medical Genetics Section, Centre for Genomic and Experimental Medicine,
185 Institute of Genetics and Molecular Medicine, University of Edinburgh,
186 Edinburgh, UK.
- 187 64. Generation Scotland, Centre for Genomic and Experimental Medicine, University
188 of Edinburgh, Edinburgh, UK.

- 189 65. Centre for Global Health Research, Usher Institute of Population Health Sciences
190 and Informatics, University of Edinburgh, Edinburgh, Scotland, UK
- 191 66. Lee Kong Chian School of Medicine, Nanyang Technological University,
192 Singapore, Singapore.
- 193 67. Department of Cardiology, Ealing Hospital, Middlesex, UK.
- 194 68. Imperial College Healthcare NHS Trust, London, UK.
- 195 69. Centre for Brain Research, Indian Institute of Science, Bangalore, India.
- 196 70. Institute of Genetics and Biophysics "A. Buzzati-Traverso", CNR, Napoli, Italy.
- 197 71. IRCCS Neuromed, Pozzilli, Isernia, Italy.
- 198 72. Institute for Maternal and Child Health IRCCS Burlo Garofolo, Trieste, Italy.
- 199 73. Medical Genomics and Metabolic Genetics Branch, National Human Genome
200 Research Institute, NIH, Bethesda, MD, USA.
- 201 74. Centre for Cognitive Ageing and Cognitive Epidemiology, University of
202 Edinburgh, 7 George Square, Edinburgh, UK.
- 203 75. Department of Psychology, University of Edinburgh, 7 George Square, Edinburgh,
204 UK.
- 205 76. Department of Internal Medicine, Division of Nephrology, University of
206 Groningen, University Medical Center Groningen, Groningen, The Netherlands.
- 207 77. Department of Molecular Epidemiology, Leiden University Medical Center,
208 Leiden, the Netherlands.
- 209 78. Institute for Biomedicine, Eurac Research, Bolzano, Italy - Affiliated Institute of
210 the University of Lübeck, Lübeck, Germany.
- 211 79. Mathematical and Statistical Computing Laboratory, Office of Intramural
212 Research, Center for Information Technology, National Institutes of Health,
213 Bethesda, MD, USA.
- 214 80. Department of Internal Medicine B, University Medicine Greifswald, Greifswald,
215 Germany.
- 216 81. DZHK (German Centre for Cardiovascular Research), partner site Greifswald,
217 Greifswald, Germany.
- 218 82. Cardiology, Department of Medicine, Geneva University Hospital, Geneva,
219 Switzerland.
- 220 83. CIBERCV & Cardiovascular Epidemiology and Genetics, IMIM. Dr Aiguader 88,
221 Barcelona, Spain.
- 222 84. Faculty of Medicine, Universitat de Vic-Central de Catalunya, Vic, Spain.
- 223 85. Department of Immunology, Genetics and Pathology, Uppsala Universitet,
224 Science for Life Laboratory, Uppsala, Sweden.
- 225 86. Wellcome Centre for Human Genetics, University of Oxford, Roosevelt Drive,
226 Oxford, UK.
- 227 87. Big Data Institute, Li Ka Shing Center for Health for Health Information and
228 Discovery, Oxford University, Old Road, Oxford, UK.
- 229 88. Cardiovascular Medicine Unit, Department of Medicine Solna, Karolinska
230 Institutet, Stockholm, Sweden.
- 231 89. Centre for Molecular Medicine, L8:03, Karolinska Universitetsjukhuset, Solna,
232 Sweden.
- 233 90. Department of Numerical Analysis and Computer Science, Stockholm University,
234 Stockholm, Sweden.
- 235 91. Department of Epidemiology, Erasmus MC, Rotterdam, the Netherlands.
- 236 92. Department of Public Health and Caring Sciences, Geriatrics, Uppsala, Sweden.

- 237 93. Research Unit of Molecular Epidemiology, Helmholtz Zentrum München, German
238 Research Center for Environmental Health, Neuherberg, Germany.
- 239 94. Institute of Epidemiology, Helmholtz Zentrum München, German Research
240 Center for Environmental Health, Neuherberg, Germany.
- 241 95. German Center for Diabetes Research (DZD e.V.), Neuherberg, Germany.
- 242 96. Department of Psychology, School of Social Sciences, Heriot-Watt University,
243 Edinburgh, UK.
- 244 97. Faculty of Medicine, University of Iceland, Reykjavik, Iceland.
- 245 98. Icelandic Heart Association, Kopavogur, Iceland.
- 246 99. The Institute for Translational Genomics and Population Sciences, Department of
247 Pediatrics, LABioMed at Harbor-UCLA Medical Center, Torrance, CA, USA.
- 248 100. Intramural Research Program, Laboratory of Epidemiology, Demography, and
249 Biometry, National Institute on Aging, Bethesda, MD, USA.
- 250 101. Department of Psychiatry, University of Groningen, University Medical Center
251 Groningen, Groningen, The Netherlands.
- 252 102. Department of Public Health Solutions, National Institute for Health and Welfare
253 (THL), Helsinki, Finland.
- 254 103. Institute for Molecular Medicine Finland (FIMM), University of Helsinki, Helsinki,
255 Finland.
- 256 104. Clinical Division of Neurogeriatrics, Department of Neurology, Medical
257 University of Graz, Graz, Austria.
- 258 105. Institute for Medical Informatics, Statistics and Documentation, Medical
259 University of Graz, Graz, Austria.
- 260 106. Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston,
261 MA, USA.
- 262 107. National Heart, Lung and Blood Institute's Framingham Heart Study,
263 Framingham, MA, USA.
- 264 108. The Population Science Branch, Division of Intramural Research, National Heart
265 Lung and Blood Institute national Institute of Health, Bethesda, MD, USA.
- 266 109. Department of Medical Sciences, Molecular Epidemiology and Science for Life
267 Laboratory, Uppsala University, Uppsala, Sweden.
- 268 110. Division of Cardiovascular Medicine, Department of Medicine, Stanford
269 University School of Medicine, Stanford, CA USA.
- 270 111. Department of Pulmonary Physiology and Sleep, Sir Charles Gairdner Hospital,
271 Hospital Avenue, Nedlands, Australia.
- 272 112. School of Medicine and Pharmacology, University of Western Australia.
- 273 113. Department of Psychiatry, VU University Medical Center, Amsterdam
274 Neuroscience, Amsterdam, the Netherlands.
- 275 114. Biocenter Oulu, University of Oulu, Oulu, Finland.
- 276 115. Center For Life-course Health Research, University of Oulu, Oulu Finland.
- 277 116. Unit of Primary Care, Oulu University Hospital, Oulu, Oulu, Finland.
- 278 117. Hebrew SeniorLife, Harvard Medical School, Boston, MA, USA.
- 279 118. Population Sciences Branch, National Heart, Lung and Blood Institute, National
280 Institutes of Health, Bethesda, MD, USA.
- 281 119. Department of Cardiology, Leiden University Medical Center, Leiden, the
282 Netherlands.
- 283 120. Department of Clinical Physiology, Tampere University Hospital, Tampere,
284 Finland.

- 285 121. Department of Clinical Physiology, Finnish Cardiovascular Research Center -
286 Tampere, Faculty of Medicine and Life Sciences, University of Tampere,
287 Tampere, Finland.
- 288 122. Broad Institute of the Massachusetts Institute of Technology and Harvard
289 University, Cambridge, MA, USA.
- 290 123. Division of Cardiovascular Sciences, Faculty of Biology, Medicine and Health, The
291 University of Manchester, Manchester, UK.
- 292 124. Division of Medicine, Manchester University NHS Foundation Trust, Manchester
293 Academic Health Science Centre, Manchester, UK
- 294 125. Department of Public Health and Primary Care, Institute of Public Health,
295 University of Cambridge, Cambridge, UK.
- 296 126. Data Science Institute and Lancaster Medical School, Lancaster, UK.
- 297 127. Department of Public Health, Faculty of Medicine, University of Split, Croatia.
- 298 128. National Heart and Lung Institute, Imperial College London, London, UK.
- 299 129. Swiss Institute of Bioinformatics, Lausanne, Switzerland.
- 300 130. Institute of Biomedicine and Translational Medicine, University of Tartu, Tartu,
301 Estonia.
- 302 131. Department of Clinical Chemistry, Fimlab Laboratories, Tampere, Finland.
- 303 132. Department of Clinical Chemistry, Finnish Cardiovascular Research Center -
304 Tampere, Faculty of Medicine and Life Sciences, University of Tampere,
305 Tampere, Finland
- 306 133. Department of Medical Sciences, Cardiovascular Epidemiology, Uppsala
307 University, Uppsala, Sweden.
- 308 134. Program in Medical and Population Genetics, Broad Institute, Cambridge, MA,
309 USA.
- 310 135. Division of Public Health Sciences, Wake Forest School of Medicine, Winston-
311 Salem, NC, USA.
- 312 136. Mindich Child health Development Institute, The Icahn School of Medicine at
313 Mount Sinai, New York, NY, USA.
- 314 137. Department of Psychiatry, Royal College of Surgeons in Ireland, Education and
315 Research Centre, Beaumont Hospital, Dublin, Ireland.
- 316 138. University College Dublin, UCD Conway Institute, Centre for Proteome Research,
317 UCD, Belfield, Dublin, Ireland.
- 318 139. Institute of Health and Society, Newcastle University, Newcastle upon Tyne, UK.
- 319 140. Department of Psychiatry, Amsterdam Public Health and Amsterdam
320 Neuroscience, VU University Medical Center/GGZ inGeest, Amsterdam, The
321 Netherlands.
- 322 141. Department of Biostatistics, University of Liverpool, Block F, Waterhouse
323 Building, Liverpool, UK.
- 324 142. Department of Epidemiology, Human Genetics and Environmental Sciences,
325 School of Public Health, University of Texas Health Science Center at Houston,
326 Houston, TX, USA.
- 327 143. Data Tecnica International, Glen Echo, MD, USA.
- 328 144. Laboratory of Neurogenetics, National Institute on Aging, Bethesda, USA.
- 329 145. Department of Medicine, Turku University Hospital and University of Turku,
330 Finland.
- 331 146. Department of Epidemiology, University of Groningen, University Medical Center
332 Groningen, Groningen, The Netherlands.

- 333 147. Interdisciplinary Center Psychopathology and Emotion regulation (ICPE),
334 University of Groningen, University Medical Center Groningen, Groningen, The
335 Netherlands.
- 336 148. SGDP Centre, Institute of Psychiatry, Psychology and Neuroscience, King's
337 College London, London, UK.
- 338 149. British Heart Foundation Glasgow Cardiovascular Research Centre, Institute of
339 Cardiovascular and Medical Sciences, College of Medical, Veterinary and Life
340 Sciences, University of Glasgow, Glasgow, UK.
- 341 150. Department of Medicine, Columbia University Medical Center, New York, NY,
342 USA.
- 343 151. Analytic and Translational Genetics Unit, Department of Medicine, Department of
344 Neurology and Department of Psychiatry Massachusetts General Hospital,
345 Boston, MA, USA.
- 346 152. The Stanley Center for Psychiatric Research and Program in Medical and
347 Population Genetics, The Broad Institute of MIT and Harvard, Cambridge, MA,
348 USA.
- 349 153. University of Tartu, Tartu, Estonia.
- 350 154. German Center for Cardiovascular Disease Research (DZHK), partner site
351 Munich, Neuherberg, Germany.
- 352 155. Psychiatric hospital "Sveti Ivan", Zagreb, Croatia.
- 353 156. Department of Neurology, General Central Hospital, Bolzano, Italy.
- 354 157. Department of Neurology, University of Lübeck, Lübeck, Germany.
- 355 158. Department of Clinical Physiology and Nuclear Medicine, Turku University
356 Hospital, Turku, Finland.
- 357 159. Research Centre of Applied and Preventive Cardiovascular Medicine, University
358 of Turku, Turku, Finland.
- 359 160. Fujian Key Laboratory of Geriatrics, Department of Geriatric Medicine, Fujian
360 Provincial Hospital, Fujian Medical University, Fuzhou, China.
- 361 161. Institute of Physiology, University Medicine Greifswald, Karlsburg, Germany.
- 362 162. Department of Biostatistics University of Washington, Seattle, WA, USA.
- 363 163. Harvard Medical School, Boston MA.
- 364 164. Public health, Faculty of Medicine, University of Helsinki, Finland
- 365 165. Centre for Global Health Research, Usher Institute of Population Health Sciences
366 and Informatics, University of Edinburgh, Scotland, UK.
- 367 166. Gottfried Schatz Research Center for Cell Signaling, Metabolism & Aging,
368 Molecular Biology and Biochemistry, Medical University of Graz, Graz, Austria.
- 369 167. The New York Academy of Medicine, New York, NY, USA.
- 370 168. Alzheimer Scotland Dementia Research Centre, University of Edinburgh,
371 Edinburgh, UK.
- 372 169. Institute of Cardiovascular and Medical Sciences, Faculty of Medicine, University
373 of Glasgow, United Kingdom.
- 374 170. Population Health Research Institute, St George's, University of London, London,
375 UK.
- 376 171. Department of Genetics, University of Groningen, University Medical Center
377 Groningen, Groningen, The Netherlands.
- 378 172. Institute for Community Medicine, University Medicine Greifswald, Greifswald,
379 Germany.
- 380 173. Department of Gerontology and Geriatrics, Leiden University Medical Center,
381 Leiden, the Netherlands.

- 382 174. Dasman Diabetes Institute, Dasman, Kuwait.
383 175. Chronic Disease Prevention Unit, National Institute for Health and Welfare,
384 Helsinki, Finland.
385 176. Department of Public Health, University of Helsinki, Helsinki, Finland.
386 177. Saudi Diabetes Research Group, King Abdulaziz University, Jeddah, Saudi Arabia.
387 178. Department of Internal Medicine, Erasmus MC, Rotterdam, the Netherlands.
388 179. Research Institute for Primordial Prevention of Non-communicable Disease,
389 Isfahan University of Medical Sciences, Isfahan, Iran.
390 180. Interfaculty Institute for Genetics and Functional Genomics, University Medicine
391 Greifswald, Greifswald, Germany.
392 181. Department of Internal Medicine, University Hospital, CHUV, Lausanne,
393 Switzerland.
394 182. Experimental Genetics Division, Sidra Medical and Research Center, Doha, Qatar.
395 183. Centre for Population Health Sciences, Usher Institute of Population Health
396 Sciences and Informatics, University of Edinburgh, Scotland, UK
397 184. Department of Biology, Faculty of Medicine, University of Split, Croatia.
398 185. The National Institute for Health Research Blood and Transplant Research Unit
399 in Donor Health and Genomics, University of Cambridge, UK.
400 186. Division of Cardiology, University Hospital, Basel, Switzerland.
401 187. Division of Cardiology, Department of Medicine, McMaster University, Hamilton,
402 Canada.
403 188. Institute of Genetic and Biomedical Research, National Research Council (CNR),
404 Monserrato, Cagliari, Italy.
405 189. Department of Biomedical Sciences, University of Sassari, Sassari, Italy.
406 190. Institute of Clinical Medicine, Internal Medicine, University of Eastern Finland
407 and Kuopio University Hospital, Kuopio, Finland.
408 191. Laboratory of Cardiovascular Science, NIA/NIH , Baltimore, MD, USA.
409 192. Department of Public Health and Primary Care, Leiden University Medical
410 Center, Leiden, the Netherlands.
411 193. Labormedizinisches Zentrum Dr. Risch, Schaan, Liechtenstein.
412 194. Private University of the Principality of Liechtenstein, Triesen, Liechtenstein.
413 195. University Institute of Clinical Chemistry, Inselspital, Bern University Hospital,
414 University of Bern, Bern, Switzerland.
415 196. Department of Cardiology, University of Groningen, University Medical Center
416 Groningen, Groningen, The Netherlands.
417 197. Center for Genomic Medicine, Massachusetts General Hospital, Boston, MA, USA.
418 198. Cardiovascular Research Center, Massachusetts General Hospital, Boston, MA,
419 USA.
420 199. Tennessee Valley Healthcare System (Nashville VA) & Vanderbilt University, TN,
421 USA.
422 200. VA Boston Healthcare, Section of Cardiology and Department of Medicine,
423 Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA.
424 201. Cardiovascular Health Research Unit, Departments of Medicine, Epidemiology
425 and Health Services, University of Washington, Seattle, WA, USA.
426 202. Kaiser Permanente Washington Health Research Institute, Seattle, WA, USA.
427 203. National Institute for Health Research Imperial Biomedical Research Centre,
428 Imperial College Healthcare NHS Trust and Imperial College London, London,
429 UK.

430 204. UK Dementia Research Institute (UK DRI) at Imperial College London, London,
431 UK
432 205. Health Data Research-UK London substantive site, London, U.K
433
434

435 Corresponding authors: Mark Caulfield (m.j.caulfield@qmul.ac.uk) and Paul Elliott
436 (p.elliott@imperial.ac.uk)

437

438 **Abstract**

439 High blood pressure is a highly heritable and modifiable risk factor for cardiovascular
440 disease. We report the largest genetic association study of blood pressure traits (systolic,
441 diastolic, pulse pressure) to date in over one million people of European ancestry. We
442 identify 535 novel blood pressure loci that not only offer new biological insights into blood
443 pressure regulation but also reveal shared genetic architecture between blood pressure and
444 lifestyle exposures. Our findings identify new biological pathways for blood pressure
445 regulation with potential for improved cardiovascular disease prevention in the future.

446

447 INTRODUCTION

448 High blood pressure (BP) is a leading heritable risk factor for stroke and coronary artery
449 disease, responsible for an estimated 7.8 million deaths and 148 million disability life years
450 lost worldwide in 2015 alone¹. Blood pressure is determined by complex interactions
451 between life-course exposures and genetic background²⁻⁴. Previous genetic association
452 studies have identified and validated variants at 274 loci with modest effects on population
453 BP, explaining in aggregate ~3% of the trait variance⁵⁻¹².

454 Here, we report genome-wide discovery analyses of BP traits - systolic (SBP), diastolic (DBP)
455 and pulse pressure (PP) - in people of European ancestry drawn from UK Biobank (UKB)¹³
456 and the International Consortium of Blood Pressure-Genome Wide Association Studies
457 (ICBP)^{11,12}. We adopted a combination of a one- and two-stage study design to test common
458 and low-frequency single nucleotide polymorphisms (SNPs) with minor allele frequency
459 (MAF) $\geq 1\%$ associated with BP traits (**Fig. 1**). In all, we studied over 1 million people of
460 European descent, including replication data from the US Million Veterans Program (MVP,
461 N=220,520)¹⁴ and the Estonian Genome Centre, University of Tartu (EGCUT, N=28,742)
462 Biobank¹⁵.

463 UKB is a prospective cohort study of ~500,000 richly phenotyped individuals, including BP
464 measurements¹³, with genotyping by customized array and imputation from the Haplotype
465 Reference Consortium (HRC) panel, yielding ~7 million SNPs (imputation quality score (INFO)
466 ≥ 0.1 and MAF $\geq 1\%$)¹⁶. We performed genome-wide association studies (GWAS) of BP traits
467 (N=458,577 Europeans) under an additive genetic model¹⁷ (**Supplementary Table 1a**).
468 Following LD-score regression¹⁸, genomic control (GC) was applied to the UKB data prior to
469 meta-analysis (Online methods).

470 In addition, we performed GWAS analyses for BP traits in newly extended ICBP GWAS data
471 comprising 77 independent studies for up to 299,024 Europeans genotyped with various
472 arrays, and imputed to either the 1,000 Genomes Reference Panel or the HRC platforms
473 (**Supplementary Table 1b**). After QC we applied GC at the individual study level and
474 obtained summary effect sizes for ~7 million SNPs with INFO ≥ 0.3 and heterogeneity
475 Cochran's Q statistic¹⁹ filtered at $P \geq 1 \times 10^{-4}$ (Online Methods).

476 We then combined the UKB and ICBP GWAS results using inverse-variance weighted fixed
477 effects meta-analysis (Online Methods), giving a total discovery sample of up to 757,601
478 individuals²⁰.

479 In our two-stage design we attempted replication (in MVP and EGCUT, **Supplementary**
480 **Table 1c**) of 1,062 SNPs at $P < 1 \times 10^{-6}$ from discovery with concordant effect direction
481 between UKB and ICBP, using the sentinel SNP (i.e. SNP with smallest P -value at the locus)
482 after excluding the HLA region (chr 6:25-34MB) and all SNPs in Linkage Disequilibrium (LD)
483 ($r^2 \geq 0.1$) or ± 500 Kb from any previously validated BP-associated SNPs at the 274 published
484 loci. Our replication criteria were genome-wide significance ($P < 5 \times 10^{-8}$) in the combined
485 meta-analysis, $P < 0.01$ in the replication data and concordant direction of effect between
486 discovery and replication.

487 We additionally undertook a one-stage design to reduce type II error from the two-stage
488 analysis. We used $P < 5 \times 10^{-9}$ as threshold from the discovery meta-analysis, i.e. an order of
489 magnitude more stringent than genome-wide significance²¹, and required an internal
490 replication $P < 0.01$ in each of the UKB and ICBP GWAS analyses, with concordant direction
491 of effect, to minimize false positive findings.

492 We carried out conditional analyses using genome-wide complex trait analysis (GCTA)²². We
493 then explored putative function of BP-associated signals using a range of *in silico* resources,
494 and evaluated co-occurrence of BP-associated loci with lifestyle exposures and other
495 complex traits and diseases. Finally, we developed a genetic risk score (GRS) and assessed
496 impact of BP-associated variants on BP level, risk of hypertension (HTN), other
497 cardiovascular diseases and in other ethnicities.

498 RESULTS

499 We present a total of 535 novel loci (**Fig.2, Supplementary Fig. 1**): 325 loci claimed from the
500 two-stage design (**Supplementary Tables 2a-c**) and an additional 210 claimed from our one-
501 stage design with internal replication (**Supplementary Tables 3a-c**). Our two-stage design
502 uniquely identified 121 variants, while 204 also met the one-stage criteria (**Fig. 3a**); large
503 numbers of loci would not have been detected by either the one- or two-stage designs
504 alone (**Fig. 3a**). For SBP, the distributions of effect sizes are similar for the one-stage
505 (median = 0.219 mmHg per allele; Inter-Quartile Range (IQR) = 0.202-0.278) and two-stage
506 loci (median = 0.224; IQR = 0.195-0.267) ($P = 0.447$) (**Supplementary Fig. 2**). Of the 210 loci
507 found only in the one-stage analysis, 186 are also genome-wide significant ($P < 5 \times 10^{-8}$) in
508 the combined meta-analysis, with all variants, except one, having concordant direction of
509 effect between discovery and replication (**Supplementary Tables 3a-c**); of the remaining 24
510 SNPs, 10 still have concordant direction of effect.

511 We find support in our data for all 274 previously published BP loci (**Supplementary Fig. 1 &**
512 **2 and Supplementary Table 4**); >95% of the previously reported SNPs covered within our
513 data are genome-wide significant. Only 6 available SNPs did not reach Bonferroni-
514 significance, likely because they were originally identified in non-European ancestries (e.g.
515 rs6749447, rs10474346, rs11564022), or from a gene-age interaction analysis (rs16833934).
516 In addition, we confirmed a further 92 previously reported, but not replicated, loci
517 (**Supplementary Table 5**)⁹; together with 274 previously reported loci confirmed, and 535
518 novel loci identified here, there are 901 BP-associated loci in total.

519 Novel genetic loci for blood pressure

520 Of the 535 independent novel loci, 363 SNPs were associated with one trait, 160 with two
521 traits and 12 with all three BP traits (**Fig. 3b, Supplementary Fig. 3**). Using GCTA we
522 additionally identified 163, genome-wide significant, independent secondary signals with
523 MAF $\geq 1\%$ associated with BP (**Supplementary Table 6**), of which 19 SNPs are in LD ($r^2 \geq 0.1$)
524 with previously reported secondary signals. This gives a total of 144 new secondary signals;
525 hence we now report over 1,000 independent BP signals.

526 The estimated SNP-wide heritability (h^2) of BP traits in our data was 0.213, 0.212 and 0.194
527 for SBP, DBP and PP respectively, with a gain in percentage of BP variance explained. For

528 example, for SBP, percentage variance explained increased from 2.8 % for the 274
529 previously published loci to 5.7% for SNPs identified at all 901 loci (**Supplementary Table 7**).

530 **Functional analyses**

531 Our functional analyses approach is summarised in **Supplementary Figure 4**. First, for each
532 of the 901 loci we annotated all SNPs (based on LD $r^2 \geq 0.8$) to the nearest gene within 5kb
533 of a SNP, identifying 1333 genes for novel loci and 1272 genes for known loci. Then we
534 investigated these loci for tissue enrichment, DNase hypersensitivity site enrichment and
535 pathway analyses. At 66 of the 535 novel loci we identified 97 non-synonymous SNPs,
536 including 8 predicted to be damaging (**Supplementary Table 8**).

537 We used chromatin interaction Hi-C data from endothelial cells (HUVEC)²³, neural
538 progenitor cells (NPC), mesenchymal stem cells (HVMSC) and tissue from the aorta (HAEC)
539 and adrenal gland²⁴ to identify distal associated genes. There were 498 novel loci that
540 contained a potential regulatory SNP and in 484 of these we identified long-range
541 interactions in at least one of the tissues or cell types. We found several potential long-
542 range target genes that do not overlap with the sentinel SNPs in the LD block. For example,
543 the *TGFB2* gene forms a 1.2Mb regulatory loop with SNPs in the *SLC30A10* locus, and the
544 *TGFBR1* promoter forms a 100kb loop with the *COL15A1* locus (**Supplementary Table 8**).

545 Our eQTL analysis identified 60 novel loci with eQTLs in arterial and 20 in adrenal tissue
546 (**Supplementary Table 9**), substantially increasing those identified in our previously
547 published GWAS on ~140K UKB individuals¹⁰. An example is SNP rs31120122 which defines
548 an aortic eQTL affecting expression of the *MED8* gene within the *SZT2* locus. In combination
549 with Hi-C interaction data in MSC, this supports a role for *MED8* in BP regulation, possibly
550 mediated through repression of smooth muscle cell differentiation. Hi-C interactions
551 provide supportive evidence for involvement of a further 36 arterial eGenes (genes whose
552 expression is affected by the eQTLs) that were distal to their eQTLs (e.g *PPHLN1*, *ERAP2*,
553 *FLRT2*, *ACVR2A*, *POU4F1*).

554 Using DeepSEA we found 198 SNPs in 121 novel loci with predicted effects on transcription
555 factor binding or on chromatin marks in tissues relevant for BP biology, such as vascular
556 tissue, smooth muscle and the kidney (**Supplementary Table 8**).

557 We used our genome-wide data at a false discovery rate (FDR) < 1% to robustly assess tissue
558 enrichment of BP loci using DEPICT and identified enrichment across 50 tissues and cells.
559 (**Supplementary Fig 5a; Supplementary Table 10a**). Enrichment was greatest for the
560 cardiovascular system especially blood vessels ($P = 1.5 \times 10^{-11}$) and the heart ($P = 2.7 \times 10^{-5}$).
561 Enrichment was high in adrenal tissue ($P = 3.7 \times 10^{-4}$) and, for the first time, we observed
562 high enrichment in adipose tissues ($P = 9.8 \times 10^{-9}$) corroborated by eQTL enrichment
563 analysis ($P < 0.05$) (**Supplementary Fig. 6; Supplementary Table 9**). Evaluation of enriched
564 mouse knockout phenotype terms also points to the importance of vascular morphology (P
565 $= 6 \times 10^{-15}$) and development ($P = 2.1 \times 10^{-18}$) in BP. With addition of our novel BP loci, we
566 identified new findings from both the gene ontology and protein-protein interaction
567 subnetwork enrichments, which highlight the TGF β ($P = 2.3 \times 10^{-13}$) and related SMAD
568 pathways ($P = 7 \times 10^{-15}$) (**Supplementary Table 10b, Supplementary Fig. 5b-d**).

569 We used FORGE²⁵ to investigate the regulatory regions for cell type specificity from DNase I
570 hypersensitivity sites, which showed strongest enrichment ($P < 0.001$) in the vasculature
571 and highly vascularised tissues, as reported in previous BP genetic studies¹⁰ (**Supplementary**
572 **Fig. 7**).

573 **Potential therapeutic targets**

574 Ingenuity pathway analysis and upstream regulator assessment showed enrichment of
575 canonical pathways implicated in cardiovascular disease including pathways targeted by
576 antihypertensive drugs (e.g. nitric oxide signalling) and also suggested some potential new
577 targets, such as relaxin signalling. Notably, upstream regulator analysis identified several BP
578 therapeutic targets such as angiotensinogen, calcium channels, progesterone, natriuretic
579 peptide receptor, angiotensin converting enzyme, angiotensin receptors and endothelin
580 receptors (**Supplementary Fig. 8**).

581 We developed a cumulative tally of functional evidence at each variant to assist in
582 variant/gene prioritisation at each locus and present a summary of the vascular expressed
583 genes contained within the 535 novel loci, including a review of their potential druggability
584 (**Supplementary Fig. 9**). The overlap between BP-associated genes and those associated
585 with antihypertensive drug targets further demonstrates new genetic support for known
586 drug mechanisms. For example, we report five novel BP associations with targets of five
587 antihypertensive drug classes (**Supplementary Table 11**), including the *PKD2L1*, *SLC12A2*,
588 *CACNA1C*, *CACNB4* and *CA7* loci - targeted by potassium-sparing diuretics (amiloride), loop
589 diuretics (bumetanide and furosemide), dihydropyridine, calcium channel blockers, non-
590 dihydropyridines and thiazide-like diuretics (chlortalidone) respectively. Notably in all but
591 the last case, functional variants in these genes are the best candidates in each locus.

592 **Concordance of BP variants and lifestyle exposures**

593 We examined association of sentinel SNPs at the 901 BP loci with BP-associated lifestyle
594 traits¹⁴ in UKB using either the Stanford Global Biobank Engine (N=327,302) or Gene ATLAS
595 (N=408,455). With corrected $P < 1 \times 10^{-6}$, we found genetic associations of BP variants with
596 daily fruit intake, urinary sodium and creatinine concentration, body mass index (BMI),
597 weight, waist circumference, and intakes of water, caffeine and tea ($P = 1.0 \times 10^{-7}$ to $P = 1.3$
598 $\times 10^{-46}$). Specifically, SNP rs13107325 in *SLC39A8* is a novel locus for frequency of drinking
599 alcohol ($P = 3.5 \times 10^{-15}$) and time spent watching TV ($P = 2.3 \times 10^{-11}$) as well as being
600 associated with BMI ($P = 1.6 \times 10^{-33}$), weight ($P = 8.8 \times 10^{-16}$) and waist circumference ($P =$
601 4.7×10^{-11}) (**Supplementary Table 12**). We used unsupervised hierarchical clustering for the
602 36 BP loci that showed at least one association at $P < 1 \times 10^{-6}$ with the lifestyle-related traits
603 in UKB (**Fig. 4**). The heatmap summarises the locus-specific associations across traits and
604 highlights heterogeneous effects with anthropometric traits across the loci examined. For
605 example, it shows clusters of associations between BP-raising alleles and either increased or
606 decreased adult height and weight. We note that some observed cross-trait associations are
607 in counter-directions to those expected epidemiologically.

608 **Association lookups with other traits and diseases**

609 We further evaluated cross-trait and disease associations using GWAS catalog²⁶,
610 PhenoScanner²⁷ and DisGeNET^{28,29}. The GWAS catalog and PhenoScanner search of
611 published GWAS showed that 77 of our 535 novel loci (using sentinel SNPs or proxies; $r^2 \geq$
612 0.8) are also significantly associated with other traits and diseases (**Fig. 5, Supplementary**
613 **Table 13**). We identified *APOE* as a highly cross-related BP locus showing associations with
614 lipid levels, cardiovascular-related outcomes and Alzheimer's disease, highlighting a
615 common link between cardiovascular risk and cognitive decline (**Fig. 5**). Other loci overlap
616 with anthropometric traits, including BMI, birth weight and height (**Fig. 5**) and with
617 DisGeNET terms related to lipid measurements, cardiovascular outcomes and obesity (**Fig.**
618 **6**).

619 We did lookups of our sentinel SNPs in ¹H NMR lipidomics data on plasma (N=2,022) and
620 data from the Metabolon platform (N=1,941) in the Airwave Study³⁰, and used
621 PhenoScanner to test SNPs against published significant ($P < 5 \times 10^{-8}$) genome vs
622 metabolome-wide associations in plasma and urine (Online Methods). Ten BP SNPs show
623 association with lipid particle metabolites and a further 31 SNPs (8 also on PhenoScanner)
624 show association with metabolites on the Metabolon platform, highlighting lipid pathways,
625 amino acids (glycine, serine, glutamine), tri-carboxylic acid cycle intermediates
626 (succinylcarnitine) and drug metabolites (**Supplementary Tables 14 and 15**). These findings
627 suggest a close metabolic coupling of BP regulation with lipid and energy metabolism.

628 **Genetic risk of increased blood pressure, hypertension and cardiovascular disease**

629 A weighted GRS for BP levels across all 901 loci was associated with a 10.4 mmHg higher,
630 sex-adjusted mean SBP in UK Biobank comparing the upper and lower quintiles of the GRS
631 distribution (95% CI: 10.2 to 10.6 mm Hg, $P < 1 \times 10^{-300}$) and with 12.9 mmHg difference in
632 SBP (95% CI: 12.6 to 13.1, $P < 1 \times 10^{-300}$) comparing the upper and lower deciles (**Fig. 7a,**
633 **Supplementary Table 16**). In addition, we observed over three-fold sex-adjusted higher risk
634 of hypertension (OR 3.34; 95% CI: 3.24 to 3.45; $P < 1 \times 10^{-300}$) between the upper and lower
635 deciles of the GRS in UK Biobank (**Fig. 7a**). Sensitivity analyses in the independent Airwave
636 cohort gave similar results (**Supplementary Table 17**).

637 We also show that the GRS is associated with increased, sex-adjusted risk of incident stroke,
638 myocardial infarction and all incident cardiovascular outcomes, comparing upper and lower
639 deciles of the GRS distribution, with odds ratios of 1.47 (95% CI: 1.35 to 1.59, $P = 1.1 \times 10^{-20}$),
640 1.50 (95% CI: 1.28 to 1.76, $P = 8.0 \times 10^{-7}$) and 1.52 (95% CI: 1.26 to 1.82, $P = 7.7 \times 10^{-6}$)
641 respectively (**Fig. 7b, Supplementary Table 16**).

642 **Extending analyses to other ancestries**

643 We examined associations with BP of both individual SNPs and the GRS among unrelated
644 individuals of African and South Asian descent in UKB, for the 901 known and novel loci.
645 Compared to Europeans, 62.4%, 62.5% and 64.8% of the variants among Africans (N=7,782),
646 and 74.2%, 72.3% and 75% South Asians (N=10,323) have concordant direction of effect for
647 SBP, DBP and PP respectively (**Supplementary Table 18; Supplementary Fig. 10**). Pearson
648 correlation coefficients with effect estimates in Europeans were $r^2 = 0.37$ and 0.78 for
649 Africans and South Asians respectively (**Supplementary Fig. 11**). We then applied the

650 European-derived GRS findings to unrelated Africans (N=6,970) and South Asians (N=8,827).
651 BP variants in combination were associated with 6.1 mmHg (95% CI: 4.5 to 7.7; $P = 4.9 \times 10^{-14}$)
652 and 7.4 mmHg (95% CI: 6.0 to 8.7; $P = 1.7 \times 10^{-26}$) higher, sex-adjusted mean systolic
653 pressure among Africans and South Asians, respectively, comparing upper and lower
654 quintiles of the GRS distribution (**Supplementary Tables 19a and 19b**).

655 **DISCUSSION**

656 Our study of over 1 million people offers an important step forward in understanding the
657 genetic architecture of BP. We identified over 1,000 independent signals at 901 loci for BP
658 traits, and the 535 novel loci more than triples the number of BP loci and doubles the
659 percentage variance explained, illustrating the benefits of large-scale biobanks. By
660 explaining 27% of the estimated heritability for BP, we make major inroads into the missing
661 heritability influencing BP level in the population³¹. The novel loci open the vista of entirely
662 new biology and highlight gene regions in systems not previously implicated in BP
663 regulation. This is particularly timely as global prevalence of people with SBP over 110-115
664 mm Hg, above which cardiovascular risk increases in a continuous graded manner, now
665 exceeds 3.5 billion, of whom over 1 billion are within the treatment range^{32,33}.

666 Our functional analysis highlights the role of the vasculature and associated pathways in the
667 genetics underpinning BP traits. We show a role for several loci in the transforming growth
668 factor beta (TGF β) pathway including SMAD family genes and the *TGF β* gene locus itself.
669 This pathway affects sodium handling in the kidney, ventricular remodelling, while plasma
670 levels of TGF β have recently been correlated with hypertension (**Fig. 8**)^{34,35}. The activin A
671 receptor type 1C (*ACVR1C*) gene mediates the effects of the TGF β family of signalling
672 molecules. A BP locus contains the Bone Morphogenetic Protein 2 (*BMP2*) gene in the TGF β
673 pathway, which prevents growth suppression in pulmonary arterial smooth muscle cells and
674 is associated with pulmonary hypertension³⁶. Another BP locus includes the Kruppel-like
675 family 14 (*KLF14*) gene of transcription factors, induced by low levels of TGF β receptor II
676 gene expression, and which has also been associated with type 2 diabetes,
677 hypercholesterolaemia and atherosclerosis³⁷.

678 Our analysis shows enrichment of BP gene expression in the adrenal tissue. Autonomous
679 aldosterone production by the adrenal glands is thought to be responsible for 5-10% of all
680 hypertension, rising to ~20% amongst people with resistant hypertension³⁸. Some of our
681 novel loci are linked functionally to aldosterone secretion^{39,40}. For example, the *CTNNA1*
682 locus encodes β -catenin, the central molecule in the canonical Wnt signalling system,
683 required for normal adrenocortical development^{41,42}. Somatic adrenal mutations of this
684 gene that prevent serine/threonine phosphorylation lead to hypertension through
685 generation of aldosterone-producing adenomas^{43,44}.

686 Our novel loci also include genes involved in vascular remodelling, such as vascular
687 endothelial growth factor A (*VEGFA*), the gene product of which induces proliferation,
688 migration of vascular endothelial cells and stimulates angiogenesis. Disruption of this gene
689 in mice resulted in abnormal embryonic blood vessel formation, while allelic variants of this
690 gene have been associated with microvascular complications of diabetes, atherosclerosis
691 and the antihypertensive response to enalapril⁴⁵. We previously reported a fibroblast

692 growth factor (*FGF5*) gene locus in association with BP¹⁰. Here, we additionally identify a
693 new BP locus encoding FGF9, which is linked to enhanced angiogenesis and vascular smooth
694 muscle cell differentiation by regulating *VEGFA* expression.

695 Several of our novel loci contain lipid-related genes consistent with the observed strong
696 associations among multiple cardio-metabolic traits. For example, the apolipoprotein E
697 gene (*APOE*) encodes the major apoprotein of the chylomicron. Recently, APOE serum levels
698 have been correlated with SBP in population-based studies and in murine knockout models;
699 disruption of this gene led to atherosclerosis and hypertension^{46,47}. A second novel BP locus
700 contains the low-density lipoprotein receptor-related protein 4 (*LRP4*) gene which may be a
701 target for APOE and is strongly expressed in the heart in mice and humans. In addition, we
702 identified a novel locus including the apolipoprotein L domain containing 1 gene (*APOLD1*)
703 that is highly expressed in the endothelium of developing tissues (particularly heart) during
704 angiogenesis.

705 Many of our novel BP loci encode proteins which may modulate vascular tone or signalling.
706 For example, the locus containing urotensin-2 receptor (*UTS2R*) gene encodes a class A
707 rhodopsin family G-protein coupled-receptor that upon activation by the neuropeptide
708 urotensin II, produces profound vasoconstriction. One novel locus for SBP contains the
709 relaxin gene, encoding a G-protein coupled receptor, with roles in vasorelaxation and
710 cardiac function; it signals by phosphatidylinositol 3-kinase (PI3K)^{48,49}, an enzyme which
711 inhibits vascular smooth muscle cell proliferation and neo-intimal formation⁵⁰. We identify
712 the *PI3K* gene here as a novel BP locus. We also identify the novel *RAMP2* locus which
713 encodes an adrenomedullin receptor⁵¹; we previously identified the adrenomedullin (*ADM*)
714 gene as a BP locus¹². Adrenomedullin is known to exert differential effects on BP in the brain
715 (vasopressor) and the vasculature (vasodilator). In addition, a locus containing Rho guanine
716 nucleotide exchange factor 25 (*ARHGEF25*) gene generates a factor that interacts with Rho
717 GTPases involved in contraction of vascular smooth muscle and regulation of responses to
718 angiotensin II⁵².

719 We evaluated the 901 BP loci for extant or potentially druggable targets. Loci encoding
720 *MARK3*, *PDGFC*, *TRHR*, *ADORA1*, *GABRA2*, *VEGFA* and *PDE3A* are within systems with
721 existing drugs not currently linked to a known antihypertensive mechanism; they may offer
722 repurposing opportunities e.g. detection of *SLC5A1* as the strongest repurposing candidate
723 in a new BP locus targeted by the type-2 diabetes drug canagliflozin. This is important as
724 between 8-12% of patients with hypertension exhibit resistance or intolerance to current
725 therapies and repositioning of a therapy with a known safety profile may reduce
726 development costs.

727 This study strengthens our previously reported GRS analysis indicating that all BP elevating
728 alleles combined could increase systolic BP by 10 mm Hg or more across quintiles or deciles
729 of the population distribution, substantially increasing risk of cardiovascular events¹⁰. We
730 previously suggested that genotyping BP elevating variants in the young may lead to
731 targeted lifestyle intervention in early life that might attenuate the BP rise at older ages¹⁰.

732 We identified several BP-associated loci that are also associated with lifestyle traits,
733 suggesting shared genetic architecture between BP and lifestyle exposures⁵³. We adjusted
734 our BP GWAS analyses for BMI to control for possible confounding effects, though we
735 acknowledge the potential for collider bias⁵⁴. Nonetheless, our findings of possible genetic
736 overlap between loci associated with BP and lifestyle exposures could support renewed
737 focus on altering specific lifestyle measures known to affect BP⁵⁵.

738 Despite smaller sample sizes, we observed high concordance with direction of effects on BP
739 traits of BP variants in Africans (> 62%) and South Asians (> 72%). The GRS analyses show
740 that, in combination, BP variants identified in European analyses are associated with BP in
741 non-European ancestries, though effect sizes were 30-40% smaller.

742 Our use of a two- and one-stage GWAS design illustrates the value of this approach to
743 minimize the effects of stochastic variation and heterogeneity. The one-stage approach
744 included signals that had independent and concordant support ($P < 0.01$) from both UKB
745 and ICBP, reducing the impact of winners' curse on our findings. Indeed, all but two of the
746 210 SNPs discovered in the one-stage analysis reach $P < 5 \times 10^{-6}$ in either UKB or ICBP. To
747 further minimize the risk of reporting false positive loci within our one-stage design, we set
748 a stringent overall discovery meta-analysis P -value threshold of $P < 5 \times 10^{-9}$, an order of
749 magnitude smaller than a genome-wide significance P -value, in line with thresholds
750 recommended for whole genome sequencing²². We found high concordance in direction of
751 effects between discovery data in the one-stage approach and the replication resources,
752 with similar distributions of effect sizes for the two approaches. We note that 24 of the
753 one-stage SNPs which reached $P < 5 \times 10^{-9}$ in discovery failed to reach genome-wide
754 significance ($P < 5 \times 10^{-8}$) in the combined meta-analysis of discovery and replication
755 resources, and hence may still require further validation in future, larger studies.

756 The new discoveries reported here more than triple the number of loci for BP to a total of
757 901 and represent a substantial advance in understanding the genetic architecture of BP.
758 The identification of many novel genes across the genome, could partly support an
759 omnigenic model for complex traits where genome-wide association of multiple
760 interconnected pathways is observed. However, our strong tissue enrichment shows
761 particular relevance to the biology of BP and cardiovascular disease⁵⁶, suggesting trait-
762 specificity, which could argue against an omnigenic model. Our confirmation of the impact
763 of these variants on BP level and cardiovascular events, coupled with identification of
764 shared risk variants for BP and adverse lifestyle could contribute to an early life precision
765 medicine strategy for cardiovascular disease prevention.

766 **URLs**

767 FORGE: http://browser.1000genomes.org/Homo_sapiens/UserData/Forge?db=core
768 Fantom5 data: <http://fantom.gsc.riken.jp/5/>
769 ENCODE DNase I data: (wgEncodeAwgDnaseMasterSites; accessed using Table browser)
770 ENCODE cell type data: <http://genome.ucsc.edu/ENCODE/cellTypes.html>.
771 GTEx: www.gtexportal.org
772 DeepSEA: <http://deepsea.princeton.edu/>
773 WebGetstalt: <http://www.webgestalt.org>

774 IPA: www.qiagen.com/ingenuity
775 Mouse Genome Informatics (MGI): <http://www.informatics.jax.org/batch>
776 Drug Gene Interaction database: www.dgidb.org
777 PhenoScanner: <http://www.phenoscanter.medschl.cam.ac.uk> (Phenoscanter integrates
778 results from the GWAS catalogue: <https://www.ebi.ac.uk/gwas/> and GRASP:
779 <https://grasp.nhlbi.nih.gov/>)
780 DisGeNET: <http://www.disgenet.org>
781 GeneAtlas: <http://geneatlas.roslin.ed.ac.uk>
782 Global Biobank Engine: <https://biobankengine.stanford.edu>
783

784 **Acknowledgements**

785
786 H.R.W. was funded by the National Institute for Health Research (NIHR) as part of the
787 portfolio of translational research of the NIHR Biomedical Research Centre at Barts and The
788 London School of Medicine and Dentistry. D.M-A is supported by the Medical Research
789 Council [grant number MR/L01632X.1]. B.M. holds an MRC eMedLab Medical Bioinformatics
790 Career Development Fellowship, funded from award MR/L016311/1. H.G. was funded by
791 the NIHR Imperial College Health Care NHS Trust and Imperial College London Biomedical
792 Research Centre. C.P.C. was funded by the National Institute for Health Research (NIHR) as
793 part of the portfolio of translational research of the NIHR Biomedical Research Center at
794 Barts and The London School of Medicine and Dentistry. S.T. was supported by Canadian
795 Institutes of Health Research; Université Laval (Quebec City, Canada). G.P. was supported by
796 Canada Research Chair in Genetic and Molecular Epidemiology and CISCO Professorship in
797 Integrated Health Biosystems. I.K. was supported by the EU PhenoMeNal project (Horizon
798 2020, 654241). C.P.K. is supported by grant U01DK102163 from the NIH-NIDDK, and by
799 resources from the Memphis VA Medical Center. C.P.K. is an employee of the US
800 Department of Veterans affairs. Opinions expressed in this paper are those of the authors'
801 and do not necessarily represent the opinion of the Department of Veterans Affairs. S.D.
802 was supported for this work by grants from the European Research Council (ERC), the EU
803 Joint Programme - Neurodegenerative Disease Research (JPND), the Agence Nationale de la
804 Recherche (ANR). T.B., J.MART., V.V., A.F.W. and C.H. were supported by a core MRC grant
805 to the MRCHGU QTL in Health and Disease research programme. M.BOE is supported by NIH
806 grant R01-DK062370. H.W. and A.G. acknowledge support of the Tripartite
807 Immunometabolism Consortium [TriC], Novo Nordisk Foundation (grant NNF15CC0018486).
808 N.V. was supported by Marie Skłodowska-Curie GF grant (661395) and ICIN-NHI. C.M. is
809 funded by the MRC AimHy (MR/M016560/1) project grant. M.A.N participation is supported
810 by a consulting contract between Data Tecnica International and the National Institute on
811 Aging, NIH, Bethesda, MD, USA. M.BR., M.CO., I.G., P.G., G.G, A.MO., A.R., D.V., C.M.B.,
812 C.F.S., M.T., D.T. were supported by the Italian Ministry of Health RF2010 to Paolo
813 Gasparini, RC2008 to Paolo Gasparini. D.I.B. is supported by the Royal Netherlands
814 Academy of Science Professor Award (PAH/6635). J.C.C. is supported by the Singapore
815 Ministry of Health's National Medical Research Council under its Singapore Translational
816 Research Investigator (STaR) Award (NMRC/STaR/0028/2017). C.C., P.B.M and M.R.B were
817 funded by the National Institutes for Health Research (NIHR) as part of the portfolio of
818 translational research of the NIHR Biomedical Research Centre at Barts. T.F. is supported by
819 the NIHR Biomedical Research Centre, Oxford. M.R. is recipient from China Scholarship
820 Council (No. 2011632047). C.L. was supported by the Medical Research Council UK

821 (G1000143; MC_UU_12015/1; MC_PC_13048; MC_U106179471), Cancer Research UK
822 (C864/A14136), EU FP6 programme (LSHM_CT_2006_037197). G.B.E is supported by the
823 Swiss National Foundation SPUM project FN 33CM30-124087, Geneva University, and the
824 Fondation pour Recherches Médicales, Genève. C.M.L is supported by the Li Ka Shing
825 Foundation, WT-SSI/John Fell funds and by the NIHR Biomedical Research Centre, Oxford,
826 by Widenlife and NIH (CRR00070 CR00.01). R.J.F.L. is supported by the NIH (R01DK110113,
827 U01HG007417, R01DK101855, R01DK107786). D.O.M-K. is supported by Dutch Science
828 Organization (ZonMW-VENI Grant 916.14.023). M.M was supported by the National
829 Institute for Health Research (NIHR) BioResource Clinical Research Facility and Biomedical
830 Research Centre based at Guy's and St Thomas' NHS Foundation Trust and King's College
831 London. H.W. and M.F. acknowledge the support of the Wellcome Trust core award
832 (090532/Z/09/Z) and the BHF Centre of Research Excellence (RE/13/1/30181). A.G, H.W.
833 acknowledge European Union Seventh Framework Programme FP7/2007-2013 under grant
834 agreement no. HEALTH-F2-2013-601456 (CVGenes@Target) & and A.G, the Wellcome Trust
835 Institutional strategic support fund. L.R. was supported by Forschungs- und Förder-Stiftung
836 INOVA, Vaduz, Liechtenstein. M.TO. is supported by British Heart Foundation
837 (PG/17/35/33001). P.S. is recipient of an NIHR Senior Investigator Award and is supported
838 by the Biomedical Research Centre Award to Imperial College Healthcare NHS Trust. P.v.d.H.
839 was supported by ICIN-NHI and Marie Skłodowska-Curie GF (call: H2020-MSCA-IF-2014,
840 Project ID: 661395). N.J.W. was supported by the Medical Research Council UK (G1000143;
841 MC_UU_12015/1; MC_PC_13048; MC_U106179471), Cancer Research UK (C864/A14136),
842 EU FP6 programme (LSHM_CT_2006_037197). E.Z. was supported by the Wellcome Trust
843 (WT098051). J.N.H. was supported by the Vanderbilt Molecular and Genetic Epidemiology
844 of Cancer (MAGEC) training program, funded by T32CA160056 (PI: X.-O. Shu) and by VA
845 grant 1I01CX000982. A.G. was supported by VA grant 1I01CX000982. T.L.E. and D.R.V.E.
846 were supported by grant R21HL121429 from NIH/NHLBI. A.M.H. was supported by VA
847 Award #I01BX003360. C.J.O. was supported by the VA Boston Healthcare, Section of
848 Cardiology and Department of Medicine, Brigham and Women's Hospital, Harvard Medical
849 School. The MRC/BHF Cardiovascular Epidemiology Unit is supported by the UK Medical
850 Research Council [MR/L003120/1]; British Heart Foundation [RG/13/13/30194]; and UK
851 National Institute for Health Research Cambridge Biomedical Research Centre. J.DA is a
852 British Heart Foundation Professor and NIHR Senior Investigator. L.V.W. holds a
853 GlaxoSmithKline/British Lung Foundation Chair in Respiratory Research. P.E. acknowledges
854 support from the NIHR Biomedical Research Centre at Imperial College Healthcare NHS
855 Trust and Imperial College London, the NIHR Health Protection Research Unit in Health
856 Impact of Environmental Hazards (HPRU-2012-10141), and the Medical Research Council
857 (MRC) and Public Health England (PHE) Centre for Environment and Health
858 (MR/L01341X/1). P.E. is a UK Dementia Research Institute (DRI) professor, UK DRI at
859 Imperial College London, funded by the MRC, Alzheimer's Society and Alzheimer's Research
860 UK. He is also associate director of Health Data Research-UK London funded by a
861 consortium led by the Medical Research Council. M.J.C. was funded by the National Institute
862 for Health Research (NIHR) as part of the portfolio of translational research of the NIHR
863 Biomedical Research Center at Barts and The London School of Medicine and Dentistry.
864 M.J.C. is a National Institute for Health Research (NIHR) senior investigator and this work is
865 funded by the MRC eMedLab award to M.J.C. and M.R.B. and the NIHR Biomedical Research
866 Centre at Barts.

867 This research has been conducted using the UK Biobank Resource under Application
868 Numbers 236 and 10035. This research was supported by the British Heart Foundation
869 (grant SP/13/2/30111). Large-scale comprehensive genotyping of UK Biobank for
870 cardiometabolic traits and diseases: UK CardioMetabolic Consortium (UKCMC).

871 Computing: This work was enabled using the computing resources of the i) UK MEDical
872 BIOinformatics partnership - aggregation, integration, visualisation and analysis of large,
873 complex data (UK MED-BIO) which is supported by the Medical Research Council [grant
874 number MR/L01632X/1] and ii) the MRC eMedLab Medical Bioinformatics Infrastructure,
875 supported by the Medical Research Council [grant number MR/L016311/1].

876 **Author contributions**

877

878 **Central analysis:** E.E., H.R.W., D.M-A., B.M., R.P., H.G., G.N., N.D., C.P.C., I.K., F.N., M.E.,
879 K.W., E.T. L.V.W.

880 **Writing of the manuscript:** E.E., H.R.W., D.M-A., B.M., R.P., H.G., I.T., M.R.B., L.V.W., P.E.,
881 M.J.C. (with group leads EE, H.R.W, L.V.W., P.E., M.J.C.)

882 **ICBP-Discovery contributor:** (3C-Dijon) S.D., M.S., P.A.M., G.C., C.T.; (AGES-Reykjavik) V.GU.,
883 L.J.L., A.V.S., T.B.H.; (ARIC) D.E.A., E.B., A.CH. A.C.M., P.N.; (ASCOT) N.R.P., D.C.S., A.S.,
884 S.THO., P.B.M., P.S., M.J.C., H.R.W.; (ASPS) E.H., Y.S., R.S., H.S.; (B58C) D.P.S., BHSA.J.,
885 N.SHR.; (BioMe (formerly IPM)) E.P.B., Y.LU., R.J.F.L.; (BRIGHT) J.C., M.F., M.J.B., P.B.M.,
886 M.J.C., H.R.W. ; (CHS) J.C.B., K.R., K.D.T., B.M.P.; (Cilento study) M.C., T.NU., D.R., R.SO.;
887 (COLAUS) M.B., Z.K., P.V.; (CROATIA_Korcula) J.MART., A.F.W.; (CROATIA_SPLIT) I.KO., O.P.,
888 T.Z.; (CROATIA_Vis) J.E.H., I.R., V.V.; (EPIC) K-T.K., R.J.F.L., N.J.W.; (EPIC-CVD) W-Y.L., P.SU.,
889 A.S.B., J.DA., J.M.M.H.; (EPIC-Norfolk, Fenland-OMICS, Fenland-GWAS) J-H.Z.; (EPIC-Norfolk,
890 Fenland-OMICS, Fenland-GWAS, InterAct-GWAS) J.L., C.L., R.A.S., N.J.W.; (ERF) N.A., B.A.O.,
891 C.M.v.D.; (Fenland-Exome, EPIC-Norfolk-Exome) S.M.W., FHSS-J.H., D.L.; (FINRISK
892 (COROGENE_CTRL)) P.J., K.K., M.P., A-P.S.; (FINRISK_PREDICT_CVD) A.S.H., A.P., S.R., V.S.;
893 (FUSION) A.U.J, M.BOE., F.C., J.T., (GAPP) S.T., G.P., D.CO., L.R.; (Generation Scotland
894 (GS:SFHS)) T.B., C.H., A.C., S.P.; (GoDARTs) N.S., A.S.F.D., A.D.M., C.N.A.P.; (GRAPHIC) P.S.B.,
895 C.P.N., N.J.SA., M.D.T.; (H2000_CTRL) A.JU., P.K., S.KO., T.N.; (HABC) Y.L., M.A.N., T.B.H.;
896 (HCS) J.R.A., E.G.H., C.O., R.J.SC.; (HTO) K.L.A., H.J.C., B.D.K., M.TO, C.MA.; (ICBP-SC) G.A.,
897 T.F., M-R.J., A.D.J., M.LA., C.N.; (INGI-CARL) I.G., G.G., A.MO., A.R.; (INGI-FVG) M.BR., M.CO.,
898 P.G., D.V.; (INGI-VB) C.M.B., C.F.S., D.T., M.T.; (JUPITER) F.G., L.M.R., P.M.R., D.I.C.; (KORA
899 S3) C.G., M.L., E.O., S.S.; (KORA S4) A.PE., J.S.R.; (LBC1921) S.E.H., D.C.M.L., A.PA., J.M.S.;
900 (LBC1936) G.D., I.J.D., A.J.G., L.M.L.; (Lifelines) N.V., M.H.d.B., M.A.S., P.v.d.H.; (LOLIPOP)
901 J.C.C., J.S.K., B.L., W.Z.; (MDC) P.A., O.M.; (MESA) X.G., W.P., J.I.R., J.Y.; (METSIM) A.U.J.,
902 M.LAA.; (MICROS) F.D.G.M. , A.A.H., P.P.P.; (MIGEN) R.E., S.K., J.M., D.SI.; (NEO) R.L., R.d.M.,
903 R.N., D.O.M-K.; (NESDA) Y.M., I.M.N., B.W.J.H.P., H.SN.; (NSPHS) S.E., U.G., Å.JO.; (NTR)
904 D.I.B., E.J.d.G., J-J.H., G.W.; (ORCADES) H.C., P.K.J., S.H.W., J.F.W.; (PIVUS) L.LI., C.M.L., J.S.,
905 A.M.; (Prevend) N.V., P.v.d.H.; (PROCARDIS) M.F., A.G., H.W.; (PROSPER) J.DE., J.W.J., D.J.S.,
906 S.TR.; (RS) O.H.F., A.HO., A.U., G.C.V.; (SardinIA) J.D., Y.Q., F.CU., E.G.L.; (SHIP) M.D., R.R.,
907 A.T., U.V.; (STR) M.FR., A.H., R.J.S., E.I.; (TRAILS) C.A.H., A.J.O., H.R., P.J.v.d.M.; (TwinsUK)

908 M.M., C.M., T.D.S.; (UKHLS) B.P.P., E.Z.; (ULSAM) V.G., A.P.M., A.M., E.I.; (WGHS) F.G.,
909 L.M.R., P.M.R., D.I.C.; (YFS) M.K., T.L., L-P.L., O.T.R.

910 **Replication study contributor:** (MVP) J.N.H., A.G., D.R.V.E., Y.V.S., K.C., J.M.G., P.W.F.W.,
911 P.S.T., C.P.K., A.M.H., C.J.O., T.L.E.; (EGCUT) T.E., R.M., L.M. A.ME.

912 **Airwave Health Monitoring Study:** E.E, H.G, A-C.V., R.P., I.K., I.T., P.E.

913 **All authors critically reviewed and approved the final version of the manuscript**

914 **Conflicts/Disclosures**

915

916 K.W. is a Commercial partnerships manager for Genomics England, a UK Government
917 Company

918 M.A.N. consults for Illumina Inc, the Michael J. Fox Foundation and University of California
919 Healthcare among others.

920 A.S.B. has received grants outside of this work from Merck, Pfizer, Novartis, AstraZeneca,
921 Biogen and Bioverativ and personal fees from Novartis

922 J.DA. has the following competing interests: Pfizer Population Research Advisory Panel
923 (grant), AstraZeneca (grant), Wellcome Trust (grant), UK Medical Research Council (grant),
924 Pfizer(grant), Novartis (grant), NHS Blood and Transplant(grant), National Institute of Health
925 Research(grant), UK MEDICAL RESEARCH COUNCIL(grant), BRITISH HEART
926 FOUNDATION(grant),UK NATIONAL INSTITUTE OF HEALTH RESEARCH (grant), EUROPEAN
927 COMMISSION (grant), Merck Sharp and Dohme UK Atherosclerosis (personal fees), Novartis
928 Cardiovascular and Metabolic Advisory Board (personal fees), British Heart Foundation
929 (grant), European Research Council (grant), Merck (grant).

930 B.M.P. serves on the DSMB of a clinical trial funded by Zoll LifeCor and on the Steering
931 Committee of the Yale Open Data Access Project funded by Johnson & Johnson.

932 M.J.C. is Chief Scientist for Genomics England, a UK Government company.

933

934 The views expressed in this manuscript are those of the authors and do not necessarily
935 represent the views of the National Heart, Lung, and Blood Institute; the National Institutes
936 of Health; or the U. S. Department of Health and Human Services. This publication does not
937 represent the views of the Department of Veterans Affairs or the United States Government.
938

939 **References**

940

- 941 1. Forouzanfar, M.H. *et al.* Global Burden of Hypertension and Systolic Blood Pressure
942 of at Least 110 to 115 mm Hg, 1990-2015. *JAMA* **317**, 165-182 (2017).
- 943 2. Munoz, M. *et al.* Evaluating the contribution of genetics and familial shared
944 environment to common disease using the UK Biobank. *Nat Genet* **48**, 980-3 (2016).
- 945 3. Poulter, N.R., Prabhakaran, D. & Caulfield, M. Hypertension. *Lancet* **386**, 801-12
946 (2015).

- 947 4. Feinleib, M. *et al.* The NHLBI twin study of cardiovascular disease risk factors:
948 methodology and summary of results. *Am J Epidemiol* **106**, 284-5 (1977).
- 949 5. Cabrera, C.P. *et al.* Exploring hypertension genome-wide association studies findings
950 and impact on pathophysiology, pathways, and pharmacogenetics. *Wiley Interdiscip*
951 *Rev Syst Biol Med* **7**, 73-90 (2015).
- 952 6. Ehret, G.B. *et al.* The genetics of blood pressure regulation and its target organs from
953 association studies in 342,415 individuals. *Nat Genet* **48**, 1171-1184 (2016).
- 954 7. Surendran, P. *et al.* Trans-ancestry meta-analyses identify rare and common variants
955 associated with blood pressure and hypertension. *Nat Genet* **48**, 1151-1161 (2016).
- 956 8. Liu, C. *et al.* Meta-analysis identifies common and rare variants influencing blood
957 pressure and overlapping with metabolic trait loci. *Nat Genet* **48**, 1162-70 (2016).
- 958 9. Hoffmann, T.J. *et al.* Genome-wide association analyses using electronic health
959 records identify new loci influencing blood pressure variation. *Nat Genet* **49**, 54-64
960 (2017).
- 961 10. Warren, H.R. *et al.* Genome-wide association analysis identifies novel blood pressure
962 loci and offers biological insights into cardiovascular risk. *Nat Genet* **49**, 403-415
963 (2017).
- 964 11. Wain, L.V. *et al.* Novel Blood Pressure Locus and Gene Discovery Using Genome-
965 Wide Association Study and Expression Data Sets From Blood and the Kidney.
966 *Hypertension* (2017).
- 967 12. International Consortium for Blood Pressure Genome-Wide Association Studies *et al.*
968 Genetic variants in novel pathways influence blood pressure and cardiovascular
969 disease risk. *Nature* **478**, 103-9 (2011).
- 970 13. Sudlow, C. *et al.* UK biobank: an open access resource for identifying the causes of a
971 wide range of complex diseases of middle and old age. *PLoS Med* **12**, e1001779
972 (2015).
- 973 14. Gaziano, J.M. *et al.* Million Veteran Program: A mega-biobank to study genetic
974 influences on health and disease. *J Clin Epidemiol* **70**, 214-23 (2016).
- 975 15. Leitsalu, L. *et al.* Cohort Profile: Estonian Biobank of the Estonian Genome Center,
976 University of Tartu. *Int J Epidemiol* **44**, 1137-47 (2015).
- 977 16. McCarthy, S. *et al.* A reference panel of 64,976 haplotypes for genotype imputation.
978 *Nat Genet* **48**, 1279-83 (2016).
- 979 17. Loh, P.R. *et al.* Efficient Bayesian mixed-model analysis increases association power
980 in large cohorts. *Nat Genet* **47**, 284-90 (2015).
- 981 18. Bulik-Sullivan, B.K. *et al.* LD Score regression distinguishes confounding from
982 polygenicity in genome-wide association studies. *Nat Genet* **47**, 291-5 (2015).
- 983 19. Ioannidis, J.P., Patsopoulos, N.A. & Evangelou, E. Heterogeneity in meta-analyses of
984 genome-wide association investigations. *PLoS One* **2**, e841 (2007).
- 985 20. Evangelou, E. & Ioannidis, J.P. Meta-analysis methods for genome-wide association
986 studies and beyond. *Nat Rev Genet* **14**, 379-89 (2013).
- 987 21. Pulit, S.L., de With, S.A. & de Bakker, P.I. Resetting the bar: Statistical significance in
988 whole-genome sequencing-based association studies of global populations. *Genet*
989 *Epidemiol* **41**, 145-151 (2017).
- 990 22. Yang, J., Lee, S.H., Goddard, M.E. & Visscher, P.M. GCTA: a tool for genome-wide
991 complex trait analysis. *Am J Hum Genet* **88**, 76-82 (2011).
- 992 23. Rao, S.S. *et al.* A 3D map of the human genome at kilobase resolution reveals
993 principles of chromatin looping. *Cell* **159**, 1665-80 (2014).

- 994 24. Schmitt, A.D. *et al.* A Compendium of Chromatin Contact Maps Reveals Spatially
995 Active Regions in the Human Genome. *Cell Rep* **17**, 2042-2059 (2016).
- 996 25. Dunham, I.K., E.; Iotchkova, V.; Morganella, S.; Birney, E. FORGE: A tool to discover
997 cell specific enrichments of GWAS associated SNPs in regulatory regions.
998 *F1000Research* **4**(2015).
- 999 26. MacArthur, J. *et al.* The new NHGRI-EBI Catalog of published genome-wide
1000 association studies (GWAS Catalog). *Nucleic Acids Res* **45**, D896-D901 (2017).
- 1001 27. Staley, J.R. *et al.* PhenoScanner: a database of human genotype-phenotype
1002 associations. *Bioinformatics* **32**, 3207-3209 (2016).
- 1003 28. Pinero, J. *et al.* DisGeNET: a discovery platform for the dynamical exploration of
1004 human diseases and their genes. *Database (Oxford)* **2015**, bav028 (2015).
- 1005 29. Pinero, J. *et al.* DisGeNET: a comprehensive platform integrating information on
1006 human disease-associated genes and variants. *Nucleic Acids Res* **45**, D833-D839
1007 (2017).
- 1008 30. Elliott, P. *et al.* The Airwave Health Monitoring Study of police officers and staff in
1009 Great Britain: rationale, design and methods. *Environ Res* **134**, 280-5 (2014).
- 1010 31. Ehret, G.B. & Caulfield, M.J. Genes for blood pressure: an opportunity to understand
1011 hypertension. *Eur Heart J* **34**, 951-61 (2013).
- 1012 32. Blood Pressure Lowering Treatment Trialists, C. *et al.* Blood pressure-lowering
1013 treatment based on cardiovascular risk: a meta-analysis of individual patient data.
1014 *Lancet* **384**, 591-8 (2014).
- 1015 33. GBD 2015 Risk Factors Collaborators. Global, regional, and national comparative risk
1016 assessment of 79 behavioural, environmental and occupational, and metabolic risks
1017 or clusters of risks, 1990-2015: a systematic analysis for the Global Burden of Disease
1018 Study 2015. *Lancet* **388**, 1659-1724 (2016).
- 1019 34. Nakao, E. *et al.* Elevated Plasma Transforming Growth Factor beta1 Levels Predict
1020 the Development of Hypertension in Normotensives: The 14-Year Follow-Up Study.
1021 *Am J Hypertens* **30**, 808-814 (2017).
- 1022 35. Feng, W., Dell'Italia, L.J. & Sanders, P.W. Novel Paradigms of Salt and Hypertension. *J*
1023 *Am Soc Nephrol* **28**, 1362-1369 (2017).
- 1024 36. International PPH Consortium *et al.* Heterozygous germline mutations in BMPR2,
1025 encoding a TGF-beta receptor, cause familial primary pulmonary hypertension. *Nat*
1026 *Genet* **26**, 81-4 (2000).
- 1027 37. Voight, B.F. *et al.* Twelve type 2 diabetes susceptibility loci identified through large-
1028 scale association analysis. *Nat Genet* **42**, 579-89 (2010).
- 1029 38. Douma, S. *et al.* Prevalence of primary hyperaldosteronism in resistant hypertension:
1030 a retrospective observational study. *Lancet* **371**, 1921-6 (2008).
- 1031 39. Rossi, G.P. *et al.* A prospective study of the prevalence of primary aldosteronism in
1032 1,125 hypertensive patients. *J Am Coll Cardiol* **48**, 2293-300 (2006).
- 1033 40. Calhoun, D.A., Nishizaka, M.K., Zaman, M.A., Thakkar, R.B. & Weissmann, P.
1034 Hyperaldosteronism among black and white subjects with resistant hypertension.
1035 *Hypertension* **40**, 892-6 (2002).
- 1036 41. Drelon, C., Berthon, A., Mathieu, M., Martinez, A. & Val, P. Adrenal cortex tissue
1037 homeostasis and zonation: A WNT perspective. *Mol Cell Endocrinol* **408**, 156-64
1038 (2015).
- 1039 42. El Wakil, A. & Lalli, E. The Wnt/beta-catenin pathway in adrenocortical development
1040 and cancer. *Mol Cell Endocrinol* **332**, 32-7 (2011).

- 1041 43. Teo, A.E. *et al.* Pregnancy, Primary Aldosteronism, and Adrenal CTNNB1 Mutations. *N Engl J Med* **373**, 1429-36 (2015).
1042
1043 44. Tissier, F. *et al.* Mutations of beta-catenin in adrenocortical tumors: activation of the
1044 Wnt signaling pathway is a frequent event in both benign and malignant
1045 adrenocortical tumors. *Cancer Res* **65**, 7622-7 (2005).
1046 45. Oliveira-Paula, G.H. *et al.* Polymorphisms in VEGFA gene affect the antihypertensive
1047 responses to enalapril. *Eur J Clin Pharmacol* **71**, 949-57 (2015).
1048 46. Yang, R. *et al.* Hypertension and endothelial dysfunction in apolipoprotein E
1049 knockout mice. *Arterioscler Thromb Vasc Biol* **19**, 2762-8 (1999).
1050 47. Sofat, R. *et al.* Circulating Apolipoprotein E Concentration and Cardiovascular
1051 Disease Risk: Meta-analysis of Results from Three Studies. *PLoS Med* **13**, e1002146
1052 (2016).
1053 48. Conrad, K.P. Unveiling the vasodilatory actions and mechanisms of relaxin.
1054 *Hypertension* **56**, 2-9 (2010).
1055 49. Sun, H.J. *et al.* Relaxin in paraventricular nucleus contributes to sympathetic
1056 overdrive and hypertension via PI3K-Akt pathway. *Neuropharmacology* **103**, 247-56
1057 (2016).
1058 50. Miyamoto, Y. *et al.* Phosphatidylinositol 3-kinase inhibition induces vasodilator
1059 effect of sevoflurane via reduction of Rho kinase activity. *Life Sci* **177**, 20-26 (2017).
1060 51. Pawlak, J.B., Wetzel-Strong, S.E., Dunn, M.K. & Caron, K.M. Cardiovascular effects of
1061 exogenous adrenomedullin and CGRP in Ramp and Calcrl deficient mice. *Peptides* **88**,
1062 1-7 (2017).
1063 52. Ohtsu, H. *et al.* Signal-crosstalk between Rho/ROCK and c-Jun NH2-terminal kinase
1064 mediates migration of vascular smooth muscle cells stimulated by angiotensin II.
1065 *Arterioscler Thromb Vasc Biol* **25**, 1831-6 (2005).
1066 53. Tzoulaki, I., Elliott, P., Kontis, V. & Ezzati, M. Worldwide Exposures to Cardiovascular
1067 Risk Factors and Associated Health Effects: Current Knowledge and Data Gaps.
1068 *Circulation* **133**, 2314-33 (2016).
1069 54. Munafo, M.R., Tilling, K., Taylor, A.E., Evans, D.M. & Davey Smith, G. Collider scope:
1070 when selection bias can substantially influence observed associations. *Int J Epidemiol*
1071 **47**, 226-235 (2017).
1072 55. Pazoki, R. *et al.* Genetic predisposition to high blood pressure and lifestyle factors:
1073 Associations with midlife blood pressure levels and cardiovascular events. *Circulation*
1074 **137**, 653-661 (2018)
1075 56. Boyle, E.A., Li, Y.I. & Pritchard, J.K. An expanded view of complex traits. From
1076 polygenic to omnigenic. *Cell* **169**, 1177-1186 (2017)
1077

1078 **Figure Legends**

1079 **Figure 1. Study design schematic for discovery and validation of loci.** ICBP; International
1080 Consortium for Blood Pressure; N, sample size; QC, quality control; PCA, principal-component
1081 analysis; GWAS, Genome-wide Association Study; 1000G 1000 Genomes; HRC, Haplotype Reference
1082 Panel; BP: blood pressure; SNPs, single nucleotide polymorphisms; BMI, body mass index; LMM;
1083 linear mixed model; UKB, UK Biobank, MAF, minor allele frequency; HLA, Human Leukocyte Antigen;
1084 MVP, Million Veterans Program; EGCTU; Estonian Genome Center, University of Tartu; SBP, systolic
1085 blood pressure; DBP, diastolic blood pressure; PP, pulse pressure.

1086 **Figure 2. Manhattan plot showing the minimum P -value for the association across all blood**
1087 **pressure traits in the discovery stage excluding known and previously reported variants.**
1088 Manhattan plot of the discovery genome-wide association meta-analysis in 757,601 individuals
1089 excluding variants in 274 known loci. The minimum P -value, computed using inverse variance fixed
1090 effects meta-analysis, across SBP, DBP and PP is presented. The y axis shows the $-\log_{10} P$ values and
1091 the x axis shows their chromosomal positions. Horizontal red and blue line represents the thresholds
1092 of $P = 5 \times 10^{-8}$ for genome-wide significance and $P = 1 \times 10^{-6}$ for selecting SNPs for replication,
1093 respectively. SNPs in blue are in LD ($r^2 > 0.8$) with the 325 novel variants independently replicated
1094 from the 2-stage design whereas SNPs in red are in LD ($r^2 > 0.8$) with 210 SNPs identified through the
1095 1-stage design with internal replication. Any loci in black or grey that exceed the significance
1096 thresholds were significant in the discovery meta-analysis, but did not meet the criteria of
1097 replication in the one- or two-stage designs.

1098 **Figure 3: Venn Diagrams of Novel Loci Results (a) “Comparison of 1-stage and 2-stage design**
1099 **analysis criteria”:** For all 535 novel loci, we compare the results according to the association criteria
1100 used for the one-stage and the two-stage design. Two-hundred and ten loci exclusively met the one-
1101 stage analysis criteria ($P < 5 \times 10^{-9}$ in the discovery meta-analysis [N=757,601], $P < 0.01$ in UKB
1102 [N=458,577], $P < 0.01$ in ICBP [N=299,024] and concordant direction of effect between UKB and
1103 ICBP). The P -values for the discovery and the ICBP meta-analyses were calculated using inverse
1104 variance fixed effects meta-analysis. The P -values in UKB were derived from linear mixed modeling
1105 using BOLT-LMM. Of the 325 novel replicated loci from the 2-stage analysis (genome-wide
1106 significance in the combined meta-analysis, $P < 0.01$ in the replication meta-analysis and concordant
1107 direction of effect), 204 loci would also have met the one-stage criteria, whereas 121 were only
1108 identified by the two-stage analysis. **(b) “Overlap of Associations across Blood Pressure Traits”.**
1109 For all 535 novel loci, we show the number of loci associated with each blood pressure trait. We
1110 present the two-stage loci first, followed by the one-stage loci. SBP: systolic blood pressure; DBP:
1111 diastolic blood pressure; PP: pulse pressure; UKB: UK Biobank; ICBP: International Consortium of
1112 Blood Pressure.

1113 **Figure 4. Association of blood pressure loci with lifestyle traits.** Plot shows unsupervised
1114 hierarchical clustering of BP loci based on associations with lifestyle-related factors. For the sentinel
1115 SNP at each BP locus (x-axis), we calculated the $-\log_{10}(P) * \text{sign}(\beta)$ (aligned to BP-raising allele) as
1116 retrieved from the Gene Atlas catalogue (<http://geneatlas.roslin.ed.ac.uk>). The P -values in Gene
1117 Atlas were calculated applying linear mixed models. BP loci and traits were clustered according to
1118 the Euclidean distance amongst $-\log_{10}(P) * \text{sign}(\beta)$. Red squares indicate direct associations with the
1119 trait of interest and blue squares inverse associations. Only SNPs with at least one association at P
1120 $< 10^{-6}$ with at least one of the traits examined are annotated in the heat-map. All 901 loci are
1121 considered, both known and novel: novel loci are printed in bold font. SNPs: Single Nucleotide
1122 Polymorphisms; BP: Blood Pressure.

1123 **Figure 5. Association of blood pressure loci with other traits.** Plot shows results from associations
1124 with other traits which were extracted from the GWAS catalog and PhenoScanner databases for the
1125 535 novel sentinel SNPs including proxies in Linkage Disequilibrium ($r^2 \geq 0.8$) with genome-wide
1126 significant associations. SBP: Systolic Blood Pressure; DBP: Diastolic Blood Pressure; PP: Pulse
1127 Pressure; HR: Heart Rate; ECG: Electrocardiographic traits; CAD: Coronary Artery Disease CHD;
1128 Coronary Heart Disease MI; Myocardial Infraction; T2D: Type II Diabetes.

1129 **Figure 6. Association of blood pressure loci with other traits.** Plots (a) and (b) show overlap
1130 between variants associated to (a) traits and (b) diseases in the manually-curated version of the
1131 DisGeNET database, and all variants in LD $r^2 > 0.8$ with the known (red bars) SNPs from the 274

1132 published loci, and all (green bars) BP variants from all 901 loci. Numbers on top of the bars denote
1133 the number of SNPs included in DisGeNET for the specific trait or disease. Traits/diseases with an
1134 overlap of at least 5 variants in LD with all markers are shown. The Y axis shows the percentage of
1135 variants associated with the diseases that is covered by the overlap. For the sake of clarity, the
1136 DisGeNET terms for blood pressure and hypertension are not displayed, whereas the following
1137 diseases have been combined: coronary artery disease (CAD), coronary heart disease (CHD) and
1138 myocardial infarction (MI); prostate and breast carcinoma; Crohn's and inflammatory bowel
1139 diseases.

1140 **Figure 7. Relationship of deciles of the genetic risk score (GRS) based on all 901 loci with blood**
1141 **pressure, risk of hypertension and cardiovascular disease in UK Biobank.** The plots show sex-
1142 adjusted (a) mean systolic blood pressure (SBP) and odds ratios of hypertension (HTN) (N=364,520)
1143 and (b) odds ratios of incident cardiovascular disease (CVD), myocardial infarction (MI) and stroke
1144 (N=392,092), comparing each of the upper nine GRS deciles with the lowest decile; dotted lines
1145 represent the upper 95% confidence intervals.

1146 **Figure 8: Known and novel BP associations in the TGF β signalling pathway.** Genes with known
1147 associations with BP are indicated in cyan. Genes with novel associations with BP reported in this
1148 study are indicated in red. TGF β pathway was derived from an ingenuity canonical pathway. BP:
1149 Blood Pressure.

1150

1151 **ONLINE METHODS**

1152 **UK Biobank (UKB) data**

1153 We performed a Genome Wide Association Study (GWAS) analysis in 458,577 UKB
1154 participants¹³ (**Supplementary Methods**). These consist of 408,951 individuals from UKB
1155 genotyped at 825,927 variants with a custom Affymetrix UK Biobank Axiom Array chip and
1156 49,626 individuals genotyped at 807,411 variants with a custom Affymetrix UK BiLEVE
1157 Axiom Array chip from the UK BiLEVE study⁵⁷, which is a subset of UKB. SNPs were imputed
1158 centrally by UKB using a reference panel that merged the UK10K and 1000 Genomes Phase
1159 3 panel as well as the Haplotype Reference Consortium (HRC) panel⁵⁸. For current analysis
1160 only SNPs imputed from the HRC panel were considered.

1161 *UKB phenotypic data*

1162 Following Quality Control (QC) (**Supplementary Methods**), we restricted our data to a
1163 subset of post-QC individuals of European ancestry combining information from self-
1164 reported and genetic data (**Supplementary Methods**) resulting in a maximum of N=458,577
1165 individuals (**Fig. 1, Supplementary Fig. 12**).

1166 Three BP traits were analysed: systolic (SBP), diastolic (DBP) and pulse pressure (PP)
1167 (difference between SBP and DBP). We calculated the mean SBP and DBP values from two
1168 automated (N=418,755) or two manual (N=25,888) BP measurements. For individuals with
1169 one manual and one automated BP measurement (N=13,521), we used the mean of these
1170 two values. For individuals with only one available BP measurement (N=413), we used this
1171 single value. After calculating BP values, we adjusted for medication use by adding 15 and
1172 10 mmHg to SBP and DBP, respectively, for individuals reported to be taking BP-lowering
1173 medication (N=94,289)⁵⁹. Descriptive summary statistics are shown in **Supplementary Table**
1174 **1a**.

1175 *UKB analysis models*

1176 For the UKB GWAS we performed linear mixed model (LMM) association testing under an
1177 additive genetic model of the three (untransformed) continuous, medication-adjusted BP
1178 traits (SBP, DBP, PP) for all measured and imputed genetic variants in dosage format using
1179 the BOLT-LMM (v2.3) software¹⁷. We also calculated the estimated SNP-wide heritability
1180 (h^2) in our data. Within the association analysis, we adjust for the following covariates: sex,
1181 age, age², BMI and a binary indicator variable for UKB vs UK BiLEVE to account for the
1182 different genotyping chips. The analysis of all HRC-imputed SNPs was restricted to variants
1183 with MAF \geq 1% and INFO > 0.1.

1184 *Genomic inflation and confounding*

1185 We applied the univariate LD score regression method (LDSR)¹⁸ to test for genomic inflation
1186 (expected for polygenic traits like BP, with large sample sizes, and especially also from
1187 analyses of such dense genetic data with many SNPs in high LD)⁶⁰. LDSR intercepts (and

1188 standard errors) were 1.217 (0.018), 1.219 (0.020) and 1.185 (0.017) for SBP, DBP and PP
1189 respectively, and were used to adjust the UKB GWAS results for genomic inflation, prior to
1190 the meta-analysis.

1191 **International Consortium for Blood Pressure (ICBP) GWAS**

1192 ICBP GWAS is an international consortium to investigate BP genetics⁶. We combined
1193 previously reported post-QC GWAS data from 54 studies (N=150,134)^{11,12,61}, with newly
1194 available GWAS data from a further 23 independent studies (N=148,890) using a fixed
1195 effects inverse variance weighted meta-analysis. The 23 studies providing new data were:
1196 ASCOT-SC, ASCOT-UK, BRIGHT, Dijon 3C, EPIC-CVD, GAPP, HCS, GS:SFHS, Lifelines, JUPITER,
1197 PREVEND, TWINSUK, GWAS-Fenland, InterAct-GWAS, OMICS-EPIC, OMICS-Fenland, UKHLS,
1198 GoDARTS-Illumina and GoDarts-Affymetrix, NEO, MDC, SardinIA, METSIM.

1199 All study participants were Europeans and were imputed to either the 1000 Genomes
1200 Project Phase 1 integrated release v.3 [March 2012] all ancestry reference panel⁶² or the
1201 HRC panel¹⁶. The final enlarged ICBP GWAS dataset included 77 cohorts (N=299,024).

1202 Full study names, cohort information and general study methods are included in
1203 **Supplementary Table 1b** and in **Supplementary Tables 20a-c**. GC was applied at study-level.
1204 The LDSR intercepts (standard error) for the ICBP GWAS meta-analysis were 1.089 (0.012),
1205 1.086 (0.012) and 1.066 (0.011) for SBP, DBP and PP, respectively.

1206 **Meta-analyses of discovery datasets**

1207 We performed a fixed-effects inverse variance weighted meta-analysis using METAL^{20,63} to
1208 obtain summary results from the UKB and ICBP GWAS, for up to N=757,601 participants and
1209 ~7.1 M SNPs with MAF \geq 1% for variants present in both the UKB data and ICBP meta-
1210 analysis for all three traits. The LDSR intercepts (standard error), in the discovery meta-
1211 analysis of UKB and ICBP were 1.156 (0.020), 1.160 (0.021) and 1.113 (0.018) for SBP, DBP
1212 and PP respectively. The LDSR intercept (standard error), after the exclusion of all published
1213 BP variants (see below) in the discovery meta-analysis of UKB and ICBP was 1.090 (0.018),
1214 1.097 (0.017) and 1.064 (0.015) for SBP, DBP and PP respectively, hence showing little
1215 inflation in the discovery GWAS after the exclusion of published loci (**Supplementary Fig.**
1216 **13**). No further correction was applied to the discovery meta-analysis of UKB and ICBP
1217 GWAS.

1218 **Previously reported variants**

1219 We compiled from the peer-reviewed literature all 357 SNPs previously reported to be
1220 associated with BP at the time that our analysis was completed, that have been identified
1221 and validated as the sentinel SNP in primary analyses from previous BP genetic association
1222 studies. These 357 published SNPs correspond to 274 distinct loci, according to locus
1223 definition of: (i) SNPs within \pm 500kb distance of each other; (ii) SNPs in Linkage
1224 Disequilibrium (LD), using a threshold of $r^2 \geq$ 0.1, calculated with PLINK (v2.0). We then

1225 augment this list to all SNPs present within our data, which are contained within these 274
1226 published BP loci, i.e. all SNPs which are located $\pm 500\text{kb}$ from each of the 357 published
1227 SNPs and/or in LD with any of the 357 previously validated SNPs ($r^2 \geq 0.1$).

1228 **Identification of novel signals: Two-stage and one-stage study designs**

1229 To identify novel signals of association with BP, two complementary study designs (which
1230 we term here “two-stage design” and “one-stage design”) were implemented in order to
1231 maximize the available data and minimize reporting of false positive associations.

1232 **Two-stage design: Overview:**

1233 All of the following criteria had to be satisfied for a signal to be reported as a novel signal of
1234 association with BP using our two-stage design:

- 1235 (i) the sentinel SNP shows significance ($P < 1 \times 10^{-6}$) in the discovery meta-analysis
1236 of UKB and ICBP, with concordant direction of effect between UKB and ICBP;
- 1237 (ii) the sentinel SNP is genome-wide significant ($P < 5 \times 10^{-8}$) in the combined meta-
1238 analysis of discovery and replication (MVP and EGCUT) (replication, described
1239 below);
- 1240 (iii) the sentinel SNP shows support ($P < 0.01$) in the replication meta-analysis of
1241 MVP and EGCUT alone (**Supplementary Methods**);
- 1242 (iv) the sentinel SNP has concordant direction of effect between the discovery and
1243 the replication meta-analyses;
- 1244 (v) the sentinel SNP must not be located within any of the 274 previously reported
1245 loci described above.

1246 The primary replicated trait was then defined as the BP trait with the most significant
1247 association from the combined meta-analysis of discovery and replication (in the case
1248 where a SNP was replicated for more than one BP trait.)

1249 **Two-stage design: Selection of variants from the discovery meta-analysis**

1250 We considered for follow-up SNPs in loci non-overlapping with previously reported loci
1251 according to both an LD threshold at r^2 of 0.1 and a 1Mb interval region, as calculated by
1252 PLINK⁶⁴. We obtained a list of such SNPs with $P < 1 \times 10^{-6}$ for any of the three BP traits,
1253 which also had concordant direction of effect between UKB vs ICBP (**Supplementary Table**
1254 **21**). By ranking the SNPs by significance in order of minimum P-value across all BP traits, we
1255 performed an iterative algorithm to determine the number of novel signals (**Supplementary**
1256 **Methods**), and identify the sentinel SNP (most significant) per locus.

1257 **Two-stage design: Replication analysis**

1258 We considered SNPs with $\text{MAF} \geq 1\%$ for an independent replication in MVP (max
1259 $N=220,520$)¹⁴ and in EGCUT Biobank ($N=28,742$)¹⁵ (**Supplementary Methods**). This provides
1260 a total of $N=249,262$ independent samples of European descent available for replication.

1261 Additional information on the analyses of the two replication datasets is provided in
1262 **Supplementary Methods** and in **Supplementary Table 1c**.

1263 The two datasets were then combined using fixed effects inverse variance weighted meta-
1264 analysis and summary results for all traits were obtained for the replication meta-analysis
1265 dataset.

1266 **Two-stage design: Combined meta-analysis of discovery and replication meta-analyses**

1267 The meta-analyses were performed within METAL software⁶³ using fixed effects inverse
1268 variance weighted meta-analysis (**Supplementary Methods**). The variants from the
1269 discovery GWAS that required proxies for replication are shown in **Supplementary Table 22**.
1270 The combined meta-analysis of both the discovery data (N=757,601) and replication meta-
1271 analysis (max N=249,262) provided a maximum sample size of N=1,006,863.

1272 **One-stage design: Overview**

1273 Variants that were looked-up but did not replicate according to the two-stage criteria were
1274 considered in a one-stage design. All of the following criteria had to be satisfied for a signal
1275 to be reported as a novel signal of association with BP using our one-stage criteria:

- 1276 i) the sentinel SNP has $P < 5 \times 10^{-9}$ in the discovery (UKB+ICBP) meta-analysis;
- 1277 ii) the sentinel SNP shows support ($P < 0.01$) in the UKB GWAS alone;
- 1278 iii) the sentinel SNP shows support ($P < 0.01$) in the ICBP GWAS alone;
- 1279 iv) the sentinel SNP has concordant direction of effect between UKB and ICBP
1280 datasets;
- 1281 v) The sentinel SNP must not be located within any of the 274 previously reported
1282 loci described above (**Supplementary Table 4**) or the recently reported non-
1283 replicated loci from Hoffman et al⁹ (**Supplementary Table 23**).

1284 We selected the one-stage P -value threshold to be an order of magnitude more stringent
1285 than a genome-wide significance P -value, so as to ensure robust results and to minimize
1286 false positive findings. The threshold of $P < 5 \times 10^{-9}$ has been proposed as a more
1287 conservative statistical significance threshold, e.g. for whole-genome sequencing-based
1288 studies²¹.

1289 Selection of variants from the meta-analysis of UKB and ICBP was performed as described
1290 above for the two-stage design.

1291 **Conditional Analysis**

1292 We performed conditional analyses using the GWAS discovery meta-analysis data, in order
1293 to identify any independent secondary signals in addition to the sentinel SNPs at the 901
1294 loci. We used two different methodological approaches, each using the Genome-wide
1295 Complex Traits Analysis (GCTA) software²²: (i) full “genome-wide conditional analysis” with
1296 joint multivariate analysis and stepwise model selection across all three BP traits; and (ii)
1297 “locus-specific conditional analysis” for the primary BP trait conditioning on the sentinel

1298 SNPs within each locus (**Supplementary Methods**). For robustness, secondary signals are
1299 only reported if obtained from both approaches. All secondary signals were selected at
1300 genome-wide significance level, with $MAF \geq 1\%$ and confirmed to be pairwise-LD-
1301 independent ($r^2 < 0.1$), as well as not being in LD with any of the published or sentinel SNPs
1302 at any of the 901 BP-associated loci ($r^2 < 0.1$). In all cases the UKB data was used as the
1303 reference genetic data for LD calculation, restricted to individuals of European ancestry
1304 only.

1305 **Functional analyses: Variants**

1306 We used an integrative bioinformatics approach to collate functional annotation at both the
1307 variant level (for each sentinel SNP within all BP loci) and the gene level (using SNPs in LD r^2
1308 ≥ 0.8 with the sentinel SNPs). At the variant level, we use Variant Effect Predictor (VEP) to
1309 obtain comprehensive characterization of variants, including consequence (e.g. downstream
1310 or non-coding transcript exon), information on nearest genomic features and, where
1311 applicable, amino acid substitution functional impact, based on SIFT and PolyPhen. The
1312 biomaRt R package is used to further annotate the nearest genes.

1313 We evaluated all SNPs in LD ($r^2 \geq 0.8$) with our novel sentinel SNPs for evidence of mediation
1314 of expression quantitative trait loci (eQTL) in all 44 tissues using the Genotype-Tissue
1315 Expression (GTEx) database, to highlight specific tissue types which show eQTLs for a larger
1316 than expected proportion of novel loci. We further seek to identify novel loci with the
1317 strongest evidence of eQTL associations in arterial tissue, in particular. A locus is annotated
1318 with a given eGene only if the most significant eQTL SNP for the given eGene is in high LD (r^2
1319 ≥ 0.8) with the sentinel SNP, suggesting that the eQTL signal co-localises with the sentinel
1320 SNP.

1321 We annotated nearest genes, eGenes (genes whose expression is affected by eQTLs) and Hi-
1322 C interactors with HUVEC, HVMSC and HAEC expression from the Fantom5 project. Genes
1323 that had higher than median expression levels in the given cell types were indicated as
1324 expressed.

1325 To identify SNPs in the novel loci that have a non-coding functional effect (influence binding
1326 of transcription factors or RNA polymerase, or influence DNase hypersensitivity sites or
1327 histone modifications), we used DeepSEA, a deep learning algorithm, that learnt the binding
1328 and modification patterns of ~ 900 cell/factor combinations⁶⁵. A change of >0.1 in the
1329 binding score predicted by DeepSEA for the reference and alternative alleles respectively
1330 was used as cut-off to find alleles with non-coding functional effect (**Supplementary**
1331 **Methods**)

1332 We identified potential target genes of regulatory SNPs using long-range chromatin
1333 interaction (Hi-C) data from HUVECs²³, aorta, adrenal glands, neural progenitor and
1334 mesenchymal stem cell, which are tissues and cell types that are considered relevant for
1335 regulating BP²⁴. We find the most significant promoter interactions for all potential

1336 regulatory SNPs (RegulomeDB score ≤ 5) in LD ($r^2 \geq 0.8$) with our novel sentinel SNPs and
1337 published SNPs, and choose the interactors with the SNPs of highest regulatory potential to
1338 annotate the loci.

1339 We then performed overall enrichment testing across all loci. Firstly, we used DEPICT⁶⁶
1340 (Data-driven Expression Prioritized Integration for Complex Traits) to identify tissues and
1341 cells which are highly expressed at genes within the BP loci (**Supplementary Methods**).
1342 Secondly, we used DEPICT to test for enrichment in gene sets associated with biological
1343 annotations (manually curated and molecular pathways, phenotype data from mouse KO
1344 studies) (**Supplementary Methods**). We report significant enrichments with a false
1345 discovery rate < 0.01 . The variants tested were i) the 357 published BP associated SNPs at
1346 the time of analysis and ii) a set including all (published and novel) variants (with novel SNPs
1347 filtered by highest significance, $P < 1 \times 10^{-12}$).

1348 Furthermore, to investigate cell type specific enrichment within DNase I sites, we used
1349 FORGE, which tests for enrichment of SNPs within DNase I sites in 123 cell types from the
1350 Epigenomics Roadmap Project and ENCODE²⁵ (**Supplementary Methods**). Two analyses
1351 were compared (i) using published SNPs only; (ii) using sentinel SNPs at all 901 loci, in order
1352 to evaluate the overall tissue specific enrichment of BP associated variants.

1353 **Functional analyses: Genes**

1354 At the gene level, we used Ingenuity Pathway Analysis (IPA) software (IPA®, QIAGEN
1355 Redwood City) to review genes with prior links to BP, based on annotation with the
1356 “Disorder of Blood Pressure”, “Endothelial Development” and “Vascular Disease” Medline
1357 Subject Heading (MESH) terms. We used the Mouse Genome Informatics (MGI) tool to
1358 identify BP and cardiovascular relevant mouse knockout phenotypes for all genes linked to
1359 BP in our study. We also used IPA to identify genes that interact with known targets of anti-
1360 hypertensive drugs. Genes were also evaluated for evidence of small molecule druggability
1361 or known drugs based on queries of the Drug Gene Interaction database.

1362 **Lookups in non-European ancestries**

1363 As a secondary analysis, we look up all known and novel BP-associated SNPs in Africans
1364 (7,782) and South Asians (10,322) from UKB using BOLT-LMM analysis for each BP trait
1365 within each ancestry (**Supplementary Methods**).

1366 **Effects on other traits and diseases**

1367 We queried SNPs against GWAS catalog²⁶ and PhenoScanner²⁷, including genetics and
1368 metabolomics databases, to investigate cross-trait effects, extracting all association results
1369 with genome-wide significance at $P < 5 \times 10^{-8}$ for all SNPs in high LD ($r^2 \geq 0.8$) with the 535
1370 sentinel novel SNPs, to highlight the loci with strongest evidence of association with other
1371 traits. We further evaluated these effects using DisGeNET^{28,29}. At the gene level,
1372 overrepresentation enrichment analysis (ORA) with WebGestalt⁶⁷ on the nearest genes to
1373 all BP loci was carried out. Moreover, we tested sentinel SNPs at all published and novel

1374 (N=901) loci for association with lifestyle related data including food, water and alcohol
1375 intake, anthropomorphic traits and urinary sodium, potassium and creatinine excretion
1376 using the recently developed Stanford Global Biobank Engine and the Gene ATLAS⁶⁸. Both
1377 are search engines for GWAS findings for multiple phenotypes in UK Biobank. We used a
1378 Bonferroni corrected significance threshold of $P < 1 \times 10^{-6}$ to deem significance.

1379 **Genetic risk scores and percentage of variance explained**

1380 We calculated a weighted genetic risk score (GRS) (**Supplementary Table 24**) to provide an
1381 estimate of the combined effect of the BP raising variants on BP and risk of hypertension
1382 and applied this to the UKB data (**Supplementary Methods**). Our analysis included 423,713
1383 unrelated individuals of European ancestry of whom 392,092 individuals were free of
1384 cardiovascular events at baseline.

1385 We assessed the association of the continuous GRS variable on BP and with the risk of
1386 hypertension, with and without adjustment for sex. We then compared BP levels and risk of
1387 hypertension, respectively, for individuals in the top vs bottom quintiles of the GRS
1388 distribution. Similar analyses were performed for the top vs bottom deciles of the GRS
1389 distribution. All analyses were restricted to the 392,092 unrelated individuals of European
1390 ancestry from UKB. As a sensitivity analysis to assess for evidence of bias in the UKB results,
1391 we also carried out similar analyses in Airwave, an independent cohort of N=14,004
1392 unrelated participants of European descent³⁰ (**Supplementary Methods**).

1393 We calculated the association of the GRS with cardiovascular disease in unrelated
1394 participants in UKB data, based on self-reported medical history, and linkage to
1395 hospitalization and mortality data (**Supplementary Table 25**). We use logistic regression
1396 with binary outcome variables for composite incident cardiovascular disease
1397 (**Supplementary Methods**), incident myocardial infarction and incident stroke (using the
1398 algorithmic UKB definitions) and GRS as explanatory variable (with and without sex
1399 adjustment).

1400 We also assessed the association of this GRS with BP in unrelated individuals Africans
1401 (N=6,970) and South Asians (N=8,827) from the UKB to see whether BP-associated SNPs
1402 identified from GWAS predominantly in Europeans are also associated with BP in
1403 populations of non-European ancestry.

1404 We calculated the percentage of variance in BP explained by genetic variants using the
1405 independent Airwave cohort (N=14,004) (**Supplementary Methods**). We considered three
1406 different levels of the GRS: (i) all pairwise-independent, LD-filtered ($r^2 < 0.1$) published SNPs
1407 within the known loci; (ii) all known SNPs and sentinel SNPs at novel loci; (iii) all
1408 independent signals at all 901 known and novel loci including the 163 secondary SNPs.

1409 **Data availability statement**

1410 The UKB GWAS data can be accessed from the UK Biobank data repository
1411 (<http://biota.osc.ox.ac.uk/>). The genetic and phenotypic UKB data are available upon

1412 application to the UK Biobank (<https://www.ukbiobank.ac.uk>). ICBP summary data can be
1413 assessed through request to ICBP steering committee. Contact Mark Caulfield
1414 (m.j.caulfield@qmul.ac.uk) or Paul Elliott (p.elliott@imperial.ac.uk) to apply for access to
1415 the data. The UKB+ICBP summary data can be assessed through request to Paul Elliott
1416 (p.elliott@imperial.ac.uk) or Mark Caulfield (m.j.caulfield@qmul.ac.uk). All replication data
1417 generated during this study are included in the published article. For example, association
1418 results of look-up variants from our replication analyses and the subsequent combined
1419 meta-analyses are contained within the Supplementary Tables provided.

1420 **Reporting Summary**

1421 Further information on experimental design is available in the Life Sciences Reporting
1422 Summary linked to this article.

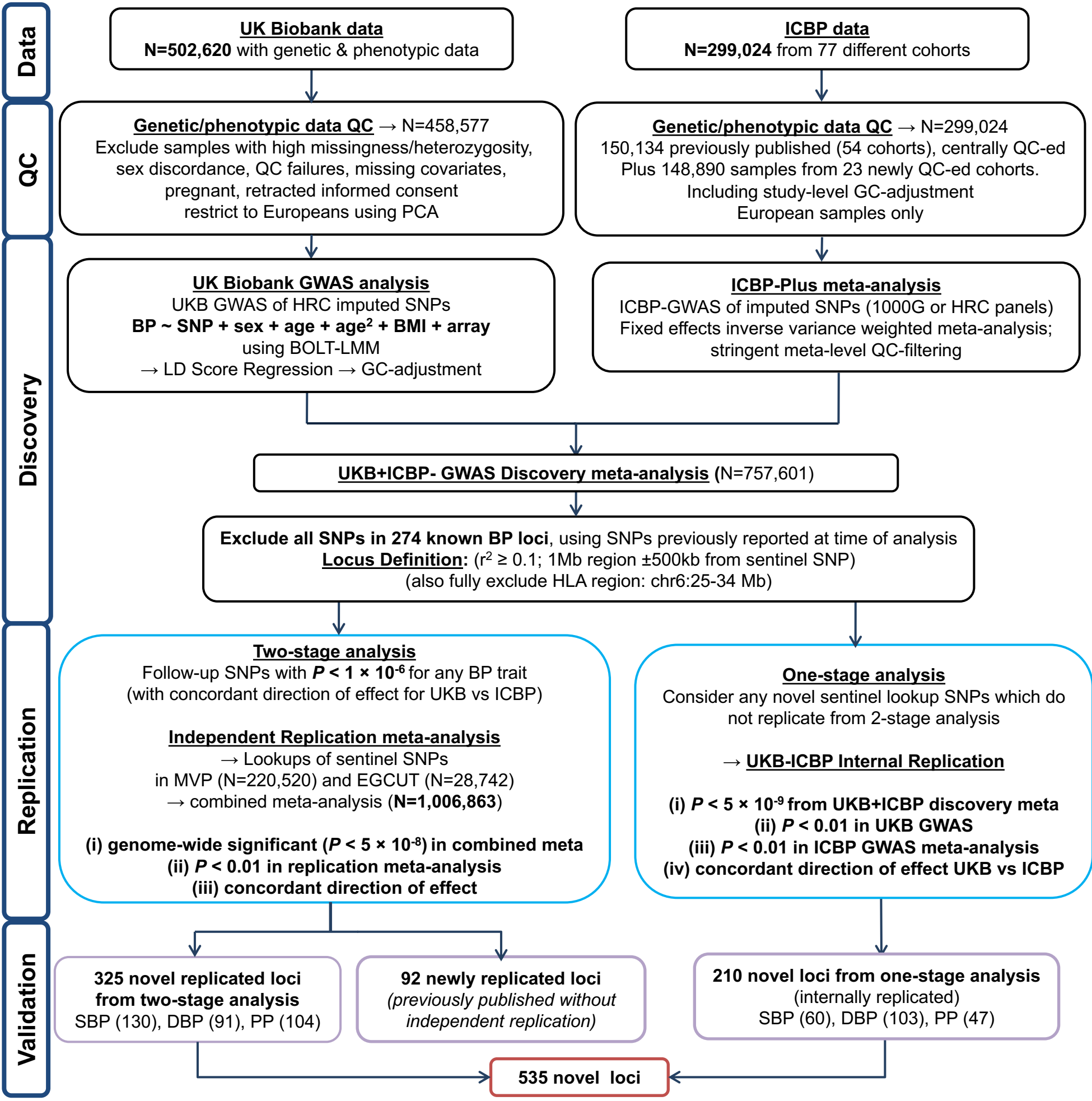
1423 **Ethics Statement**

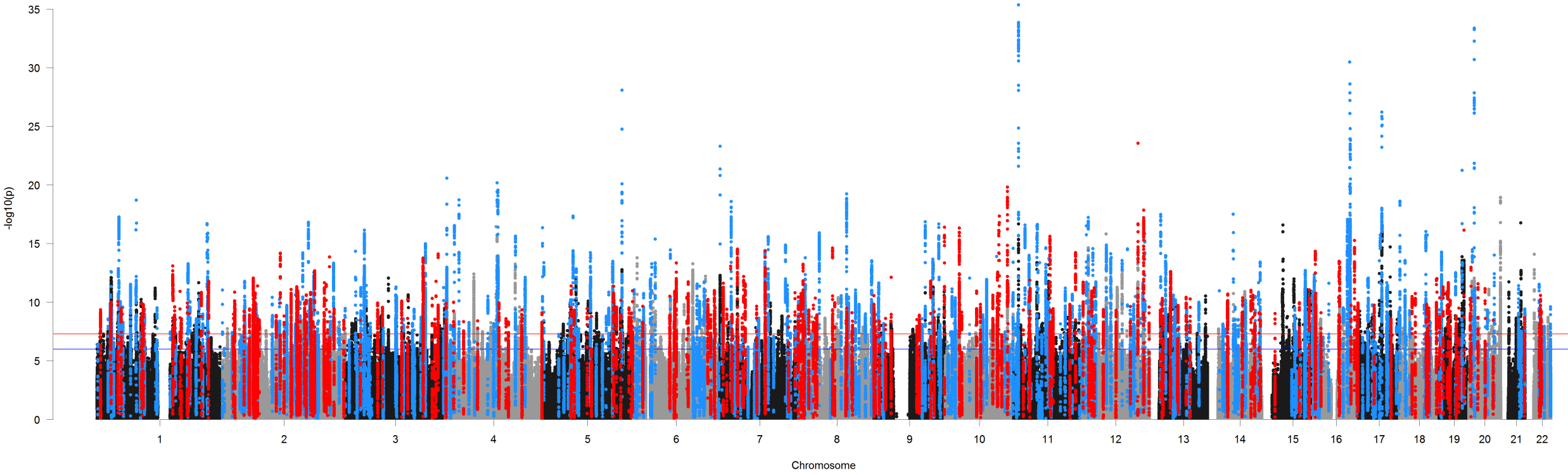
1424 The UKB study has approval from the North West Multi-Centre Research Ethics Committee.
1425 Any participants from UKB who withdrew consent have been removed from our analysis.
1426 Each cohort within the ICBP meta-analysis as well as our independent replication cohorts of
1427 MVP and EGCUT had ethical approval locally. More information on the participating cohorts
1428 is available in **Supplementary Methods**.

1429 **References**

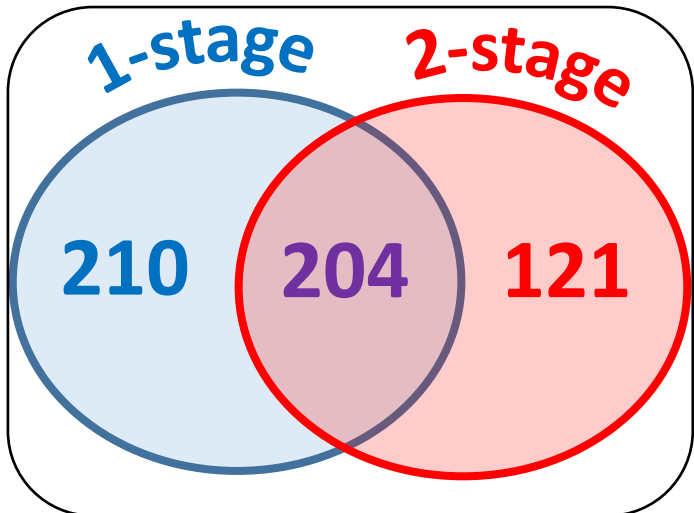
- 1430 57. Wain, L.V. *et al.* Novel insights into the genetics of smoking behaviour, lung function,
1431 and chronic obstructive pulmonary disease (UK BiLEVE): a genetic association study
1432 in UK Biobank. *Lancet Respir Med* **3**, 769-81 (2015).
- 1433 58. Bycroft, C.F., C; Petkova, D; Band, G; Elliot, LT; Sharp, K; Motyer, A; Vukcevic, D;
1434 Delaneau, O; O'Connell, J; Cortes, A; Welsh, S; McVean, G; Leslie, S; Donnelly, P;
1435 Marchini, J. Genome-wide genetic data on 500,000 UK Biobank Participants. *bioRxiv*
1436 **166298** (2017).
- 1437 59. Tobin, M.D., Sheehan, N.A., Scurrah, K.J. & Burton, P.R. Adjusting for treatment
1438 effects in studies of quantitative traits: antihypertensive therapy and systolic blood
1439 pressure. *Stat Med* **24**, 2911-35 (2005).
- 1440 60. Marouli, E. *et al.* Rare and low-frequency coding variants alter human adult height.
1441 *Nature* **542**, 186-190 (2017).
- 1442 61. Wain, L.V. *et al.* Genome-wide association study identifies six new loci influencing
1443 pulse pressure and mean arterial pressure. *Nat Genet* **43**, 1005-11 (2011).
- 1444 62. 1000 Genomes Project Consortium *et al.* A global reference for human genetic
1445 variation. *Nature* **526**, 68-74 (2015).
- 1446 63. Willer, C.J., Li, Y. & Abecasis, G.R. METAL: fast and efficient meta-analysis of
1447 genomewide association scans. *Bioinformatics* **26**, 2190-1 (2010).
- 1448 64. Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-
1449 based linkage analyses. *Am J Hum Genet* **81**, 559-75 (2007).
- 1450 65. Zhou, J. & Troyanskaya, O.G. Predicting effects of noncoding variants with deep
1451 learning-based sequence model. *Nat Methods* **12**, 931-4 (2015).
- 1452 66. Pers, T.H. *et al.* Biological interpretation of genome-wide association studies using
1453 predicted gene functions. *Nat Commun* **6**, 5890 (2015).

- 1454 67. Wang, J., Vasaikar, S., Shi, Z., Greer, M. & Zhang, B. WebGestalt 2017: a more
1455 comprehensive, powerful, flexible and interactive gene set enrichment analysis
1456 toolkit. *Nucleic Acids Res* (2017).
1457 68. Canela-Xandri, O.R., Konrad; Tenesa, Albert. An atlas of genetic associations in UK
1458 Biobank. *bioRxiv* **176834** (2017).
1459

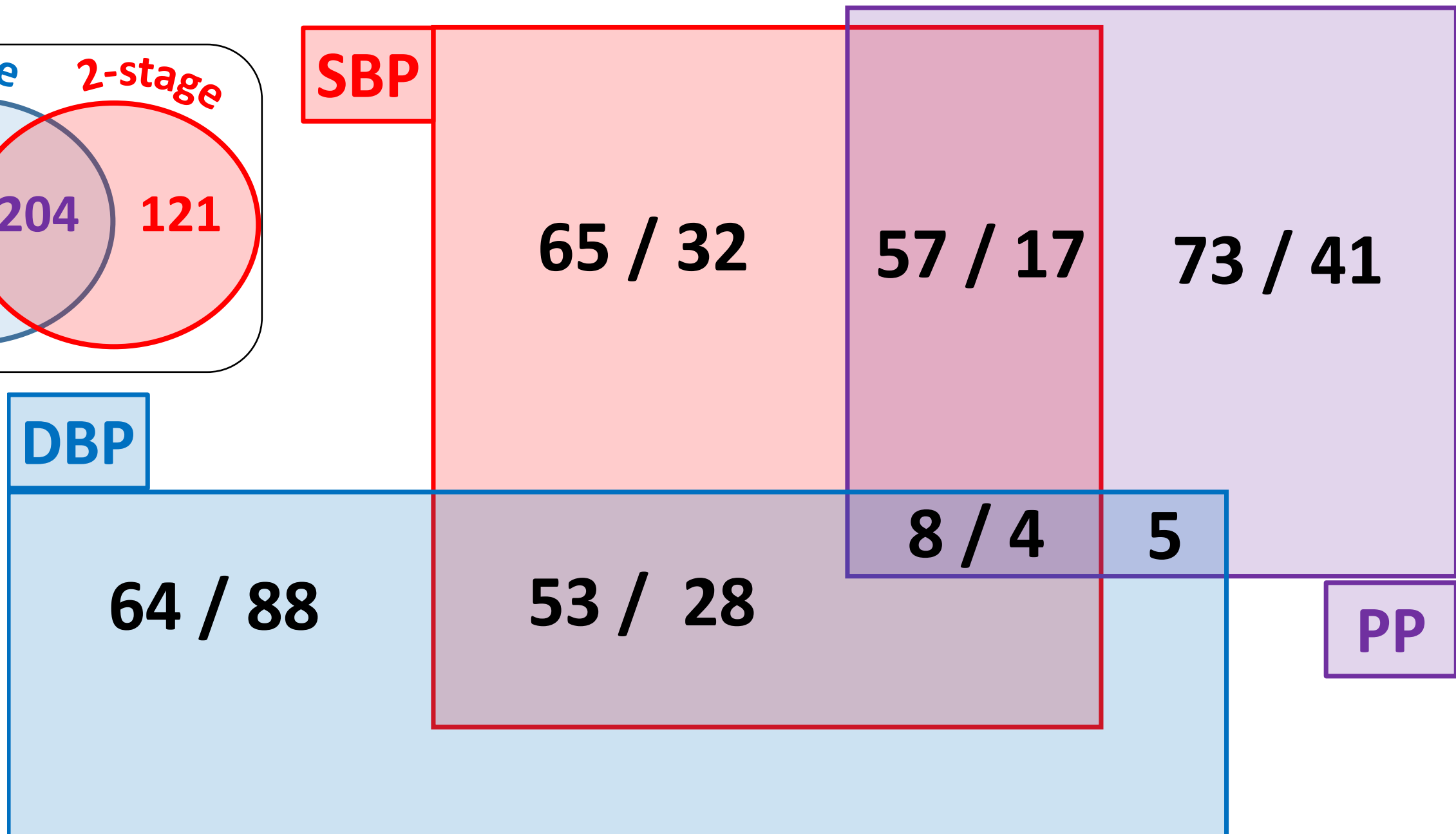


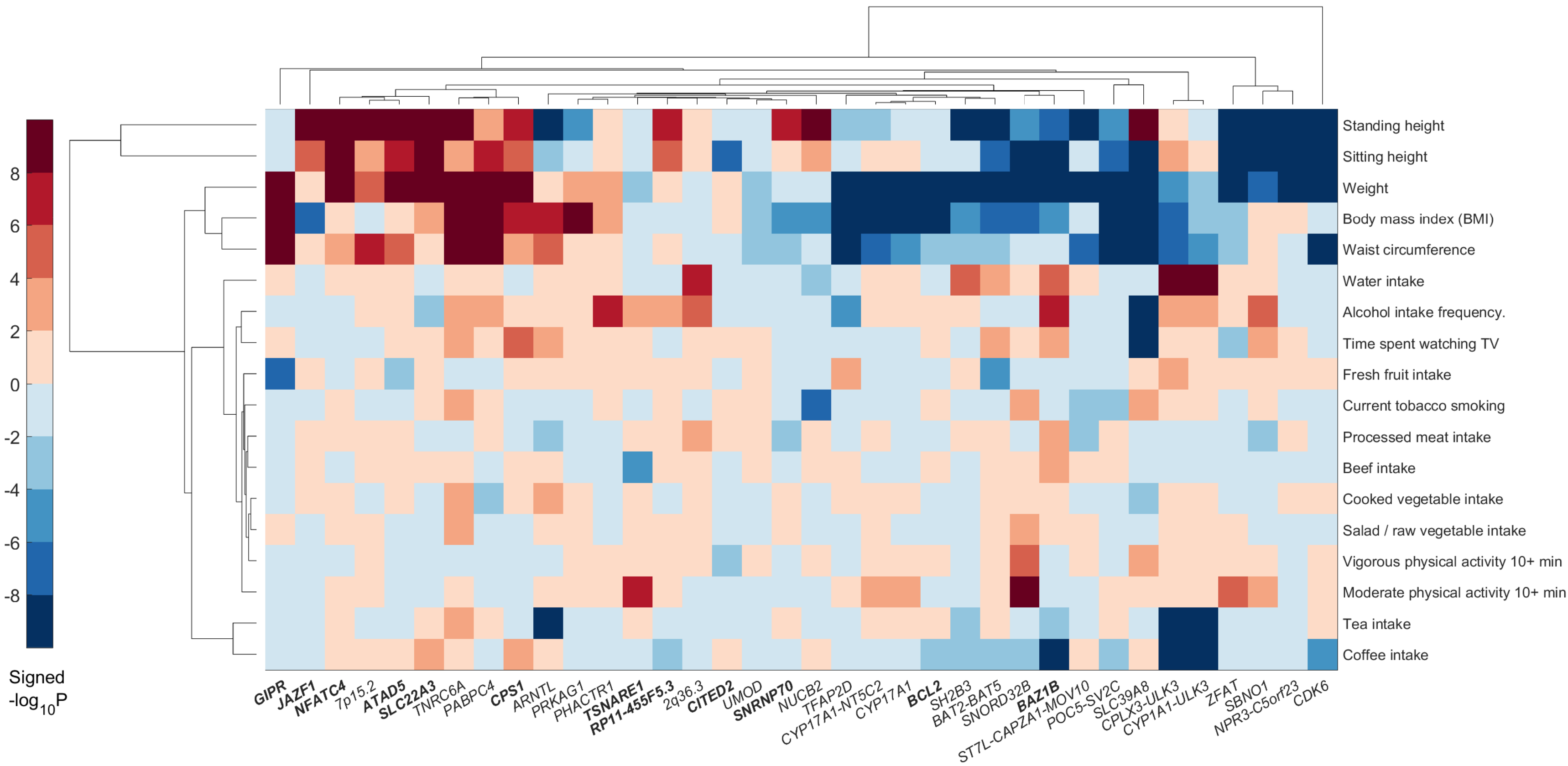


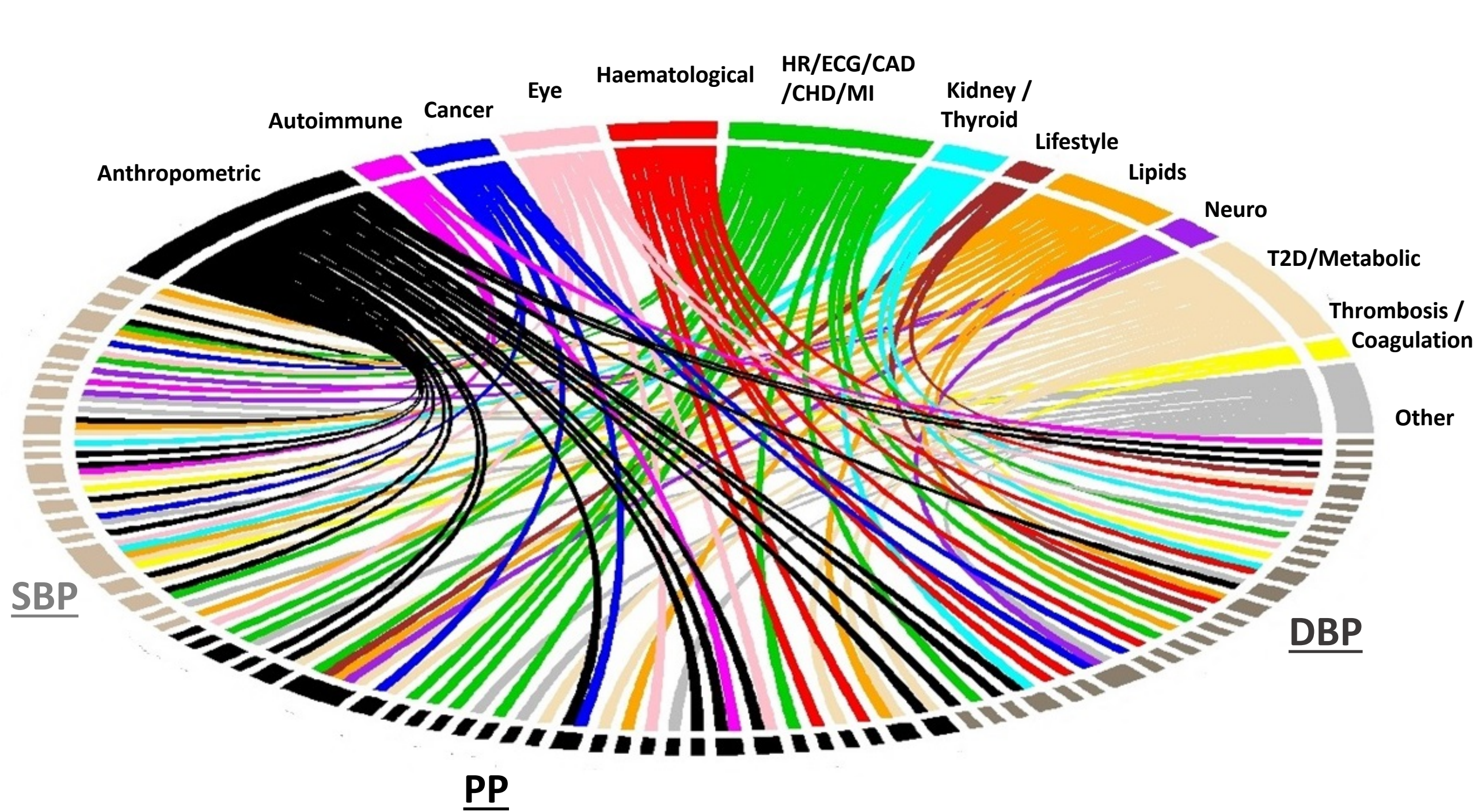
(A)

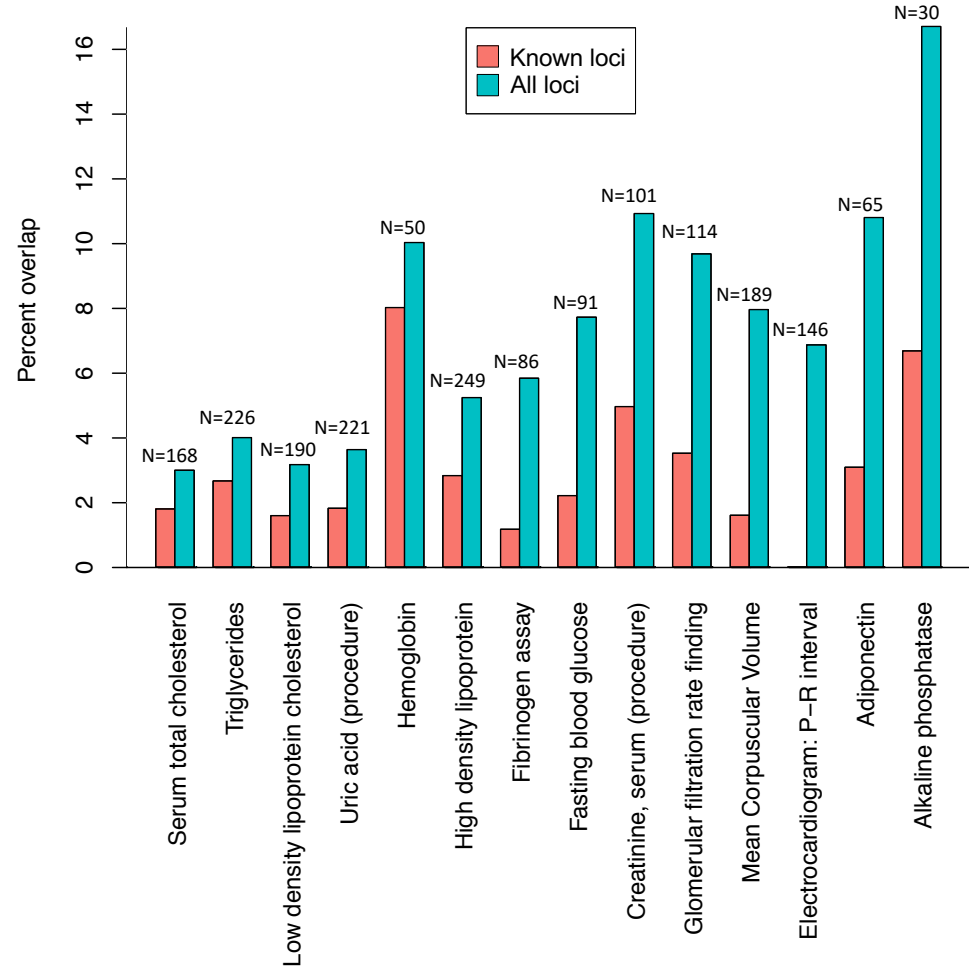
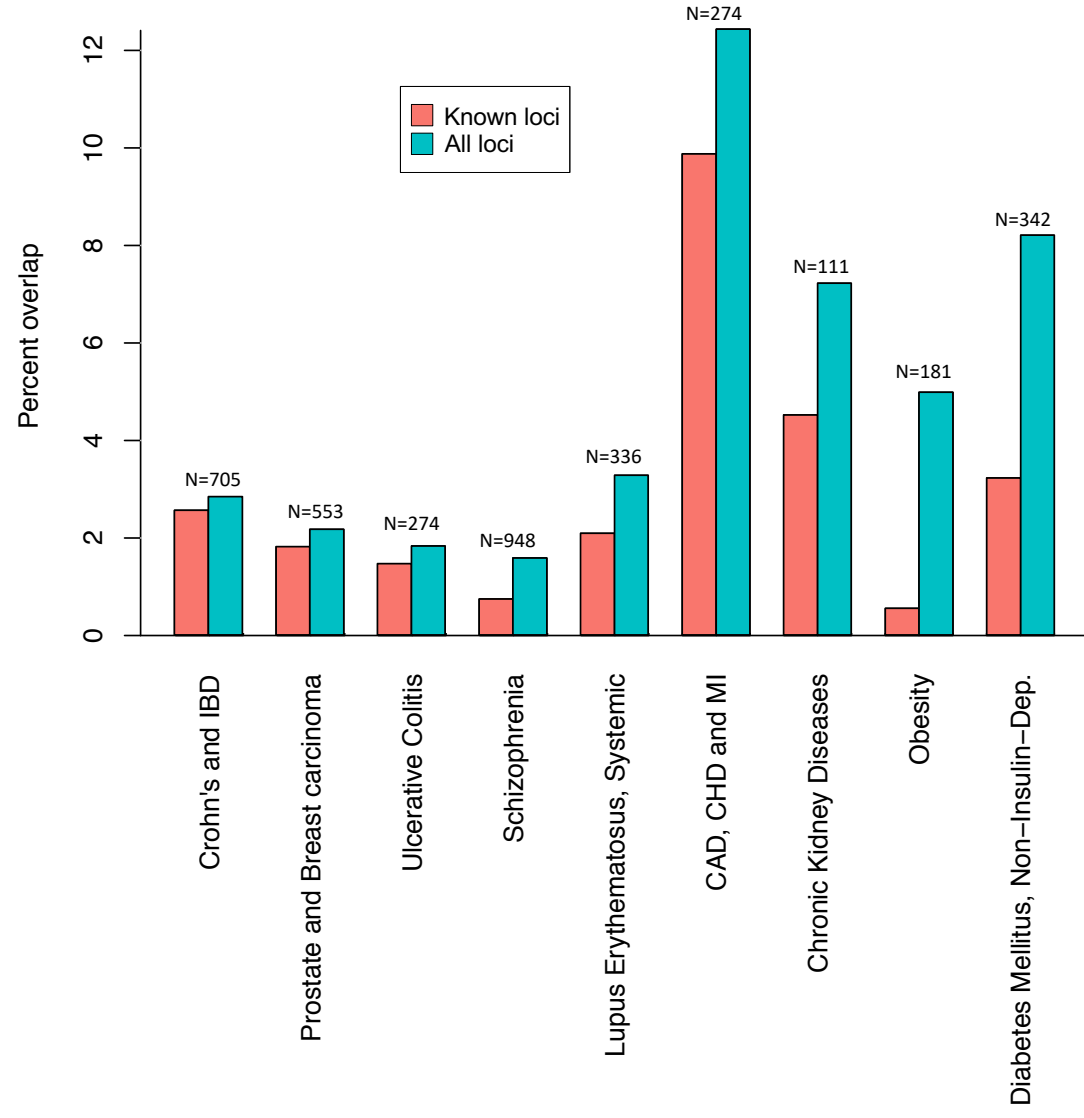


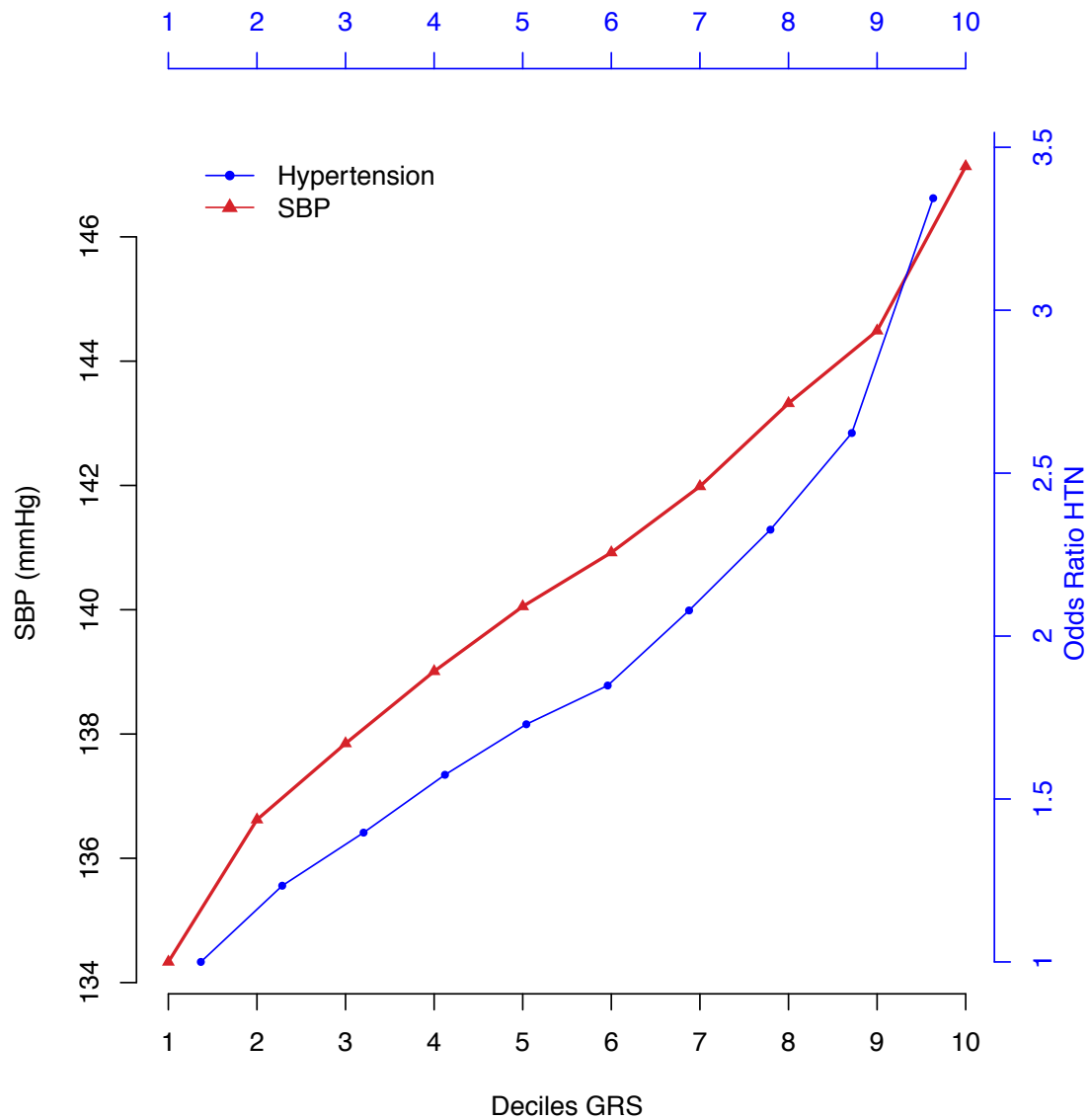
(B)







a**b**

a**b**