Multi-ancestry GWAS of the electrocardiographic PR interval identifies 210 loci underlying cardiac conduction

Ioanna Ntalla[1*], Lu-Chen Weng[2, 3*], James H. Cartwright[1], Amelia Weber Hall[2, 3], Gardar Sveinbjornsson[4], Nathan R. Tucker[2, 3], Seung Hoan Choi[3], Mark D. Chaffin[3], Carolina Roselli[3, 5], Michael R. Barnes[1, 6], Borbala Mifsud[1, 7], Helen R. Warren[1, 6], Caroline Hayward[8], Jonathan Marten[8], James J. Cranley[1], Maria Pina Concas[9], Paolo Gasparini[9, 10], Thibaud Boutin[8], Ivana Kolcic[11], Ozren Polasek[11-13], Igor Rudan[14], Nathalia M. Araujo[15], Maria Fernanda Lima-Costa[16], Antonio Luiz P. Ribeiro[17], Renan P. Souza[15], Eduardo Tarazona-Santos[15], Vilmantas Giedraitis[18], Erik Ingelsson[19-22], Anubha Mahajan[23], Andrew P. Morris[23-25], Fabiola Del Greco M.[26], Luisa Foco[26], Martin Gögele[26], Andrew A. Hicks[26], James P. Cook[24], Lars Lind[27], Cecilia M. Lindgren[28-30], Johan Sundström[31], Christopher P. Nelson[32, 33], Muhammad B. Riaz[32, 33], Nilesh J. Samani[32, 33], Gianfranco Sinagra[34], Sheila Ulivi[9], Mika Kähönen[35, 36], Pashupati P. Mishra[37, 38], Nina Mononen[37, 38], Kjell Nikus[39, 40], Mark J. Caulfield[1, 6], Anna Dominiczak[41], Sandosh Padmanabhan[41, 42], May E. Montasser[43, 44], Jeff R. O'Connell[43, 44], Kathleen Ryan[43, 44], Alan R. Shuldiner[43, 44], Stefanie Aeschbacher[45], David Conen[45, 46], Lorenz Risch[47-49], Sébastien Thériault[46, 50], Nina Hutri-Kähönen[51, 52], Terho Lehtimäki[37, 38], Leo-Pekka Lyytikäinen[37-39], Olli T. Raitakari[53-55], Catriona L. K. Barnes[14], Harry Campbell[14], Peter K. Joshi[14], James F. Wilson[8, 14], Aaron Isaacs[56], Jan A. Kors[57], Cornelia M. van Duijn[58], Paul L. Huang[2], Vilmundur Gudnason[59, 60], Tamara B. Harris[61], Lenore J. Launer[61], Albert V. Smith[59, 62], Erwin P. Bottinger[63], Ruth J. F. Loos[63, 64], Girish N. Nadkarni[63], Michael H. Preuss[63], Adolfo Correa[65], Hao Mei[66], James Wilson[67], Thomas Meitinger[68-70], Martina Müller-Nurasyid[68, 71-73], Annette Peters[68, 74, 75], Melanie Waldenberger[68, 75, 76], Massimo Mangino[77, 78], Timothy D. Spector[77], Michiel Rienstra[5], Yordi J.

van de Vegte[5], Pim van der Harst[5], Niek Verweij[5, 79], Stefan Kääb[68, 73], Katharina Schramm[68, 71, 73], Moritz F. Sinner[68, 73], Konstantin Strauch[71, 72], Michael J. Cutler[80], Diane Fatkin[81-83], Barry London[84], Morten Olesen[85, 86], Dan M. Roden[87], M. Benjamin Shoemaker[88], J. Gustav Smith[89], Mary L. Biggs[90, 91], Joshua C. Bis[90], Jennifer A. Brody[90], Bruce M. Psaty[90, 92, 93], Ken Rice[91], Nona Sotoodehnia[90, 92, 94], Alessandro De Grandi[26], Christian Fuchsberger[26], Cristian Pattaro[26], Peter P. Pramstaller[26], Ian Ford[95], J. Wouter Jukema[96, 97], Peter W. Macfarlane[98], Stella Trompet[99], Marcus Dörr[100, 101], Stephan B. Felix[100, 101], Uwe Völker[100, 102], Stefan Weiss[100, 102], Aki S. Havulinna[103, 104], Antti Jula[103], Katri Sääksjärvi[103], Veikko Salomaa[103], Xiuqing Guo[105], Susan R. Heckbert[106], Henry J. Lin[105], Jerome I. Rotter[107], Kent D. Taylor[105], Jie Yao[108], Renée de Mutsert[109], Arie C. Maan[96], Dennis O. Mook-Kanamori[109, 110], Raymond Noordam[99], Francesco Cucca[111], Jun Ding[112], Edward G. Lakatta[113], Yong Qian[112], Kirill V. Tarasov[113], Daniel Levy[114, 115], Honghuang Lin[115, 116], Christopher H. Newton-Cheh[3, 117], Kathryn L. Lunetta[115, 118], Alison D. Murray[119], David J. Porteous[120, 121], Blair H. Smith[122], Bruno H. Stricker[123], André Uitterlinden[124], Marten E. van den Berg[123], Jeffrey Haessler[125], Rebecca D. Jackson[126], Charles Kooperberg[125], Ulrike Peters[125], Alexander P. Reiner[125, 127], Eric A. Whitsel[128], Alvaro Alonso[129], Dan E. Arking[130], Eric Boerwinkle[131], Georg B. Ehret[132], Elsayed Z. Soliman[133], Christy L. Avery[134, 135], Stephanie M. Gogarten[91], Kathleen F. Kerr[91], Cathy C. Laurie[91], Amanda A. Seyerle[136], Adrienne Stilp[91], Solmaz Assa[5], M. Abdullah Said[5], M. Yldau van der Ende[5], Pier D. Lambiase[137, 138], Michele Orini[137, 139], Julia Ramirez[1, 138], Stefan Van Duijvenboden[1, 138], David O. Arnar[4, 60, 140], Daniel F. Gudbjartsson[4, 141], Hilma Holm[4], Patrick Sulem[4], Gudmar Thorleifsson[4], Rosa B. Thorolfsdottir[4, 60], Unnur Thorsteinsdottir[4, 60], Emelia J. Benjamin[115, 142, 143], Andrew Tinker[1, 6], Kari Stefansson[4, 60], Patrick T. Ellinor[2, 3, 144], Yalda Jamshidi[145], Steven A. Lubitz[2, 3, 144#], and Patricia B. Munroe[1, 6#]

1. William Harvey Research Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, UK.
2. Cardiovascular Research Center, Massachusetts General Hospital, Boston, MA, USA.
3. Program in Medical and Population Genetics, The Broad Institute of MIT and Harvard, Cambridge, MA, USA.
4. deCODE genetics/Amgen, Inc., Reykjavik, Iceland.
5. University of Groningen, University Medical Center Groningen, Department of Cardiology, Groningen, the Netherlands.
6. National Institute for Health Research, Barts Cardiovascular Biomedical Research Centre, Queen Mary University of London, London, UK.
7. College of Health and Life Sciences, Hamad Bin Khalifa University, Education City, Doha, Qatar.
8. Medical Research Council Human Genetics Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, UK.
9. Institute for Maternal and Child Health-IRCCS 'Burlo Garofolo', Trieste, Italy.
10. University of Trieste, Department of Medicine, Surgery and Health Sciences, Trieste, Italy.
11. University of Split School of Medicine, Split, Croatia.
12. Clinical Hospital Centre Split, Split, Croatia.
13. Psychiatric Hospital Sveti Ivan, Zagreb, Croatia.
14. Centre for Global Health Research, Usher Institute of Population Health Sciences and Informatics, University of Edinburgh, Edinburgh, UK.

15. Departamento de Biologia Geral, Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brazil.

16. Rene Rachou Reserch Institute, Oswaldo Cruz Foundation, Belo Horizonte, Minas Gerais, Brazil.

17. Hospital das Clínicas e Faculdade de Medicina, Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brazil.

18. Department of Public Health, Geriatrics, Uppsala University, Uppsala, Sweden.

19. Department of Medicine, Division of Cardiovascular Medicine, Stanford University School of Medicine, Stanford, CA, USA.

20. Stanford Cardiovascular Institute, Stanford University, Stanford, CA, USA.

21. Stanford Diabetes Research Center, Stanford University, Stanford, CA, USA.

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

23. Wellcome Centre for Human Genetics, University of Oxford, Oxford, UK.

24. Department of Biostatistics, University of Liverpool, Liverpool, UK.

25. Division of Musculoskeletal and Dermatological Sciences, University of Manchester, Manchester, UK.

26. Institute for Biomedicine, Eurac Research, Affiliated Institute of the University of Lübeck, Bolzano, Italy.

27. Medical Sciences, Uppsala University Hospital, Uppsala, Sweden.

28. Li Ka Shing Centre for Health Information and Discovery, Big Data Institute, Nuffield Department of Medicine, University of Oxford , Oxford, UK.

29. The Wellcome Centre for Human Genetics, Nuffield Department of Medicine, University of Oxford , Oxford, UK.

30. The Broad Institute of MIT and Harvard, Cambridge, MA, USA.

31. Department of Medical Sciences, Uppsala University, Uppsala, Sweden.

32. Department of Cardiovascular Sciences, University of Leicester, Cardiovascular Research Centre, Glenfield Hospital, Leicester, UK.

33. NIHR Leicester Biomedical Research Centre, Glenfield Hospital, Groby Road, Leicester, UK.

34. Cardiovascular Department, Azienda Ospedaliera Universitaria Integrata of Trieste, Trieste, Italy.

35. Department of Clinical Physiology, Tampere University Hospital, Tampere, Finland.

36. Department of Clinical Physiology, Finnish Cardiovascular Research Center - Tampere, Faculty of Medicine and Health Technology, Tampere University, Tampere, Finland.

37. Department of Clinical Chemistry, Fimlab Laboratories, Tampere, Finland.

38. Department of Clinical Chemistry, Finnish Cardiovascular Research Center - Tampere, Faculty of Medicine and Health Technology, Tampere University, Tampere, Finland.

39. Department of Cardiology, Heart Center, Tampere University Hospital, Tampere, Finland.

40. Department of Cardiology, Finnish Cardiovascular Research Center - Tampere, Faculty of Medicine and Health Technology, Tampere University, Tampere, Finland.

41. Institute of Cardiovascular and Medical Sciences, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow, UK.

42. Institute of Cardiovascular and Medical Sciences, University of Glasgow, Glasgow, UK.

43. Division of Endocrinology, Diabetes, and Nutrition, University of Maryland School of Medicine, Baltimore, MD, USA.

44. Program for Personalized and Genomic Medicine, University of Maryland School of Medicine, Baltimore, MD, USA.

45. Cardiology Division, University Hospital, Basel, Switzerland.

46. Population Health Research Institute, McMaster University, Hamilton, Canada.

47. Institute of Clinical Chemistry, Inselspital Bern, University Hospital, University of Bern, Bern, Switzerland.

48. Labormedizinisches Zentrum Dr. Risch , Vaduz, Liechtenstein.

49. Private University of the Principality of Liechtenstein, Triesen, Liechtenstein.

50. Department of Molecular Biology, Medical Biochemistry and Pathology, Laval University, Quebec City, Canada.

51. Department of Pediatrics, Tampere University Hospital, Tampere, Finland.

52. Department of Pediatrics, Faculty of Medicine and Health Technology, Tampere University, Tampere, Finland.

53. Department of Clinical Physiology and Nuclear Medicine, Turku University Hospital, Turku, Finland.

54. Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, Turku, Finland.

55. Centre for Population Health Research, University of Turku and Turku University Hospital, Turku, Finland.

56. CARIM School for Cardiovascular Diseases, Maastricht Center for Systems Biology (MaCSBio), Department. of Biochemistry, and Dept. of Physiology, Maastricht University, Maastricht, Netherlands.

57. Department of Medical Informatics Erasmus MC - University Medical Center Rotterdam, Rotterdam, the Netherlands.

58. Genetic Epidemiology Unit, Dept. of Epidemiology, Erasmus University Medical Center, Rotterdam, Netherlands.

59. Icelandic Heart Association, Kopavogur, Iceland.

60. Faculty of Medicine, University of Iceland, Reykjavik, Iceland.

61. Laboratory of Epidemiology and Population Sciences, National Institute on Aging, NIH, Baltimore, MD, USA.

62. School of Public Health, Department of Biostatistics, University of Michigan, Ann Arbor, MI, USA.

63. The Charles Bronfman Institute for Personalized Medicine, Icahn School of Medicine at Mount Sinai, New York, NY, USA.

64. The Mindich Child Health and Development Institute, Icahn School of Medicine at Mount Sinai, New York, NY 10029, New York, NY, USA.

65. Jackson Heart Study, Department of Medicine, University of Mississippi Medical Center, Jackson, MS, USA.

66. Department of Data Science, University of Mississippi Medical Center, Jackson, MS, USA.

67. Department of Physiology and Biophysics, University of Mississippi Medical Center, Jackson, MS, USA.

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

69. Institute of Human Genetics, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg, Germany.

70. Institute of Human Genetics, Klinikum rechts der Isar, Technische Universität München, Munich, Germany.

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

72. Chair of Genetic Epidemiology, IBE, Faculty of Medicine, LMU Munich, Munich, Germany.

73. Department of Internal Medicine I (Cardiology), Hospital of the Ludwig-Maximilians-University (LMU) Munich, Munich, Germany.

74. German Center for Diabetes Research, Neuherberg, Germany.

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

76. Research Unit of Molecular Epidemiology, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg, Germany.

77. Department of Twin Research and Genetic Epidemiology, Kings College London, London, UK.

78. NIHR Biomedical Research Centre at Guy's and St Thomas' Foundation Trust, London, UK.

79. Genomics plc, Oxford, UK.

80. Intermountain Heart Institute, Intermountain Medical Center, Murray, UT, USA.

81. Molecular Cardiology and Biophysics Division, Victor Chang Cardiac Research Institute, Darlinghurst, Australia.

82. Cardiology Department, St. Vincent's Hospital, Darlinghurst, Australia.

83. St Vincent's Clinical School, Faculty of Medicine, UNSW Sydney, Kensington, Australia.

84. Department of Cardiovascular Medicine, University of Iowa, Iowa City, Iowa, USA.

85. Laboratory for Molecular Cardiology, Department of Cardiology, The Heart Centre, Rigshospitalet, University Hospital of Copenhagen, Copenhagen, Denmark.

86. Department of Biomedical Sciences, University of Copenhagen, Copenhagen, Denmark.

87. Departments of Medicine, Pharmacology, and Biomedical Informatics, Vanderbilt University Medical Center, Nashville, TN, USA.

88. Department of Medicine, Vanderbilt University Medical Center, Nashville, TN, USA.

89. Department of Cardiology, Clinical Sciences; Wallenberg Center for Molecular Medicine; Lund University Diabetes Center; Lund University and Skane University Hospital, Lund, Sweden, Lund, Sweden.

90. Cardiovascular Health Research Unit, Department of Medicine, University of Washington, Seattle, WA, USA.

91. Department of Biostatistics, University of Washington, Seattle, WA, USA.

92. Department of Epidemiology, University of Washington, Seattle, WA, USA.

93. Kaiser Permanente Washington Health Research Institute, Seattle, WA, USA.

94. Cardiology Division, University of Washington, Seattle, WA, USA.

95. Robertson Center for Biostatistics, University of Glasgow, Glasgow, UK.

96. Department of Cardiology, Leiden University Medical Center, Leiden, the Netherlands.

97. Einthoven Laboratory for Experimental Vascular Medicine, Leiden University Medical Center, Leiden, the Netherlands.

98. Institute of Health and Wellbeing, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow, UK.

99. Department of Internal Medicine, section of Gerontology and Geriatrics, Leiden University Medical Center, Leiden, The Netherlands.

100. DZHK (German Centre for Cardiovascular Research); partner site Greifswald, Greifswald, Germany.

101. Department of Internal Medicine B - Cardiology, Pneumology, Infectious Diseases, Intensive Care Medicine; University Medicine Greifswald, Greifswald, Germany.

102. Interfaculty Institute for Genetics and Functional Genomics; Department of Functional Genonics; University Medicine and University of Greifswald, Greifswald, Germany.

103. National Institute for Health and Welfare, Helsinki, Finland.

104. Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland.

105. Institute for Translational Genomics and Population Sciences and Department of Pediatrics, Los Angeles Biomedical Research Institute at Harbor-UCLA Medical Center, Torrance, CA, USA.

106. Cardiovascular Health Research Unit and Department of Epidemiology, University of Washington, Seattle, WA, USA.

107. Institute for Translational Genomics and Population Sciences and Departments of Pediatrics and Medicine, Los Angeles Biomedical Research Institute at Harbor-UCLA Medical Center, Torrance, CA, USA.

108. Institute for Translational Genomics and Population Sciences, Los Angeles Biomedical Research Institute at Harbor-UCLA Medical Center, Torrance, CA, USA.

109. Department of Clinical Epidemiology, Leiden University Medical Center, Leiden, the Netherlands.

110. Department of Public Health and Primary Care, Leiden University Medical Center, Leiden, the Netherlands.

111. Department of Biomedical Sciences, University of Sassari, Sassari, Italy.

112. Laboratory of Genetics and Genomics, NIA/NIH , Baltimore, MD, USA.

113. Laboratory Of Cardiovascular Science, NIA/NIH, Baltimore, MD, USA.

114. Population Sciences Branch, Division of Intramural Research, National Heart, Lung, and Blood Institute, Bethesda, MD, USA.

115. National Heart Lung and Blood Institute's and Boston University's Framingham Heart Study, Framingham, MA, USA.

116. Section of Computational Biomedicine, Department of Medicine, Boston University School of Medicine, Boston, MA, USA.

117. Center for Human Genetic Research and Cardiovascular Research Center, Harvard Medical School and Massachusetts General Hospital, Boston, MA, USA.

118. Department of Biostatistics, Boston University School of Public Health, Boston, MA, USA.

119. The Institute of Medical Sciences, Aberdeen Biomedical Imaging Centre, University of Aberdeen, Aberdeen, UK.

120. Centre for Genomic and Experimental Medicine, Institute of Genetics & Molecular Medicine, University of Edinburgh, Western General Hospital, Edinburgh, UK.

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

122. Division of Population Health and Genomics, Ninewells Hospital and Medical School, University of Dundee, Dundee, UK.

123. Department of Epidemiology Erasmus MC - University Medical Center Rotterdam, Rotterdam, the Netherlands.

124. Human Genotyping Facility Erasmus MC University Medical Center Rotterdam, Rotterdam, the Netherlands.

125. Fred Hutchinson Cancer Research Center, Division of Pubic Health Sciences, Seattle, WA, USA.

126. Ohio State University, Division of Endocrinology, Diabetes and Metabolism, Columbus, OH, USA.

127. University of Washington, Department of Epidemiology, Seattle, WA, USA.

128. University of North Carolina, Gillings School of Global Public Health and School of Medicine, Departments of Epidemiology and Medicine, Chapel Hill, NC, USA.

129. Department of Epidemiology, Rollins School of Public Health, Emory University, Atlanta, GA, USA.

130. McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD, USA.

131. Human Genetics Center, University of Texas Health Science Center at Houston, Houston, TX, USA.

132. Cardiology, Geneva University Hospitals, Geneva, Switzerland.

133. Epidemiological Cardiology Research Center, Wake Forest School of Medicine, Winston-Salem, NC, USA.

134. Department of Epidemiology, University of North Carolina, Chapel Hill, NC, USA.

135. Carolina Population Center, University of North Carolina, Chapel Hill, NC, USA.

136. Division of Pharmaceutical Outcomes and Policy, University of North Carolina, Chapel Hill, NC, USA.

137. Barts Heart Centre, St Bartholomews Hospital, London, United Kingdom, London, UK.

138. Institute of Cardiovascular Science, University College London, London, UK.

139. Department of Mechanical Engineering, University College London, London, UK.

140. Department of Medicine, Landspitali University Hospital, Reykjavik, Iceland.

141. School of Engineering and Natural Sciences, University of Iceland, Reykjavik, Iceland.

142. Section of Cardiovascular Medicine and Section of Preventive Medicine, Department of Medicine, Boston University School of Medicine, Boston, MA, USA.

143. Department of Epidemiology, Boston University School of Public Health, Boston, MA, USA.

144. Cardiac Arrhythmia Service, Massachusetts General Hospital, Boston, MA, USA.

145. Genetics Research Centre, Molecular and Clinical Sciences Institute, St George's, University of London, London, UK.

* These authors contributed equally. # These authors jointly directed this work.

Correspondence to: slubitz@mgh.harvard.edu and p.b.munroe@qmul.ac.uk

Abstract

The electrocardiographic PR interval reflects atrioventricular conduction, and is associated with conduction abnormalities, pacemaker implantation, atrial fibrillation (AF), and cardiovascular mortality[1,2]. We performed multi-ancestry (N=293,051) and European only (N=271,570) genome-wide association (GWAS) meta-analyses for the PR interval, discovering 210 loci of which 149 are novel. Variants at all loci nearly doubled the percentage of heritability explained, from 33.5% to 62.6%. We observed enrichment for genes involved in cardiac muscle development/contraction and the cytoskeleton highlighting key regulation processes for atrioventricular conduction. Additionally, 19 novel loci harbour genes underlying inherited monogenic heart diseases suggesting the role of these genes in cardiovascular pathology in the general population. We showed that polygenic predisposition to PR interval duration is an endophenotype for cardiovascular disease risk, including distal conduction disease, AF, atrioventricular pre-excitation, non-ischemic cardiomyopathy, and coronary heart disease. These findings advance our understanding of the polygenic basis of cardiac conduction, and the genetic relationship between PR interval duration and cardiovascular disease.

Main text

The electrocardiogram is among the most common clinical tests ordered to assess cardiac abnormalities. Reproducible waveforms indicating discrete electrophysiologic processes were described over 100 years ago, yet the biological underpinnings of conduction and repolarization remain incompletely defined. The electrocardiographic PR interval reflects conduction from the atria to ventricles, across specialised conduction tissues such as the atrioventricular node and the His-Purkinje system. Pathological variation in the PR interval may indicate heart block or pre-excitation, both of which can lead to sudden death[2]. The PR interval also serves as a risk factor for AF and cardiovascular mortality[1-3]. Prior genetic association studies have identified 64 PR interval loci[4-13]. To enhance our understanding of the genetic and biological mechanisms of atrioventricular conduction, we performed GWAS meta-analyses of autosomal and X chromosome variants imputed mainly with the 1000 Genomes Project reference panel[14] using an additive model and increased sample size. Our primary meta-analysis included 293,051 individuals of European (92.6%), African (2.7%), Hispanic (4%), and Brazilian (<1%) ancestries from 40 studies (**Supplementary Tables 1-3**). We also performed ancestry-specific meta-analyses (**Fig. 1**).

We identified a total of 210 genome-wide significant loci ($P<5\times10^{-8}$), of which 149 were not previously reported (**Table 1, Fig. 2**). Of the 149 novel loci, 141 were discovered in the multi-ancestry analysis, and 8 additional novel loci were identified in the European ancestry analysis (**Table 1, Fig. 2, Supplementary Tables 4-5, Supplementary Fig. 1-4**). We considered only variants present in >60% of the maximum sample size, a filtering criterion used to ensure robustness of associated loci (**Online Methods**). There was strong support for all 64 previously reported loci (61 at $P<5\times10^{-8}$ and 3 at $P<1.1\times10^{-4}$; **Supplementary Tables 6-7**). No additional novel loci were identified in African or Hispanic/Latino ancestry meta-analyses (**Supplementary**

Table 8, Supplementary Fig. 1 and 3) or X chromosome meta-analyses (**Supplementary Fig. 5**). In secondary analyses, we examined the rank-based inverse normal transformed residuals of PR interval. Results of absolute and transformed trait meta-analyses were highly correlated (ρ>0.94, **Supplementary Tables 5, 9-10, Supplementary Fig. 6-7**).

By applying joint and conditional analyses in the European meta-analysis data, we identified multiple independently associated variants ($P_{joint}<5\times10^{-8}$ and $r^2<0.1$) at 12 novel and 25 previously reported loci (**Supplementary Table 11**). The overall variant-based heritability (h^2_g) for the PR interval estimated in 59,097 unrelated European participants from the UK Biobank (UKB) with electrocardiograms was 18.2% (**Online Methods**). In the UKB, the proportion of h^2_g explained by variation at all loci discovered in our analysis was 62.6%, compared to 33.5% when considering previously reported loci only.

The majority of the lead variants at the 149 novel loci were common (minor allele frequency, MAF>5%). We observed 6 low-frequency (MAF 1-5%) variants, and one rare (MAF<1%) predicted damaging missense variant (rs35816944, p.Ser171Leu) in *SPSB3* encoding SplA/Ryanodine Receptor Domain and SOCS Box-containing 3. SPSB3 is involved in degradation of SNAIL transcription factor, which regulates the epithelial-mesenchymal transition[15], and has not been previously associated with cardiovascular traits. In total, we identified missense variants in genes at 12 novel and 6 previously reported loci (**Supplementary Table 12**). At *MYH6*, a previously described locus for PR interval[6,10], sick sinus syndrome[16], AF and other cardiovascular traits[17], we observed a novel predicted damaging missense variant in *MYH6* (rs28711516, p.Gly56Arg). *MYH6* encodes the α-heavy chain subunit of cardiac myosin.

PR interval lead variants (or best proxy [r^2>0.8]) at 39 novel and 23 previously reported loci were significant cis-eQTLs (at a 5% false discovery rate (FDR) in left ventricle (LV) and right atrial appendage (RAA) tissue samples from the Genotype-Tissue Expression (GTEx) project[18] (**Supplementary Table 13)**. Variants at 21 novel loci were significant eQTLs in both tissues with consistent directionality of gene expression. We also performed a transcriptome-wide analysis to evaluate associations between predicted gene expression in LV and RAA with the PR interval. We identified 120 genes meeting our significance threshold ($P<4.4\times10^{-6}$, after Bonferroni correction); 26 genes were not localised at PR interval loci (≥500kb from a lead variant) representing potentially novel regions (**Supplementary Table 14, Supplementary Fig. 8**). Longer PR interval duration was associated with decreased levels of predicted gene expression for 61 genes, and increased levels for 59 genes (**Fig. 3**).

Most PR interval variants were annotated as non-coding. We therefore explored whether associated variants or proxies were located in transcriptionally active genomic regions. We observed enrichment for DNase I-hypersensitive sites in fetal heart tissue ($P<9.36\times10^{-5}$, **Supplementary Fig. 9**). Analysis of chromatin states indicated variants at 103 novel and 52 previously reported loci were located within regulatory elements that are present in heart tissues (**Supplementary Table 15**), providing support for gene regulatory mechanisms in specifying the PR interval. To identify distal candidate genes at PR interval loci, we assessed the same set of variants for chromatin interactions in a LV tissue Hi-C dataset[19]. Forty-eight target genes were identified (**Supplementary Table 16**). Variants at 38 novel loci were associated with other traits, including AF and coronary heart disease (**Supplementary Table 17, Supplementary Fig. 10**).

Candidate genes indicated by bioinformatics and *in silico* functional annotations at each novel locus are summarised in **Supplementary Tables 18-19,** and include 19 genes known to underlie

monogenic cardiovascular diseases. Enrichment analysis of genes at PR interval loci using DEPICT[20] indicated heart development ($P=1.87\times10^{-15}$) and actin cytoskeleton organisation ($P=2.20\times10^{-15}$) as the most significantly enriched processes (**Supplementary Table 20**). Ingenuity Pathway Analysis (IPA) supported heart development, ion channel signaling and cell-junction/cell-signaling amongst the most significant canonical pathways (**Supplementary Table 21**).

Finally, we evaluated associations between genetic predisposition to PR interval duration and 16 cardiac phenotypes chosen *a priori* using ~309,000 unrelated UKB European participants not included in our meta-analyses[21]. We created a polygenic risk score (PRS) for PR interval using the multi-ancestry meta-analysis results (**Fig. 4**, **Supplementary Table 22**). Genetically determined PR interval prolongation was associated with higher risk of distal conduction disease (atrioventricular block; odds ratio [OR] per standard deviation 1.11, $P=3.18\times10^{-8}$) and pacemaker implantation (OR 1.06, P=0.0005). In contrast, genetically determined PR interval prolongation was associated with reduced risk of AF (OR 0.94, $P=1.30\times10^{-11}$) and atrioventricular pre-excitation (Wolff-Parkinson-White syndrome; OR 0.83, $P=8.36\times10^{-4}$). Genetically determined PR interval prolongation was marginally associated with a reduced risk of non-ischemic cardiomyopathy (OR=0.95, P=0.046) and coronary heart disease (OR 0.99, P=0.035). Results were similar when using a PRS derived using the European ancestry meta-analysis results (**Supplementary Fig. 11, Supplementary Table 22**).

To summarise, in meta-analyses of nearly 300,000 individuals we identified 210 loci, of which 149 were novel, underlying cardiac conduction as manifested by the electrocardiographic PR interval. Apart from confirming well-established associations in loci harbouring ion-channel genes, our findings further underscore the central importance of heart development and

cytoskeletal components in atrioventricular conduction[10,12,13]. We also highlight the role of common variation at loci harboring genes underlying monogenic forms of heart disease in cardiac conduction.

We report signals in/near 13 candidate genes at novel loci with functional roles in cytoskeletal assembly (*DSP*, *DES*, *OBSL1*, *MYH11*, *PDLIM5*, *LDB3*, *FHL2*, *CEFIP*, *SSPN*, *TLN*, *PTK2*, *GJA5* and *CDH2*; **Fig. 5**). *DSP* and *DES* encode components of the cardiac desmosome, a complex involved in ionic communication between cardiomyocytes and maintenance of cellular integrity. Mutations in the desmosome are implicated in arrhythmogenic cardiomyopathy (ACM) and dilated cardiomyopathy (DCM)[22-26]. Conduction slowing is a major component of the pathophysiology of arrhythmia in ACM and other cardiomyopathies[27,28]. *OBSL1* encodes obscurin-like 1, which together with obscurin (OBSCN) is involved in sarcomerogenesis by bridging titin (TTN) and myomesin at the M-band[29]. *PDLIM5* encodes a scaffold protein that tethers protein kinases to the Z-disk, and has been associated with DCM in homozygous murine cardiac knockouts[30]. *FHL2* encodes calcineurin-binding protein four and a half LIM domains 2, which is involved in cardiac development by negatively regulating calcineurin/NFAT signaling in cardiomyocytes[31]. Missense mutations in *FHL2* have been associated with hypertrophic cardiomyopathy[32]. *CEFIP* encodes the cardiac-enriched FHL2-interacting protein located at the Z-disc, which interacts with *FHL2*. It is also involved in calcineurin–NFAT signaling, but its overexpression leads to cardiomyocyte hypertrophy[33].

Common variants in/near genes associated with inherited arrhythmia syndromes were also observed, suggesting these genes also affect atrioventricular conduction and cardiovascular pathology in the general population. Apart from *DSP*, *DES*, *MYH11* and *GJA5* listed above, our analyses indicate 15 additional candidate genes (*ADRB1*, *ALPK3*, *BMPR1*, *BMPR2*, *CRYAB*,

DERL3, DNAH11, DTNA, ETV1, HCN4, MYOZ2, PDE3A, RYR2, SPEG, LDB3) at novel loci causing Mendelian or other inherited forms of cardiovascular disease. Two genes we highlight are *HCN4* and *RYR2*. *HCN4* encodes a component of the hyperpolarization-activated cyclic nucleotide-gated potassium channel which specifies the sinoatrial pacemaker "funny" current, and is implicated in sinus node dysfunction, AF, and left ventricular noncompaction[34-36]. *RYR2* encodes a calcium channel component in the cardiac sarcoplasmic reticulum and is implicated in catecholaminergic polymorphic ventricular tachycardia[37].

Genes with roles in autonomic signaling in the heart (*CHRM2, ADCY5*) were indicated from expression analyses. *CHRM2* encodes the M2 muscarinic cholinergic receptors that bind acetylcholine and are expressed in the heart[38]. Their stimulation results in inhibition of adenylate cyclase encoded by *ADCY5*, which in turn inhibits ion channel function. Ultimately, the signaling cascade can result in reduced levels of the pacemaker "funny" current in the sinoatrial and atrioventricular nodes, reduced L-type calcium current in all myocyte populations, and increased inwardly rectifying $I_{\mathrm{K.Ach}}$ potassium current in the conduction tissues and atria causing cardiomyocyte hyperpolarization[39]. Stimulation has also been reported to shorten atrial action potential duration and thereby facilitate re-entry, which may lead to AF[40-42].

By constructing PRSs, we also observed that genetically determined PR interval duration is an endophenotype for several adult-onset complex cardiovascular diseases, the most significant of which are arrhythmias and conduction disorders. For example, our findings are consistent with previous epidemiologic data supporting a U-shaped relationship between PR interval duration and AF risk[1]. Although aggregate genetic predisposition to PR interval prolongation is associated with reduced AF risk, top PR interval prolonging alleles are associated with decreased AF risk (e.g., localized to the *SCN5A/SCN10A* locus) whereas others are associated with increased AF risk (e.g.,

localized to the *TTN* locus), consistent with prior reports[8]. These findings suggest that genetic determinants of the PR interval may identify distinct pathophysiologic mechanisms leading to AF, perhaps via specifying differences in tissue excitability, conduction velocity, or refractoriness. Future efforts are warranted to better understand the relations between genetically determined PR interval and specific arrhythmia mechanisms.

In conclusion, our study more than triples the reported number of PR interval loci, which collectively explain ~62% of trait-related heritability. Our findings highlight important biological processes underlying atrioventricular conduction which include both ion channel function, and specification of cytoskeletal components. Our study also indicates that common variation in Mendelian cardiovascular disease genes contributes to population-based variation in the PR interval. Lastly, we observed that genetic determinants of the PR interval provide novel insights into the etiology of several complex cardiac diseases, including AF. Collectively, our results represent a major advance in understanding the polygenic nature of cardiac conduction, and the genetic relationship between PR interval duration and arrhythmias.

Online Methods

Contributing studies

A total of 40 studies (**Supplementary Note**) comprising 293,051 individuals of European (N=271,570), African (N=8,173), Hispanic (N=11,686), and Brazilian (N=485) ancestries contributed GWAS summary statistics for PR interval. All participating institutions and co-ordinating centres approved this project, and informed consent was obtained from all study participants. Study-specific design, sample quality control and descriptive statistics are provided in **Supplementary Tables 1-3**. For the majority of the studies imputation was performed for autosomal chromosomes and X chromosome using the 1000 Genomes (1000G) project[14] reference panel or a most recently released haplotype version (**Supplementary Table 2**).

PR interval phenotype and exclusions

The PR interval was measured in milliseconds from standard 12-lead electrocardiograms (ECGs), except in the UK-Biobank in which it was obtained from 4-lead ECGs (CAM-USB 6.5, Cardiosoft v6.51) recorded during a 15 second rest period prior to an exercise test (**Supplementary Note**). We excluded individuals with extreme PR interval values (<80ms or >320ms), second/third degree heart block, AF on the ECG, or a history of myocardial infarction or heart failure, Wolff-Parkinson-White syndrome, pacemakers, receiving class I and class III antiarrhythmic medications, digoxin, and pregnancy.

Study-level association analyses

We regressed the absolute PR interval on each genotype dosage using multiple linear regression with an additive genetic effect and adjusted for age, sex, height, body mass index, heart rate and any other study specific covariates. To account for relatedness, linear mixed effects models were used for family studies. To account for population structure, analyses were also adjusted for principal components of ancestry derived from genotyped variants after excluding related individuals. Analyses of autosomal variants were conducted separately for each ancestry group. X chromosome analyses were performed separately for males and females. Analyses using rank-based inverse normal transformed residuals of PR interval corrected for the aforementioned covariates were also conducted. Residuals were calculated separately by ancestral group for autosomal variants, and separately for males and females for X chromosome variants.

Centralized quality control

We performed quality control centrally for each result file using EasyQC version 11.4[43]. We removed variants that were monomorphic, had a minor allele count (MAC) <6, imputation quality metric <0.3 (imputed by MACH) or 0.4 (imputed by IMPUTE2), had invalid or mismatched alleles, were duplicated, or if they were allele frequency outliers (difference >0.2 from the allele frequency in 1000G project). We inspected PZ plots, effect allele frequency plots, effect size distributions, QQ plots, and compared effect sizes in each study to effect sizes from prior reports for established PR interval loci to identify genotype and study level anomalies. Variants with effective MAC (=2×N×MAF×imputation quality metric) <10 were omitted from each study prior to meta-analysis.

Meta-analyses

We aggregated summary level associations between genotypes and absolute PR interval from all individuals (N=293,051), and only from Europeans (N=271,570), African Americans (N=8,173), and Hispanic/Latinos (N=12,823) using a fixed-effects meta-analysis approach implemented in METAL (release on 2011/03/25)[44]. For the X chromosome, meta-analyses were conducted in a sex-stratified fashion. Genomic control was applied (if inflation factor $\lambda_{GC}>1$) at the study level. Quantile–quantile (QQ) plots of observed versus expected $-\log_{10}(P)$ did not show substantive inflation (**Supplementary Figs. 1-2**).

Given the large sample size we undertook a one-stage discovery study design. To ensure the robustness of this approach we considered for further investigation only variants reaching genome-wide significance ($P<5\times10^{-8}$) present in at least 60% of the maximum sample size (N_{max}). We declared as novel any variants mapping outside the 64 loci previously reported (**Supplementary Note, Supplementary Table 6**). We grouped genome-wide significant variants into independent loci based on both distance (±500kb) and linkage disequilibrium (LD, $r^2<0.1$) (**Supplementary Note**). We assessed heterogeneity in allelic effect sizes among studies contributing to the meta-analysis and among ancestral groups by the I^2 inconsistency index[45] for the lead variant in each novel locus. LocusZoom[46] was used to create region plots of identified loci.

Meta-analyses (multi-ancestry [N=282,128], European only [N=271,570], and African [N=8,173]) of rank-based inverse normal transformed residuals of PR interval were also performed. Because not all studies contributed summary level association statistics of the transformed PR interval, we considered as primary the meta-analysis of absolute PR interval for

which we achieved the maximum sample size. Any loci that met our significance criteria in the meta-analyses of transformed PR interval were not taken forward for downstream analyses.

Conditional and heritability analysis

Conditional and joint GWAS analyses were implemented in GCTA v1.91.3[47] using summary level variant statistics from the European ancestry meta-analysis to identify independent association signals within PR interval loci. We used 59,097 unrelated (kinship coefficient >0.0884) UK Biobank participants of European ancestry as the reference sample to model patterns of LD between variants. We declared as conditionally independent any genome-wide significant variants in conditional analysis ($P_{joint}<5\times10^{-8}$) not in LD ($r^2<0.1$) with the lead variant in the locus.

Using the same set of individuals from UK Biobank, we estimated the aggregate genetic contributions to PR interval with restricted maximum likelihood as implemented in BOLT-REML[48]. We calculated the additive overall variant-heritability (h^2_g) based on 333,167 LD-pruned genotyped variants, as well as the h^2_g of variants at PR interval associated loci only. Loci windows were based on both distance (±500kb) and LD ($r^2>0.1$) around novel and previously reported variants (**Supplementary Note**). We then calculated the proportion of total h^2_g explained at PR interval loci by dividing the h^2_g estimate of PR interval loci by the total h^2_g.

Bioinformatics and in silico functional analyses

We use Variant Effect Predictor (VEP)[49] to obtain functional characterization of variants including consequence, information on nearest genes and, where applicable, amino acid substitution and functional impact , based on SIFT[50] and PolyPhen-2[51] prediction tools. For non-coding variants, we assessed overlap with DNase I–hypersensitive sites (DHS) and chromatin states as determined by Roadmap Epigenomics Project [52] across all tissues and in cardiac tissues (E083, fetal heart; E095, LV; E104, right atrium; E105, right ventricle) using HaploReg v4[53].

We assessed whether any PR interval variants were related to cardiac gene expression using GTEx[18] version 7 cis-eQTL LV (N=272) and RAA (N=264) data. If the variant at a locus was not available in GTEx, we used proxy variants (r^2>0.8). We report results only for associations at a false discovery rate (FDR) of 5%. We then evaluated the effects of predicted gene expression levels on PR interval duration using S-PrediXcan[54]. GTEx[18] genotypes (variants with MAF>0.01) and normalized expression data in LV and RAA provided by the software developers were used as the training datasets for the prediction models. The prediction models between each gene-tissue pair were performed by Elastic-Net, and only significant (FDR 5%) models for prediction were included in our analysis. We used the European meta-analysis summary-level results (variants with at least 60% of maximum sample size) as the study dataset and then performed the S-PrediXcan calculator to estimate the expression-PR interval associations. In total, we tested 5,366 and 5,977 associations in LV and RAA, respectively. Significance threshold was set at $P=4.4\times10^{-6}$ (=0.05/(5,977+5,366)) to account for multiple testing corrections.

We applied GARFIELD (GWAS analysis of regulatory or functional information enrichment with LD correction)[55] to analyse the enrichment patterns for functional annotations of the European

meta-analysis summary statistics, using regulatory maps from the Encyclopedia of DNA Elements (ENCODE)[56] and Roadmap Epigenomics[52] projects. This method calculates odds ratios and enrichment P-values at different GWAS P-value thresholds (denoted T) for each annotation by using a logistic regression model accounting for LD, matched genotyping variants and local gene density with the application of logistic regression to derive statistical significance. Threshold for significant enrichment was set to $P=9.36\times10^{-5}$ (after multiple-testing correction for the number of effective annotations).

We identified potential target genes of regulatory variants using long-range chromatin interaction (Hi-C) data from the LV[19]. Hi-C data was corrected for genomic biases and distance using the Hi-C Pro and Fit-Hi-C pipelines according to Schmitt *et al.* (40kb resolution – correction applied to interactions with 50kb-5Mb span). We identified the promoter interactions for all potential regulatory variants in LD ($r^2>0.8$) with our lead and conditionally independent PR interval variants and report the interactors with the variants with the highest regulatory potential (RegulomeDB$\geq$2) to annotate the loci.

We performed a literature review, and queried the Online Mendelian Inheritance in Man (OMIM) and the International Mouse Phenotyping Consortium databases for all genes in regions defined by $r^2>0.5$ from the lead variant at each novel locus. We further expanded the gene listing with any genes that were indicated by gene expression or chromatin interaction analyses. We performed look-ups for each lead variant or their proxies ($r^2>0.8$) for associations ($P<5\times10^{-8}$) for common traits using both GWAS catalog[57] and PhenoScanner v2[58] databases. For AF, we supplemented the variant listing with a manually curated list of all overlapping variants ($r^2>0.7$) with PR interval from two recently published GWASs[59,60].

583

584 ***Gene set enrichment and pathway analyses***

585 We used DEPICT (Data-driven Expression-Prioritized Integration for Complex Traits)[20] to
586 identify enriched pathways and tissues/cell types where genes from associated loci are highly
587 expressed using all genome-wide significant ($P<5\times10^{-8}$) variants in our multi-ancestry meta-
588 analysis present in at least 60% of N_{max} (N=20,076). To identify uncorrelated variants for PR
589 interval, DEPICT performed LD-clumping (r^2=0.1, window size=250kb) using LD estimates
590 between variants from the 1000G reference data on individuals from all ancestries after excluding
591 the major histocompatibility complex region on chromosome 6. Gene-set enrichment analysis was
592 conducted based on 14,461 predefined reconstituted gene sets from various databases and data
593 types, including Gene ontology, Kyoto encyclopedia of genes and genomes (KEGG),
594 REACTOME, phenotypic gene sets derived from the Mouse genetics initiative, and protein
595 molecular pathways derived from protein-protein interaction. Finally, tissue and cell type
596 enrichment analysis was performed based on expression information in any of the 209 Medical
597 Subject Heading (MeSH) annotations for the 37,427 human Affymetrix HGU133a2.0 platform
598 microarray probes.

599 Ingenuity Pathway Analysis (IPA) was conducted using an extended list comprising 593 genes
600 located in regions defined by r^2>0.5 with the lead or conditionally independent variants for all PR
601 interval loci, or the nearest gene. We further expanded this list by adding genes indicated by gene
602 expression analyses. Only molecules and/or relationships for human or mouse or rat and
603 experimentally verified results were considered. The significance P-value associated with
604 enrichment of functional processes is calculated using the right-tailed Fisher's exact test by

considering the number of query molecules that participate in that function and the total number of molecules that are known to be associated with that function in the IPA.

Associations between genetically determined PR interval and cardiovascular conditions

We examined associations between genetic determinants of atrioventricular conduction and candidate cardiovascular diseases in unrelated individuals of European ancestry from UK Biobank (N~309,000 not included in our GWAS meta-analyses) by creating PRSs for PR interval based on our GWAS results. We derived two PRSs. One was derived from the multi-ancestry meta-analysis results, and the other from the European meta-analysis results. We used the LD-clumping feature in PLINK v1.90[61] (r^2=0.1, window size=250kb, $P=5\times10^{-8}$) to select variants for each PRS. Referent LD structure was based on 1000G all ancestry, and European only data. In total, we selected 743 and 582 variants from multi-ancestry and European only meta-analysis results, respectively. We calculated the PRSs for PR interval by summing the dosage of PR interval prolonging alleles weighted by the corresponding effect size from the meta-analysis results. A total of 743 variants for the PRS derived from multi-ancestry results and 581 variants for the PRS derived from European results (among the variants with imputation quality >0.6) were included in our PRS calculations.

We selected candidate cardiovascular conditions *a priori*, which included various cardiac conduction and structural traits such as bradyarrhythmia, AF, atrioventricular pre-excitation, heart failure, cardiomyopathy, and congenital heart disease. We ascertained disease status based on data from baseline interviews, hospital diagnosis codes (ICD-9 and ICD-10), cause of death codes

(ICD-10), and operation codes. Details of individual selections and disease definitions are described in **Supplementary Table 23**.

We tested the PRSs for association with cardiovascular conditions using logistic regression. We adjusted for enrolled age, sex, genotyping array, and phenotype-related principal components of ancestry. Given correlation between traits, we did not establish a pre-specified significance threshold for the analysis and report nominal associations ($P<0.05$).

Tables

Table 1 Novel genome-wide significant loci associated with PR interval (N = 149).

Locus ID	Nearest gene(s)	rsID	Chr	Position	EA/OA	EAF	N	Beta	SE	P
Multi-ancestry meta-analysis										
1	*HSPG2*	rs7529220	1	22282619	C/T	0.84	293050	0.58	0.09	2.1×10^{-10}
2	*AIM1L*	rs12031946	1	26679041	C/T	0.13	293051	0.59	0.10	2.4×10^{-9}
3	*MAP7D1*	rs1475267	1	36620801	G/C	0.16	293050	0.50	0.09	2.1×10^{-8}
4	*EDN2*	rs12751675	1	41955714	G/A	0.75	293050	0.56	0.08	1.8×10^{-13}
5	*SSBP3*	rs603901	1	54741767	T/C	0.58	293051	0.43	0.06	3.3×10^{-11}
6	*NFIA*	rs6587924	1	61895257	A/C	0.49	293051	0.35	0.06	2.7×10^{-8}
7	*CDC7*	rs13447455	1	91966445	A/G	0.64	293051	0.38	0.07	1.7×10^{-8}
8	*GJA5*	rs1692144	1	147281349	C/T	0.79	293051	0.65	0.08	1.7×10^{-16}
9	*DPT*	rs531706	1	168692137	C/G	0.28	293051	0.39	0.07	3.4×10^{-8}
10	*PRRX1*	rs61824886	1	170615660	C/G	0.85	293051	0.67	0.09	6.2×10^{-13}
11	*C1orf98*	rs819636	1	200271408	C/T	0.33	293051	0.38	0.07	1.7×10^{-8}
12	*HLX*	rs6678632	1	221138612	T/C	0.44	293051	0.47	0.06	4.9×10^{-13}
13	*ADCK3*	rs3768419	1	227173477	C/G	0.48	291546	0.49	0.06	1.6×10^{-14}
14	*SIPA1L2*	rs1285678	1	232712145	A/G	0.47	287628	0.52	0.07	2.6×10^{-15}
15	*RYR2*	rs10802580	1	237194922	G/A	0.76	286413	0.45	0.08	6.9×10^{-9}
16	*SMYD3*	rs28468565	1	246157144	A/G	0.66	287628	0.49	0.07	8.9×10^{-13}
17	*LINC01249*	rs12616546	2	4824622	A/G	0.68	293051	0.53	0.07	9.2×10^{-15}
18	*STRN*	rs17496249	2	37102249	A/G	0.55	293051	0.64	0.06	1.4×10^{-23}
19	*EML4*	rs6728830	2	42537995	C/A	0.96	291132	1.02	0.18	1.9×10^{-8}
20	*EPAS1*	rs11894252	2	46533376	T/C	0.42	293047	0.45	0.06	2.3×10^{-12}
21	*FBXO11*	rs7588761	2	48150587	T/C	0.07	288153	0.75	0.13	5.3×10^{-9}
22	*SPTBN1*	rs4519566	2	54824815	G/A	0.79	293051	0.54	0.08	4.7×10^{-12}
23	*LINC01812/C1D*	rs7584373	2	68079211	A/G	0.35	293051	0.38	0.07	1.5×10^{-8}
24	*FHL2/LOC285000*	rs13006682	2	106104856	C/T	0.34	293051	0.51	0.07	6.7×10^{-14}
25	*NCKAP5*	rs17816356	2	134326085	A/C	0.05	289723	0.96	0.16	7.9×10^{-10}
26	*TEX41*	rs76909456	2	145453968	G/A	0.24	293051	0.48	0.08	1.8×10^{-10}
27	*LINC01473/ZC3H15*	rs138711926	2	187033804	G/A	0.04	280792	0.98	0.18	4.7×10^{-8}
28	*SDPR*	rs58577564	2	192723128	A/T	0.10	291546	0.78	0.11	7.5×10^{-13}

Locus ID	Nearest gene(s)	rsID	Chr	Position	EA/OA	EAF	N	Beta	SE	P
Multi-ancestry meta-analysis										
31	*TMEM198*	rs13023533*	2	220414019	T/C	0.55	293051	0.41	0.06	1.1×10^{-10}
32	*LSM3*	rs6442433	3	14275759	C/G	0.79	291836	0.71	0.08	1.9×10^{-19}
33	*THRB*	rs60325252	3	24465080	C/T	0.71	293051	0.75	0.07	9.1×10^{-27}
34	*TRAK1*	rs11921457	3	42103045	T/G	0.81	292301	0.69	0.09	8.0×10^{-16}
35	*LAMB2*	rs9865051*	3	49166069	T/C	0.78	293051	0.49	0.08	3.7×10^{-10}
36	*ADCY5*	rs1000368	3	123117165	T/C	0.26	293051	0.43	0.07	2.5×10^{-9}
37	*TSC22D2*	rs201481721	3	150176904	D/I	0.03	223845	1.32	0.22	1.8×10^{-9}
38	*RAP2B*	rs4680046	3	153000092	T/C	0.49	293051	0.47	0.06	1.2×10^{-13}
40	*FNDC3B*	rs4894803	3	171800256	G/A	0.39	293046	0.49	0.07	1.8×10^{-13}
41	*FGF12*	rs4687352	3	192373761	A/C	0.41	293051	0.53	0.06	1.3×10^{-16}
42	*DLG1*	rs143879787	3	196799232	I/D	0.73	226107	0.51	0.08	1.4×10^{-9}
43	*SRD5A3*	rs77422711	4	56123105	A/G	0.02	273824	1.85	0.30	6.4×10^{-10}
44	*LPHN3*	rs28540500	4	62409801	C/G	0.38	293051	0.42	0.07	2.4×10^{-10}
45	*FGF5*	rs36034102	4	81202048	T/G	0.27	292217	0.43	0.07	3.5×10^{-9}
46	*PDLIM5*	rs2172448	4	95506214	A/G	0.55	288153	0.37	0.06	9.4×10^{-9}
48	*SLC12A7*	rs4975572	5	1054197	T/C	0.46	293051	0.62	0.07	5.5×10^{-21}
49	*SUB1*	rs17441816	5	32629419	G/A	0.29	293051	0.51	0.07	4.7×10^{-13}
50	*HCN1*	rs10039283	5	45864843	A/G	0.41	293051	0.64	0.06	3.2×10^{-23}
51	*NR2F1*	rs4869412	5	92455655	G/A	0.49	293051	0.40	0.06	2.3×10^{-10}
53	*STARD4*	rs67968533	5	111046342	C/T	0.09	293051	0.66	0.11	4.8×10^{-9}
54	*LOC101927421/ ZNF608*	rs12654442	5	124343851	T/C	0.27	293051	0.46	0.07	2.8×10^{-10}
55	*SLC27A6*	rs2577531	5	128299279	C/T	0.59	293051	0.38	0.06	3.8×10^{-09}
56	*FGF18*	rs78810186	5	170868622	T/C	0.11	290821	0.74	0.10	1.4×10^{-12}
57	*LINC01411*	rs4868384	5	173779209	T/A	0.47	290336	0.47	0.06	2.5×10^{-13}
58	*DSP*	rs72825038	6	7527269	A/G	0.09	293051	0.94	0.11	2.7×10^{-16}
59	*DEK*	rs214502	6	18227546	A/C	0.58	291546	0.42	0.07	9.9×10^{-11}
60	*HDGFL1*	rs6922960	6	22570189	C/T	0.28	291546	0.61	0.07	5.6×10^{-18}
61	*LRRC16A*	rs139915396	6	25351477	I/D	0.10	226107	0.69	0.13	4.2×10^{-8}
62	*CDKN1A*	rs730506	6	36645968	C/G	0.20	293051	0.62	0.08	6.5×10^{-15}
63	*TFEB*	rs1015149	6	41658889	T/C	0.47	293051	0.45	0.06	1.6×10^{-12}
64	*RCAN2*	rs871728	6	46452619	C/T	0.42	293051	0.52	0.07	1.3×10^{-15}

Locus ID	Nearest gene(s)	rsID	Chr	Position	EA/OA	EAF	N	Beta	SE	P
Multi-ancestry meta-analysis										
65	*LOC101927686*	rs111739590	6	113978255	C/T	0.81	293051	0.50	0.08	1.5×10^{-9}
66	*TCF21*	rs12190287	6	134214525	G/C	0.37	290979	0.43	0.07	2.9×10^{-10}
67	*RP1-155D22.1*	rs206708	6	164532059	A/T	0.70	293051	0.47	0.07	1.9×10^{-11}
68	*GET4*	rs10226357	7	925949	G/A	0.59	293050	0.39	0.07	3.1×10^{-9}
69	*DGKB*	rs56352403	7	14453835	G/A	0.64	291623	0.56	0.07	1.4×10^{-16}
70	*PRPS1L1*	rs6961768	7	18040476	A/C	0.43	293051	0.38	0.06	3.0×10^{-9}
71	*DNAH11*	rs62441680*	7	21622494	C/T	0.17	293051	0.62	0.08	2.3×10^{-13}
72	*ELMO1*	rs4720244	7	37398113	C/G	0.64	293051	0.42	0.07	3.6×10^{-10}
73	*SEMA3A*	rs62472627	7	83998676	C/T	0.14	293051	0.61	0.09	6.7×10^{-11}
74	*CHRM2*	rs1424569	7	136569416	C/T	0.53	293051	0.36	0.07	4.6×10^{-8}
75	*DLC1*	rs1188285	8	13130478	C/T	0.56	293050	0.45	0.06	2.4×10^{-12}
76	*MTUS1*	rs4921804	8	17550623	G/A	0.63	289672	0.39	0.07	4.4×10^{-9}
77	*XPO7*	rs56317071	8	21775838	C/G	0.12	293050	0.57	0.10	1.2×10^{-8}
78	*RBPMS*	rs4545054	8	30302465	C/T	0.49	293050	0.36	0.06	2.0×10^{-8}
80	*RP11-1082L8.3*	rs35006907	8	125859817	A/C	0.31	293050	0.48	0.07	1.6×10^{-12}
81	*PTK2*	rs10106406	8	142006198	C/G	0.45	282729	0.40	0.07	1.4×10^{-9}
82	*TRPM3*	rs6560168	9	73482647	T/A	0.45	292407	0.45	0.06	2.2×10^{-12}
83	*SPATA31D5P/ RASEF*	rs7043482	9	85135915	A/C	0.65	293050	0.42	0.07	8.3×10^{-10}
84	*ASTN2*	rs1407243	9	119314851	C/T	0.60	293051	0.37	0.06	9.1×10^{-9}
85	*PLPP7*	rs4584185	9	134203545	C/T	0.45	278484	0.48	0.07	6.1×10^{-13}
86	*BEND7*	rs7916672	10	13534234	T/C	0.58	293051	0.35	0.06	3.8×10^{-8}
87	*CCDC7*	rs2947080	10	32847962	G/C	0.64	293051	0.41	0.07	4.6×10^{-10}
88	*CEFIP*	rs10776558*	10	50510406	C/T	0.53	293050	0.42	0.06	2.8×10^{-11}
89	*TMEM26*	rs74813029	10	63194576	A/G	0.17	293051	0.54	0.09	3.3×10^{-10}
90	*COL13A1*	rs2642608	10	71559723	T/C	0.27	293051	0.42	0.07	3.8×10^{-9}
91	*ZMIZ1*	rs1769758	10	80898969	T/G	0.50	267464	0.50	0.07	1.0×10^{-11}
92	*U3*	rs117443987	10	88509088	T/A	0.92	288153	1.00	0.12	6.1×10^{-17}
93	*ADRB1*	rs67234920	10	115782061	G/A	0.89	293051	0.67	0.11	2.3×10^{-10}
94	*FGFR2*	rs2912774	10	123348662	T/G	0.42	288153	0.48	0.07	1.3×10^{-13}
95	*MPPED2*	rs553951	11	30432176	C/T	0.73	293051	0.39	0.07	3.8×10^{-8}
96	*WT1*	rs11031737	11	32372772	G/A	0.52	293051	0.35	0.06	4.5×10^{-8}

Locus ID	Nearest gene(s)	rsID	Chr	Position	EA/OA	EAF	N	Beta	SE	P
Multi-ancestry meta-analysis										
97	*PCNXL3*	rs12801636	11	65391317	A/G	0.24	293051	0.41	0.07	4.1×10^{-8}
98	*CRYAB*	rs12808601	11	111776066	G/A	0.70	293050	0.51	0.07	3.1×10^{-13}
99	*USP28*	rs144789148	11	113666335	G/A	0.05	290495	0.96	0.17	6.4×10^{-9}
100	*PDE3A*	rs10770646	12	20544361	T/C	0.79	288153	0.53	0.08	6.4×10^{-11}
101	*SSPN*	rs78518764	12	26306484	T/C	0.86	293051	0.61	0.10	2.1×10^{-10}
102	*ARID2*	rs76611452	12	46209520	T/C	0.04	286416	1.19	0.19	1.5×10^{-10}
103	*SRGAP1*	rs17099893	12	64283014	A/G	0.06	291836	0.95	0.14	9.9×10^{-12}
104	*MIR6074*	rs4026608	12	66394664	T/C	0.62	293051	0.38	0.07	6.5×10^{-9}
105	*SLC6A15*	rs10862858	12	84806298	A/G	0.43	293051	0.46	0.06	1.1×10^{-12}
106	*HCFC2*	rs2629745	12	104503806	A/G	0.88	293051	0.69	0.10	1.7×10^{-12}
107	*RIC8B*	rs3759310	12	107166122	G/C	0.36	293051	0.56	0.07	4.5×10^{-17}
108	*UBE3B*	rs2004359*	12	109976893	G/T	0.47	291836	0.42	0.06	5.0×10^{-11}
109	*TESC*	rs7972416	12	117491824	A/G	0.66	293051	0.45	0.07	1.8×10^{-11}
110	*FREM2*	rs9634754*	13	39261151	G/T	0.69	293051	0.41	0.07	4.4×10^{-9}
111	*FGF14*	rs9513995	13	102878269	T/C	0.74	287628	0.50	0.07	2.1×10^{-11}
112	*ARHGEF40*	rs12885183	14	21545230	G/A	0.22	283907	0.49	0.08	8.5×10^{-10}
113	*RP11-562L8.1*	rs7146955	14	29750244	G/A	0.59	293051	0.44	0.06	6.7×10^{-12}
114	*AKAP6*	rs3784192	14	32923336	A/G	0.20	293051	0.55	0.08	3.0×10^{-11}
115	*NFKBIA*	rs8904	14	35871217	G/A	0.63	287252	0.40	0.07	1.6×10^{-9}
116	*SYNE2*	rs1255908	14	64457638	T/G	0.69	291546	0.52	0.07	6.0×10^{-14}
117	*FLRT2*	rs17712080	14	86041160	G/A	0.75	293050	0.47	0.07	2.0×10^{-10}
118	*RP11-1070N10.3*	rs179145	14	95983975	A/G	0.38	287627	0.41	0.07	5.2×10^{-10}
119	*MARK3*	rs3759579	14	103851272	A/G	0.41	287627	0.42	0.07	1.2×10^{-10}
120	*RBPMS2*	rs3935716*	15	65035979	A/G	0.15	293051	0.61	0.10	1.2×10^{-10}
121	*CORO2B*	rs11330601	15	69021265	I/D	0.55	222818	0.43	0.08	1.4×10^{-8}
122	*HCN4*	rs8039168	15	73664723	A/T	0.83	293051	0.60	0.09	2.0×10^{-12}
123	*ALPK3*	rs6496452	15	85372645	A/T	0.55	287628	0.55	0.06	1.3×10^{-17}
124	*LINC00924/NR2F2*	rs62008078	15	96460899	C/T	0.44	285649	0.48	0.07	4.4×10^{-13}
125	*SPSB3*	rs35816944*	16	1828030	G/A	0.99	247100	2.70	0.44	1.3×10^{-9}
126	*SRL*	rs79321945	16	4282284	C/A	0.78	293051	0.50	0.08	2.1×10^{-10}
128	*LOC101927480/ LINC02140*	rs1186818	16	54598337	G/A	0.24	293051	0.43	0.07	9.0×10^{-9}

Locus ID	Nearest gene(s)	rsID	Chr	Position	EA/OA	EAF	N	Beta	SE	P
Multi-ancestry meta-analysis										
129	*CNOT1*	rs7199856*	16	58584772	G/T	0.26	292217	0.56	0.07	8.6×10^{-15}
130	*LINC01082/IRF8*	rs904199	16	86184639	G/A	0.08	287629	0.73	0.12	7.9×10^{-10}
131	*ZFPM1*	rs28634651	16	88553198	T/C	0.61	261197	0.51	0.07	2.1×10^{-12}
132	*MINK1*	rs7774	17	4801163	A/C	0.33	293051	0.43	0.07	8.6×10^{-10}
133	*EFCAB5*	rs55866125*	17	28312993	T/C	0.52	293051	0.42	0.06	3.2×10^{-11}
134	*CACNA1G*	rs757416	17	48666064	T/C	0.63	293051	0.69	0.07	2.5×10^{-25}
135	*CSHL1*	rs2006122	17	61987405	T/A	0.27	293051	0.42	0.07	5.8×10^{-9}
136	*PRKCA*	rs9909004	17	64306133	C/T	0.42	291623	0.38	0.06	2.9×10^{-9}
138	*AC100791.2*	rs745570*	17	77781725	G/A	0.53	293051	0.35	0.06	3.4×10^{-8}
139	*CDH2*	rs11083300	18	26339589	G/C	0.46	293050	0.35	0.06	3.8×10^{-8}
140	*GAREM*	rs982521	18	30029141	C/T	0.18	293050	0.75	0.09	1.7×10^{-18}
141	*DTNA*	rs1786595	18	32399259	C/T	0.74	293050	0.47	0.07	6.9×10^{-11}
142	*CCBE1*	rs12961264	18	57138957	C/T	0.23	293050	0.51	0.08	1.1×10^{-11}
143	*STK11*	rs3795063	19	1217560	C/G	0.65	268324	0.49	0.08	7.3×10^{-11}
144	*ZNF358*	rs113394178	19	7581244	A/C	0.60	212667	0.46	0.08	2.1×10^{-8}
145	*TMEM59L*	rs111551996	19	18733355	G/T	0.95	290902	0.95	0.16	8.0×10^{-10}
146	*RNF24/SMOX*	rs16989138	20	4031653	G/A	0.43	291546	0.53	0.07	7.1×10^{-16}
147	*KIAA1755*	rs6023939	20	36832526	C/A	0.54	293051	0.44	0.06	6.7×10^{-12}
148	*DERL3*	rs2070464	22	24183875	G/A	0.38	291836	0.54	0.07	6.2×10^{-16}
149	*PHF5A*	rs9607805	22	41854446	T/C	0.70	287628	0.42	0.07	5.6×10^{-9}
European meta-analysis										
29	*BMPR2*	rs2103208	2	203373030	G/A	0.49	271570	0.36	0.07	4.8×10^{-8}
30	*AC007563.5*	rs6435953	2	217628087	C/T	0.16	271570	0.51	0.09	2.1×10^{-8}
39	*MLF1*	rs6799180*	3	158333891	A/G	0.47	271570	0.37	0.07	2.2×10^{-8}
47	*MYOZ2*	rs78277783	4	120070079	A/T	0.27	266672	0.42	0.08	2.1×10^{-8}
52	*FER*	rs6889995	5	108210304	G/A	0.22	266672	0.44	0.08	4.3×10^{-8}
79	*AZIN1*	rs565720	8	103914366	A/C	0.77	271570	0.44	0.08	3.1×10^{-8}
127	*MYH11*	rs72772025	16	15834729	T/C	0.27	269591	0.41	0.08	3.5×10^{-8}
137	*CASKIN2*	rs7501873	17	73505172	G/A	0.22	271570	0.45	0.08	4.9×10^{-8}

There was no evidence of heterogeneity for any of the newly identified loci across individual studies ($P_{heterogeneity} \geq 0.001$) or across ancestry groups ($P_{heterogeneity} > 0.01$).

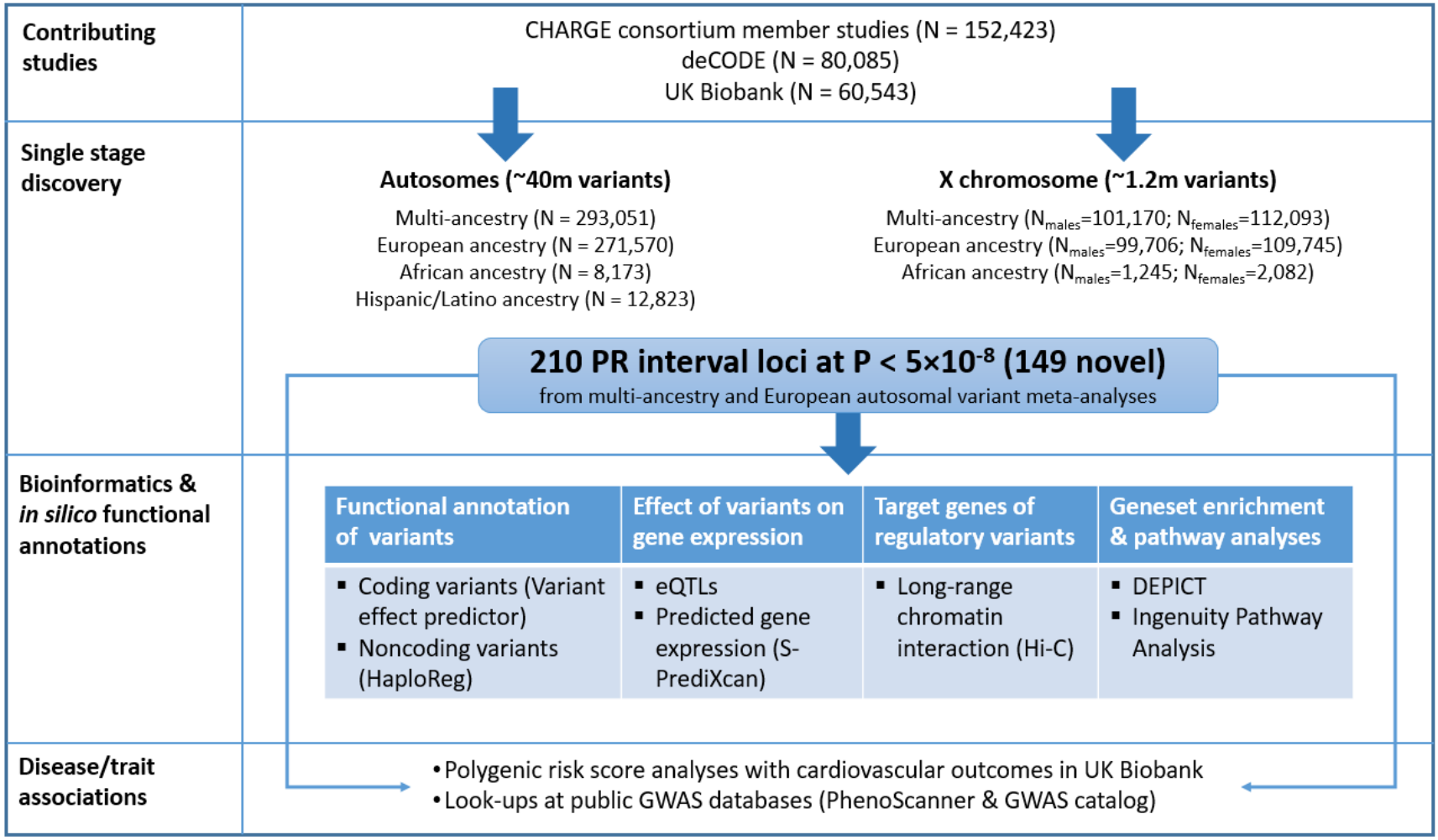
Locus ID: unique locus identifier; Nearest gene(s): Nearest annotated gene(s) to the lead variant; rsID, variant accession number; Chr, chromosome; Position, physical position in build 37; EA, effect allele; OA, other allele; EAF, effect allele frequency; N, total sample size analyzed; beta, effect estimate is milliseconds; SE, standard error; P, P-value.

* Missense variant or variant in high LD ($r^2 > 0.8$) with missense or splice site variant(s).

Figures

Figure 1 Overview of the study design.

Figure includes overview of contributing studies, single-stage discovery approach, and downstream bioinformatics and *in silico* annotations we performed to link variants to genes, and polygenic risk score analysis to link variants to cardiovascular disease risk.


Contributing studies
CHARGE consortium member studies (N = 152,423)
deCODE (N = 80,085)
UK Biobank (N = 60,543)
Single stage discovery
Autosomes (~40m variants)
Multi-ancestry (N = 293,051)
European ancestry (N = 271,570)
African ancestry (N = 8,173)
Hispanic/Latino ancestry (N = 12,823)
X chromosome (~1.2m variants)
Multi-ancestry (N_{males}=101,170; $N_{females}$=112,093)
European ancestry (N_{males}=99,706; $N_{females}$=109,745)
African ancestry (N_{males}=1,245; $N_{females}$=2,082)
210 PR interval loci at P < 5×10^{-8} (149 novel)
from multi-ancestry and European autosomal variant meta-analyses
Bioinformatics & *in silico* functional annotations
Functional annotation of variants
Coding variants (Variant effect predictor)
Noncoding variants (HaploReg)
Effect of variants on gene expression
eQTLs
Predicted gene expression (S-PrediXcan)
Target genes of regulatory variants
Long-range chromatin interaction (Hi-C)
Geneset enrichment & pathway analyses
DEPICT
Ingenuity Pathway Analysis
Disease/trait associations
Polygenic risk score analyses with cardiovascular outcomes in UK Biobank
Look-ups at public GWAS databases (PhenoScanner & GWAS catalog)

648

649

Figure 2 Manhattan plot of the multi-ancestry meta-analysis for PR interval. P values are plotted on the -$\log_{10}$ scale for all variants present in at least 60% of the maximum sample size. Associations of genome-wide significant ($P < 5 \times 10^{-8}$) variants at novel (N = 141) and previously reported loci (N = 61) are plotted in dark and light blue colours respectively.

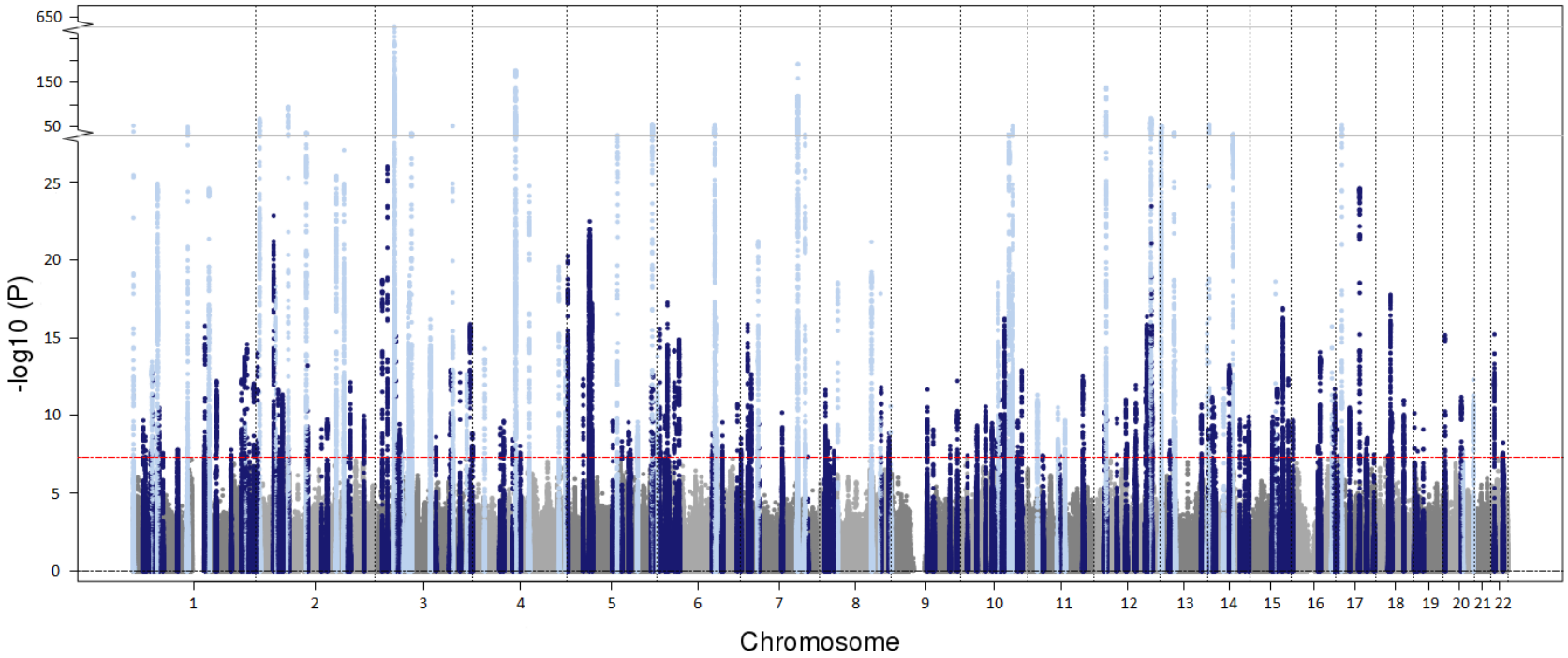
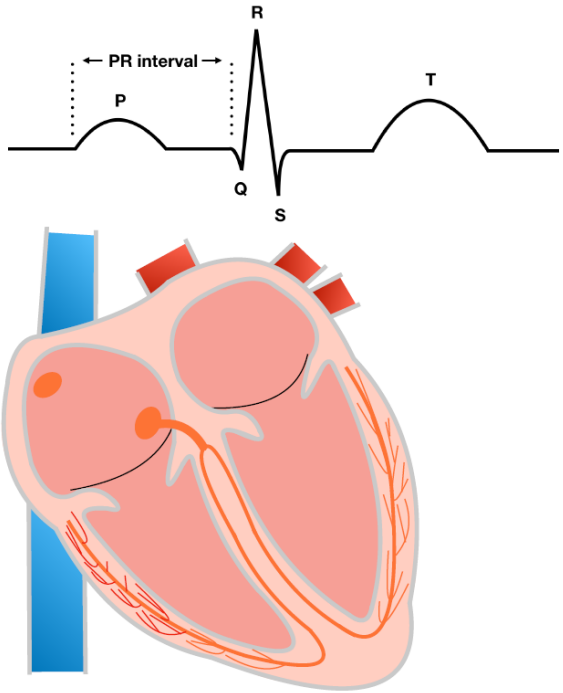


Figure 3 Plausible candidate genes of PR interval from S-PrediXcan

Diagram of standard electrocardiographic intervals and the heart. The electrocardiographic features are illustratively aligned with the corresponding cardiac conduction system structures (orange) reflected on the tracing. The PR interval (labeled) indicates conduction through the atria, atrioventricular node, His bundle, and Purkinje fibers. Right: The tables show 120 genes whose expression in the left ventricle (N=272) or right atrial appendage (N=264) was associated with PR interval duration in a transcriptome-wide analysis using S-PrediXcan and GTEx v7. Displayed genes include those with significant associations after Bonferroni correction for all tested genes at the two tissues with a $P < 4.4\times10^{-6}$ (=0.05/(5,977+5,366)). Longer PR intervals were associated with increased predicted expression of 59 genes (blue) and reduced expression of 61 genes (orange).



Increased gene expression

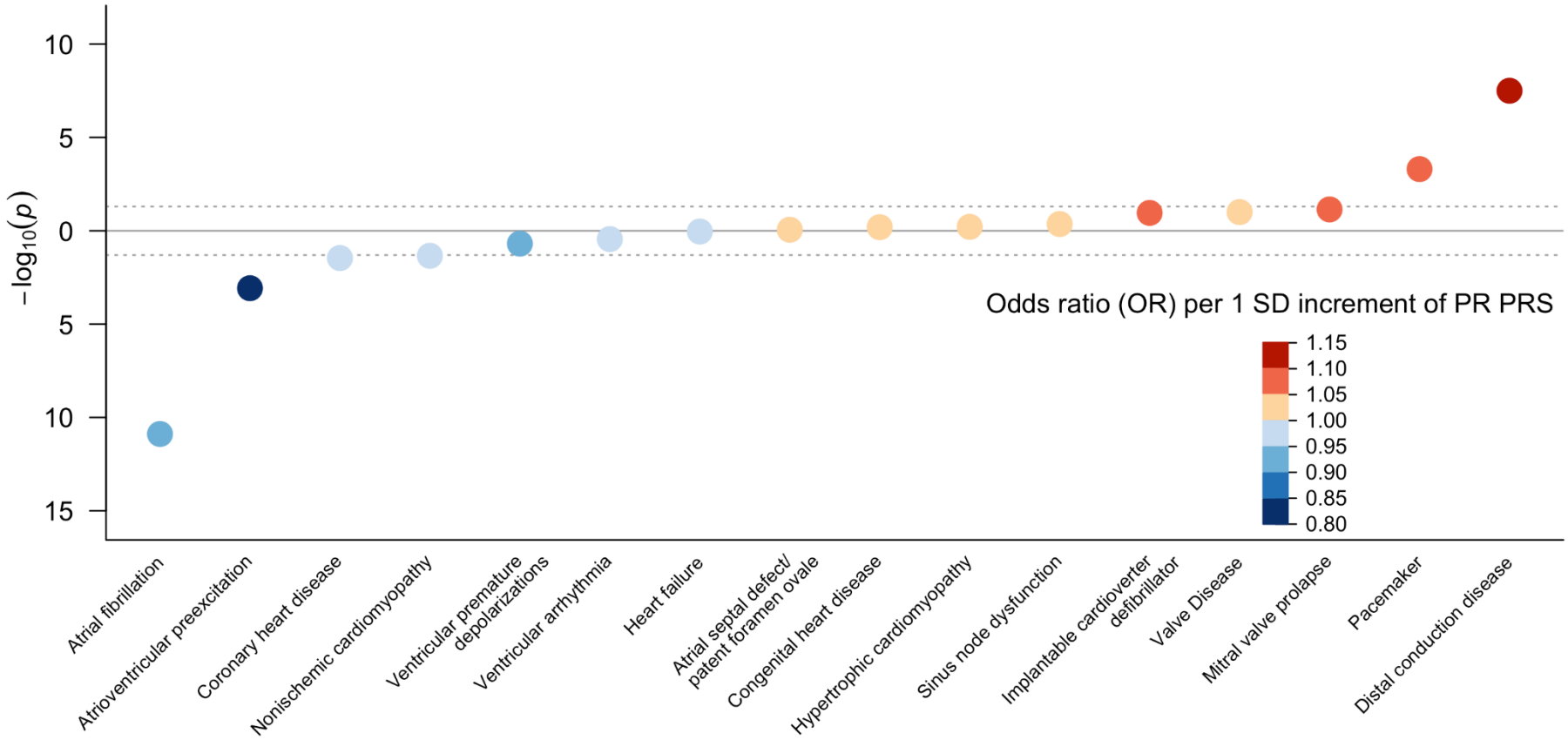
ACP6	*DEK*	*LRCH1*	*SNX1*
AL590822.1	*DNM1P51*	*MSTO2P*	*SYNPO2L*
ALPK3	*EDN2*	*MYO15A*	*TCTN3*
ATP5D	*EEFSEC*	*NPIPA5*	*TMEM182*
BMPR1A	*EF-1*	*NUDT13*	*TMEM72*
CEFIP	*FADS1*	*PDZRN3*	*TPMT*
C11orf1	*FAM211B*	*PHACTR1*	*TRAK1*
CALHM2	*FAT1*	*RP11-29H23.5*	*TRIP4*
CAMK2D	*FKBP7*	*RP11-399K21.11*	*TTC18*
CCDC36	*FUT11*	*RP11-3B7.1*	*VDAC2*
CDH13	*GBAP1*	*RP4-764O22.2*	*VPREB3*
CEP85	*HMGA1P5*	*RPSA*	*XIRP1*
CFDP1	*IFRD2*	*SLC25A26*	*ZCCHC24*
CHRM2	*KCND3*	*SLC6A6*	*ZNF503-AS1*
DAG1	*KDM1B*	*SLK*	

Reduced gene expression

ABHD12	*HCN1*	*NPIPA1*	*SH3PXD2A*
AC011747.4	*IL17D*	*PHLDB2*	*SLC2A11*
AC103965.1	*IL25*	*PLCD1*	*SMARCB1*
AGAP5	*KP-3*	*PPAPDC3*	*SPATA20*
BEND7	*LINC00964*	*QRICH1*	*SPTBN1*
C1orf86	*MALAT1*	*RCAN2*	*SSBP3*
CAB39L	*MEI1*	*RP11-1070N10.3*	*SSXP10*
CBX8	*MLF1*	*RP11-182J1.16*	*STRN*
CMTM5	*MMP11*	*RP11-344N10.5*	*SYNE2*
CSPG4P11	*MRPL37*	*RP11-379F4.7*	*SYPL2*
DDX42	*MTSS1*	*RP11-397E7.4*	*TFEC*
DNAH11	*MYBPHL*	*RP11-724N1.1*	*THRB*
EMB	*MYOZ1*	*SCN5A / SCN10A*	*UBE3B*
GBF1	*NDST2*	*SELM*	*WDR73*
GORASP1	*NEURL*	*SH3D21*	*ZHX1*

Figure 4 Bubble plot of phenome-wide association analysis of multi-ancestry PR interval polygenic risk score.

Polygenic risk score was derived from the multi-ancestry meta-analysis results. Orange circles indicate that higher polygenic risk score of prolonged PR interval is associated with an increased risk of the condition, whereas blue circles indicate that higher score is associated with lower risks. The darkness of the colour reflects the effect size (odds ratio, OR) changes per 1 standard deviation increment of the polygenic risk score. Given correlation between traits, we did not establish a pre-specified significance threshold for the analysis and report nominal associations ($P < 0.05$).

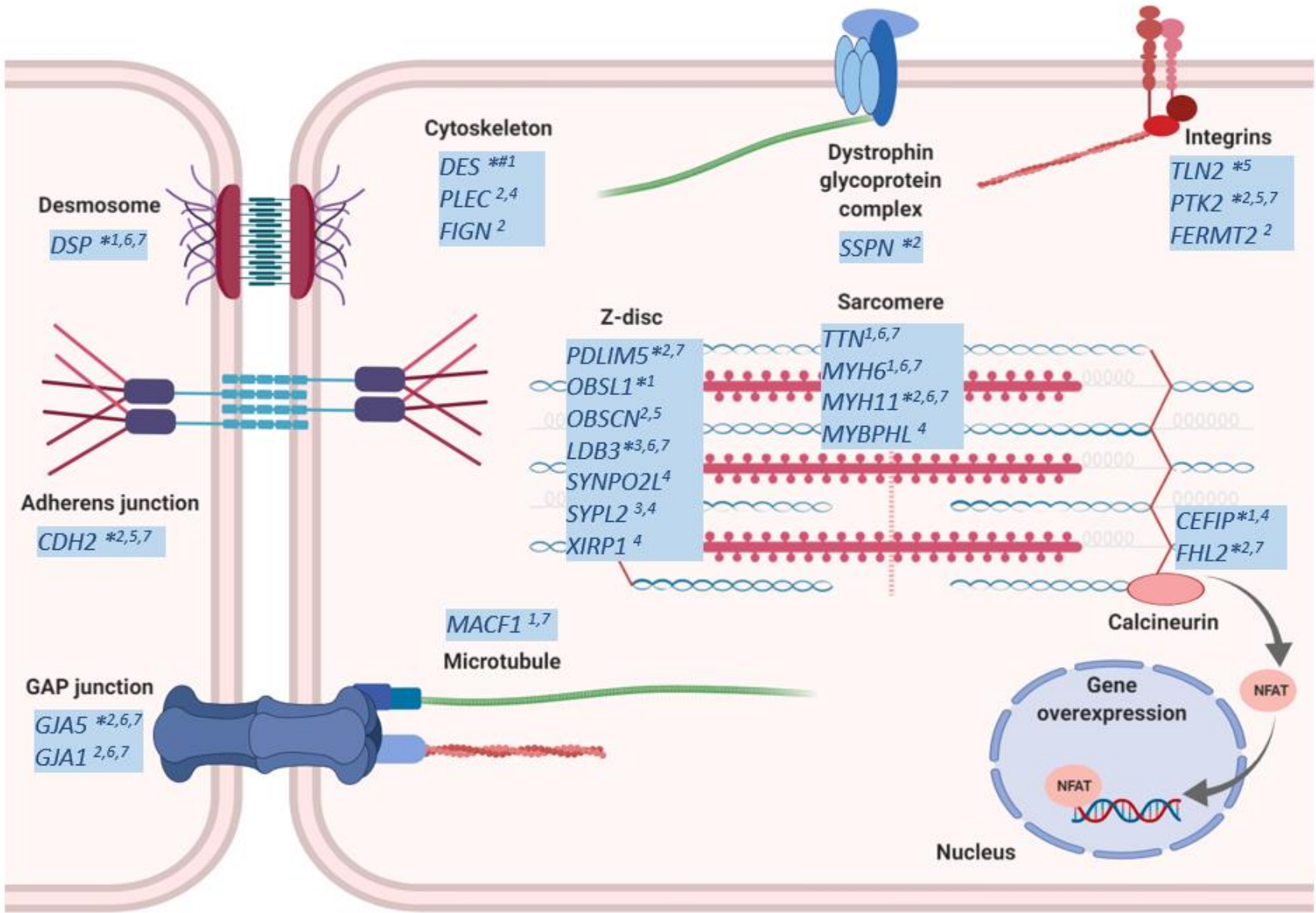

10
5
0
5
10
15
−log10(p)
Odds ratio (OR) per 1 SD increment of PR PRS
1.15
1.10
1.05
1.00
0.95
0.90
0.85
0.80
Atrial fibrillation
Atrioventricular preexcitation
Coronary heart disease
Nonischemic cardiomyopathy
Ventricular premature depolarizations
Ventricular arrhythmia
Heart failure
Atrial septal defect/ patent foramen ovale
Congenital heart disease
Hypertrophic cardiomyopathy
Sinus node dysfunction
Implantable cardioverter defibrillator
Valve Disease
Mitral valve prolapse
Pacemaker
Distal conduction disease

673

674

Figure 5 Candidate genes in PR interval loci encoding proteins involved in cardiac muscle cytoskeleton. Candidate genes or encoded proteins are indicated by a star symbol in the figure and listed in the table. More information about the genes is provided in Supplementary Tables 18-19. *Novel locus, # genome-wide significant locus in transformed trait meta-analysis.

[1] Missense variant; [2] Nearest gene to the lead variant; [3] Gene within the region (r^2>0.5); [4] Variant(s) in the locus are associated with gene expression in left ventricle and/or right atrial appendage; [5] Left ventricle best HiC locus interactor (RegulomeDB score ≤ 2); [6] Animal model; [7] Monogenic cardiovascular disease.


Cytoskeleton
DES *#1
PLEC 2,4
FIGN 2
Dystrophin glycoprotein complex
SSPN *2
Integrins
TLN2 *5
PTK2 *2,5,7
FERMT2 2
Desmosome
DSP *1,6,7
Z-disc
PDLIM5*2,7
OBSL1*1
OBSCN2,5
LDB3*3,6,7
SYNPO2L4
SYPL2 3,4
XIRP1 4
Sarcomere
TTN1,6,7
MYH61,6,7
MYH11*2,6,7
MYBPHL 4
Adherens junction
CDH2 *2,5,7
CEFIP*1,4
FHL2*2,7
Calcineurin
MACF1 1,7
Microtubule
GAP junction
GJA5 *2,6,7
GJA1 2,6,7
Gene overexpression
NFAT
NFAT
Nucleus

681

URLs

1000 Genome Project: http://www.internationalgenome.org

BOLT-LMM: https://data.broadinstitute.org/alkesgroup/BOLT-LMM/

DEPICT: https://data.broadinstitute.org/mpg/depict/

DGIdb: http://www.dgidb.org

EasyQC: https://www.uni-regensburg.de/medizin/epidemiologie-praeventivmedizin/genetische-epidemiologie/software/#

FORGE: https://github.com/iandunham/Forge

GCTA: https://cnsgenomics.com/software/gcta/#Overview

GTEx: https://gtexportal.org/home/

HRC: http://www.haplotype-reference-consortium.org

IMPUTE2: http://mathgen.stats.ox.ac.uk/impute/impute_v2.html

Ingenuity Pathway Analysis software: https://www.qiagenbioinformatics.com/products/ingenuity-pathway-analysis/

International Mouse Phenotyping Consortium: https://www.mousephenotype.org/

IPA: https://www.qiagenbioinformatics.com/products/ingenuity-pathway-analysis

LocusZoom: http://locuszoom.org/

MACH: http://csg.sph.umich.edu/abecasis/mach/tour/imputation.html

METAL: http://csg.sph.umich.edu/abecasis/metal/

OMIM: https://www.omim.org/

RegulomeDB: http://www.regulomedb.org

S-PrediXcan: https://github.com/hakyimlab/MetaXcan

UK Biobank: https://www.ukbiobank.ac.uk

Author contributions

Interpreted results, writing and editing the manuscript: I.N., L.-C.W., S.A.L., and P.B.M. Conceptualisation and supervision of project: S.A.L. and P.B.M. Contributed to GWAS analysis plan: I.N., L.-C.W., H.R.W., Y.J., S.A.L., and P.B.M. Performed meta-analyses: I.N. and L.-C.W. Performed GCTA, heritability, geneset enrichment and pathway analyses, variant annotations: I.N. Performed polygenic risk score and gene expression analyses: S.H.C., M.D.C., and L.-C.W. Performed HiC analyses: I.N., M.R.B., B.M., and P.B.M. Performed gene literature review: I.N., L.-C.W., A.W.Hall, N.R.T., M.D.C., J.H.C., J.J.C., A.T., Y.J., S.A.L., and P.B.M. Contributed to study specific GWAS by providing phenotype, genotype and performing data analyses: J.M., I.R., C.H., P.G., M.Concas, T.B., O.P., I.K., E.T., N.M.A., R.P.S., M.F.L., A.L.P.R., A.M., V.Giedraitis, E.I., A.P.M., F.D.M., L.F., M.G., A.A.Hicks, J.P.C., L.Lind, C.M.L., J.Sundström, N.J.S., C.P.N., M.B.R., S.U., G.S., P.P.M., M.K., N.M., K.N., I.N., M.Caulfield, A.Dominiczak, S.P., M.E.M., J.R.O., A.R.S., K.Ryan, D.C., L.R., S.Aeschbacher, S.Thériault, T.L., O.T.R., N.H., L.Lyytikäinen, J.F.W., P.K.J., C.L.K.B., H.C., C.M.v., J.A.K., A.I., P.L.H., L.-C.W., S.A.L., P.T.E., T.B.H., L.J.L., A.V.S., V.Gudnason, E.P.B., R.J.F.L., G.N.N., M.H.P., A.C., H.M., J.W., M.Müller-Nurasyid, A.P., T.M., M.W., T.D.S., Y.J., M.Mangino, M.R., Y.J.V., P.H., N.V., K.Schramm, S.K., K.Strauch, M.F.S., B.L., C.R., D.F., M.J.C., M.Olesen, D.M.R., M.B.S., J.Smith, J.A.B., M.L.B., J.C.B., B.M.P., N.S., K.Rice, C.P., P.P.P., A.De Grandi, C.F., J.W.J., I.F., P.W.M., S.Trompet, S.W., M.D., S.B.F., U.V., A.S.Havulinna, A.J., K.Sääksjärvi, V.S., S.R.H., J.I.R., X.G., H.J.L., J.Y., K.D.T., R.N., R.d., D.O.M., A.C.M., F.C., J.D., E.G.L., Y.Q., K.V.T., E.J.B., D.L., H.L., C.H.N., K.L.L., A.D.M., D.J.P., B.H.Smith, B.H.Stricker, M.E.v, A.U., J.H., R.D.J., U.P., A.P.R., E.A.W., C.K., E.B., D.E.A., G.B.E., A.A., E.Z.S., C.L.A., S.M.G., K.F.K., C.C.L., A.A.S., A.S., S.Assa, M.A.S., M.Y.v., P.D.L., A.T., M.Orini, J.R., S.V.D.,

P.B.M., K.Stefansson, H.H., P.S., G.S., G.T., R.B.T., U.T., D.O.A., D.F.G. All authors read, revised and approved the manuscript.

Competing Interests

S.A.L. receives sponsored research support from Bristol Myers Squibb / Pfizer, Bayer AG, and Boehringer Ingelheim, and has consulted for Bristol Myers Squibb / Pfizer and Bayer AG. P.T.E. is supported by a grant from Bayer AG to the Broad Institute focused on the genetics and therapeutics of cardiovascular diseases. P.T.E. has also served on advisory boards or consulted for Bayer AG, Quest Diagnostics, and Novartis. M.J.C. is Chief Scientist for Genomics England, a UK Government company. B.M.P. serves on the DSMB of a clinical trial funded by Zoll LifeCor and on the Steering Committee of the Yale Open Data Access Project funded by Johnson & Johnson. V.S. has participated in a conference trip sponsored by Novo Nordisk and received a modest honorarium for participating in an advisory board meeting K.Stefansson, H.H., P.S., G.S., G.T., R.B.T., U.T., D.O.A., D.F.G. are employed by deCODE genetics/Amgen Inc.

References

1. Alonso, A. *et al.* Simple risk model predicts incidence of atrial fibrillation in a racially and geographically diverse population: the CHARGE-AF consortium. *J Am Heart Assoc* **2**, e000102 (2013).
2. Cheng, S. *et al.* Long-term outcomes in individuals with prolonged PR interval or first-degree atrioventricular block. *JAMA* **301**, 2571-7 (2009).
3. Rasmussen, P.V. *et al.* Electrocardiographic PR Interval Duration and Cardiovascular Risk: Results From the Copenhagen ECG Study. *Can J Cardiol* **33**, 674-681 (2017).
4. Butler, A.M. *et al.* Novel loci associated with PR interval in a genome-wide association study of 10 African American cohorts. *Circ Cardiovasc Genet* **5**, 639-46 (2012).
5. Chambers, J.C. *et al.* Genetic variation in SCN10A influences cardiac conduction. *Nat Genet* **42**, 149-52 (2010).
6. Holm, H. *et al.* Several common variants modulate heart rate, PR interval and QRS duration. *Nat Genet* **42**, 117-22 (2010).
7. Hong, K.W. *et al.* Identification of three novel genetic variations associated with electrocardiographic traits (QRS duration and PR interval) in East Asians. *Hum Mol Genet* **23**, 6659-67 (2014).
8. Pfeufer, A. *et al.* Genome-wide association study of PR interval. *Nat Genet* **42**, 153-9 (2010).
9. Sano, M. *et al.* Genome-wide association study of electrocardiographic parameters identifies a new association for PR interval and confirms previously reported associations. *Hum Mol Genet* **23**, 6668-76 (2014).
10. van Setten, J. *et al.* PR interval genome-wide association meta-analysis identifies 50 loci associated with atrial and atrioventricular electrical activity. *Nat Commun* **9**, 2904 (2018).
11. Verweij, N. *et al.* Genetic determinants of P wave duration and PR segment. *Circ Cardiovasc Genet* **7**, 475-81 (2014).
12. van Setten, J. *et al.* Genome-wide association meta-analysis of 30,000 samples identifies seven novel loci for quantitative ECG traits. *Eur J Hum Genet* (2019).
13. Lin, H. *et al.* Common and Rare Coding Genetic Variation Underlying the Electrocardiographic PR Interval. *Circ Genom Precis Med* **11**, e002037 (2018).
14. Genomes Project, C. *et al.* A global reference for human genetic variation. *Nature* **526**, 68-74 (2015).
15. Liu, Y. *et al.* SPSB3 targets SNAIL for degradation in GSK-3beta phosphorylation-dependent manner and regulates metastasis. *Oncogene* **37**, 768-776 (2018).
16. Holm, H. *et al.* A rare variant in MYH6 is associated with high risk of sick sinus syndrome. *Nat Genet* **43**, 316-20 (2011).
17. Thorolfsdottir, R.B. *et al.* A Missense Variant in PLEC Increases Risk of Atrial Fibrillation. *J Am Coll Cardiol* **70**, 2157-2168 (2017).
18. Consortium, G.T. *et al.* Genetic effects on gene expression across human tissues. *Nature* **550**, 204-213 (2017).
19. Schmitt, A.D. *et al.* A Compendium of Chromatin Contact Maps Reveals Spatially Active Regions in the Human Genome. *Cell Rep* **17**, 2042-2059 (2016).
20. Pers, T.H. *et al.* Biological interpretation of genome-wide association studies using predicted gene functions. *Nat Commun* **6**, 5890 (2015).
21. Sudlow, C. *et al.* UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med* **12**, e1001779 (2015).

22. Bermudez-Jimenez, F.J. *et al.* Novel Desmin Mutation p.Glu401Asp Impairs Filament Formation, Disrupts Cell Membrane Integrity, and Causes Severe Arrhythmogenic Left Ventricular Cardiomyopathy/Dysplasia. *Circulation* **137**, 1595-1610 (2018).
23. Norgett, E.E. *et al.* Recessive mutation in desmoplakin disrupts desmoplakin-intermediate filament interactions and causes dilated cardiomyopathy, woolly hair and keratoderma. *Hum Mol Genet* **9**, 2761-6 (2000).
24. Rampazzo, A. *et al.* Mutation in human desmoplakin domain binding to plakoglobin causes a dominant form of arrhythmogenic right ventricular cardiomyopathy. *Am J Hum Genet* **71**, 1200-6 (2002).
25. Taylor, M.R. *et al.* Prevalence of desmin mutations in dilated cardiomyopathy. *Circulation* **115**, 1244-51 (2007).
26. van Tintelen, J.P. *et al.* Severe cardiac phenotype with right ventricular predominance in a large cohort of patients with a single missense mutation in the DES gene. *Heart Rhythm* **6**, 1574-83 (2009).
27. Glukhov, A.V. *et al.* Conduction remodeling in human end-stage nonischemic left ventricular cardiomyopathy. *Circulation* **125**, 1835-47 (2012).
28. Gomes, J. *et al.* Electrophysiological abnormalities precede overt structural changes in arrhythmogenic right ventricular cardiomyopathy due to mutations in desmoplakin-A combined murine and human study. *Eur Heart J* **33**, 1942-53 (2012).
29. Fukuzawa, A. *et al.* Interactions with titin and myomesin target obscurin and obscurin-like 1 to the M-band: implications for hereditary myopathies. *J Cell Sci* **121**, 1841-51 (2008).
30. Cheng, H. *et al.* Loss of enigma homolog protein results in dilated cardiomyopathy. *Circ Res* **107**, 348-56 (2010).
31. Hojayev, B., Rothermel, B.A., Gillette, T.G. & Hill, J.A. FHL2 binds calcineurin and represses pathological cardiac growth. *Mol Cell Biol* **32**, 4025-34 (2012).
32. Friedrich, F.W. *et al.* FHL2 expression and variants in hypertrophic cardiomyopathy. *Basic Res Cardiol* **109**, 451 (2014).
33. Dierck, F. *et al.* The novel cardiac z-disc protein CEFIP regulates cardiomyocyte hypertrophy by modulating calcineurin signaling. *J Biol Chem* **292**, 15180-15191 (2017).
34. Duhme, N. *et al.* Altered HCN4 channel C-linker interaction is associated with familial tachycardia-bradycardia syndrome and atrial fibrillation. *Eur Heart J* **34**, 2768-75 (2013).
35. Milanesi, R., Baruscotti, M., Gnecchi-Ruscone, T. & DiFrancesco, D. Familial sinus bradycardia associated with a mutation in the cardiac pacemaker channel. *N Engl J Med* **354**, 151-7 (2006).
36. Milano, A. *et al.* HCN4 mutations in multiple families with bradycardia and left ventricular noncompaction cardiomyopathy. *J Am Coll Cardiol* **64**, 745-56 (2014).
37. Priori, S.G. *et al.* Mutations in the cardiac ryanodine receptor gene (hRyR2) underlie catecholaminergic polymorphic ventricular tachycardia. *Circulation* **103**, 196-200 (2001).
38. Kubo, T. *et al.* Cloning, sequencing and expression of complementary DNA encoding the muscarinic acetylcholine receptor. *Nature* **323**, 411-6 (1986).
39. Kurachi, Y. G protein regulation of cardiac muscarinic potassium channel. *Am J Physiol* **269**, C821-30 (1995).
40. Aistrup, G.L. *et al.* Targeted G-protein inhibition as a novel approach to decrease vagal atrial fibrillation by selective parasympathetic attenuation. *Cardiovasc Res* **83**, 481-92 (2009).

41. Dobrev, D. *et al.* Molecular basis of downregulation of G-protein-coupled inward rectifying K(+) current (I(K,ACh) in chronic human atrial fibrillation: decrease in GIRK4 mRNA correlates with reduced I(K,ACh) and muscarinic receptor-mediated shortening of action potentials. *Circulation* **104**, 2551-7 (2001).
42. Stavrakis, S. *et al.* Activating autoantibodies to the beta-1 adrenergic and m2 muscarinic receptors facilitate atrial fibrillation in patients with Graves' hyperthyroidism. *J Am Coll Cardiol* **54**, 1309-16 (2009).
43. Winkler, T.W. *et al.* Quality control and conduct of genome-wide association meta-analyses. *Nat Protoc* **9**, 1192-212 (2014).
44. Willer, C.J., Li, Y. & Abecasis, G.R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190-1 (2010).
45. Higgins, J.P., Thompson, S.G., Deeks, J.J. & Altman, D.G. Measuring inconsistency in meta-analyses. *BMJ* **327**, 557-60 (2003).
46. Pruim, R.J. *et al.* LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics* **26**, 2336-7 (2010).
47. Yang, J., Lee, S.H., Goddard, M.E. & Visscher, P.M. GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet* **88**, 76-82 (2011).
48. Loh, P.R. *et al.* Contrasting genetic architectures of schizophrenia and other complex diseases using fast variance-components analysis. *Nat Genet* **47**, 1385-92 (2015).
49. McLaren, W. *et al.* The Ensembl Variant Effect Predictor. *Genome Biol* **17**, 122 (2016).
50. Kumar, P., Henikoff, S. & Ng, P.C. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc* **4**, 1073-81 (2009).
51. Adzhubei, I., Jordan, D.M. & Sunyaev, S.R. Predicting functional effect of human missense mutations using PolyPhen-2. *Curr Protoc Hum Genet* **Chapter 7**, Unit7 20 (2013).
52. Bernstein, B.E. *et al.* The NIH Roadmap Epigenomics Mapping Consortium. *Nat Biotechnol* **28**, 1045-8 (2010).
53. Ward, L.D. & Kellis, M. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Res* **40**, D930-4 (2012).
54. Barbeira, A.N. *et al.* Exploring the phenotypic consequences of tissue specific gene expression variation inferred from GWAS summary statistics. *Nat Commun* **9**, 1825 (2018).
55. Iotchkova, V. *et al.* GARFIELD classifies disease-relevant genomic features through integration of functional annotations with association signals. *Nat Genet* **51**, 343-353 (2019).
56. Consortium, E.P. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57-74 (2012).
57. Buniello, A. *et al.* The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res* **47**, D1005-D1012 (2019).
58. Staley, J.R. *et al.* PhenoScanner: a database of human genotype-phenotype associations. *Bioinformatics* **32**, 3207-3209 (2016).
59. Nielsen, J.B. *et al.* Genome-wide Study of Atrial Fibrillation Identifies Seven Risk Loci and Highlights Biological Pathways and Regulatory Elements Involved in Cardiac Development. *Am J Hum Genet* **102**, 103-115 (2018).
60. Roselli, C. *et al.* Multi-ethnic genome-wide association study for atrial fibrillation. *Nat Genet* **50**, 1225-1233 (2018).
61. Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* **81**, 559-75 (2007).

892