A Large-Scale Multi-ancestry Genome-wide Study Accounting for Smoking Behavior Identifies Multiple Significant Loci for Blood Pressure

Yun J. Sung,[1,216,*] Thomas W. Winkler,[2,216] Lisa de las Fuentes,[3,216] Amy R. Bentley,[4,216] Michael R. Brown,[5,216] Aldi T. Kraja,[6,216] Karen Schwander,[1,216] Ioanna Ntalla,[7,216] Xiuqing Guo,[8] Nora Franceschini,[9] Yingchang Lu,[10] Ching-Yu Cheng,[11,12,13] Xueling Sim,[14] Dina Vojinovic,[15] Jonathan Marten,[16] Solomon K. Musani,[17] Changwei Li,[18] Mary F. Feitosa,[6] Tuomas O. Kilpeläinen,[19,20] Melissa A. Richard,[21] Raymond Noordam,[22] Stella Aslibekyan,[23] Hugues Aschard,[24,25] Traci M. Bartz,[26] Rajkumar Dorajoo,[27] Yongmei Liu,[28] Alisa K. Manning,[29,30] Tuomo Rankinen,[31] Albert Vernon Smith,[32,33] Salman M. Tajuddin,[34] Bamidele O. Tayo,[35] Helen R. Warren,[7,36] Wei Zhao,[37] Yanhua Zhou,[38] Nana Matoba,[39] Tamar Sofer,[40] Maris Alver,[41] Marzyeh Amini,[42]

(Author list continued on next page)

Genome-wide association analysis advanced understanding of blood pressure (BP), a major risk factor for vascular conditions such as coronary heart disease and stroke. Accounting for smoking behavior may help identify BP loci and extend our knowledge of its genetic architecture. We performed genome-wide association meta-analyses of systolic and diastolic BP incorporating gene-smoking interactions in 610,091 individuals. Stage 1 analysis examined ~18.8 million SNPs and small insertion/deletion variants in 129,913 individuals from four ancestries (European, African, Asian, and Hispanic) with follow-up analysis of promising variants in 480,178 additional individuals from five ancestries. We identified 15 loci that were genome-wide significant ($p < 5 \times 10^{-8}$) in stage 1 and formally replicated in stage 2. A combined stage 1 and 2 meta-analysis identified 66 additional genome-wide significant loci (13, 35, and 18 loci in European, African, and trans-ancestry, respectively). A total of 56 known BP loci were also identified by our results ($p < 5 \times 10^{-8}$). Of the newly identified loci, ten showed significant interaction with smoking status, but none of them were replicated in stage 2. Several loci were identified in African ancestry, highlighting the importance of genetic studies in diverse populations. The identified loci show strong evidence for regulatory features and support shared pathophysiology with cardiometabolic and addiction traits. They also highlight a role in BP regulation for biological candidates such as modulators of vascular structure and function (*CDKN1B*, *BCAR1-CFDP1*, *PXDN*, *EEA1*), ciliopathies (*SDCCAG8*, *RPGRIP1L*), telomere maintenance (*TNKS*, *PINX1*, *AKTIP*), and central dopaminergic signaling (*MSRA*, *EBF2*).

Introduction

The management of blood pressure (BP) is a major public health priority with implications for the prevention of coronary heart disease, heart failure, stroke, and other vascular conditions. BP is partly under genetic control with moderately high heritability (30%–60%),[1] although only a small fraction of the heritability has been explained by variants identified through genome-wide association studies (GWASs).[2] Specifically, the common variants

[1]Division of Biostatistics, Washington University School of Medicine, St. Louis, MO 63110, USA; [2]Department of Genetic Epidemiology, University of Regensburg, Regensburg 93051, Germany; [3]Cardiovascular Division, Department of Medicine, Washington University, St. Louis, MO 63110, USA; [4]Center for Research on Genomics and Global Health, National Human Genome Research Institute, NIH, Bethesda, MD 20892, USA; [5]Department of Epidemiology, Human Genetics, and Environmental Sciences, The University of Texas School of Public Health, Houston, TX 77030, USA; [6]Division of Statistical Genomics, Department of Genetics, Washington University School of Medicine, St. Louis, MO 63108, USA; [7]William Harvey Research Institute, Clinical Pharmacology, Queen Mary University of London, London EC1M 6BQ, UK; [8]Genomic Outcomes, Department of Pediatrics, LABioMed at Harbor-UCLA Medical Center, Torrance, CA 90502, USA; [9]Department of Epidemiology, University of North Carolina Gillings School of Global Public Health, Chapel Hill, NC 27514, USA; [10]Icahn School of Medicine at Mount Sinai, The Charles Bronfman Institute for Personalized Medicine, New York, NY 10029, USA; [11]Singapore Eye Research Institute, Singapore National Eye Centre, Singapore, Singapore 169856, Singapore; [12]Ophthalmology & Visual Sciences Academic Clinical Program (Eye ACP), Duke-NUS Medical School, Singapore, Singapore 169857, Singapore; [13]Department of Ophthalmology, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore 117597, Singapore; [14]Saw Swee Hock School of Public Health, National University Health System and National University of Singapore, Singapore, Singapore 117549, Singapore; [15]Department of Epidemiology, Erasmus University Medical Center, Rotterdam, the Netherlands; [16]Medical Research Council Human Genetics Unit, Institute of Genetics & Molecular Medicine, University of Edinburgh, Edinburgh EH4 2XU, UK; [17]Jackson Heart Study, Department of Medicine, University of Mississippi Medical Center, Jackson, MS 39213, USA; [18]Department of Epidemiology and Biostatistics, University of Giorgia at Athens College of Public Health, Athens, GA 30602, USA; [19]Section of Metabolic Genetics, Novo Nordisk Foundation Center for Basic Metabolic Research, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen 2100, Denmark; [20]Department of Environmental Medicine and Public Health, The Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA; [21]Brown Foundation Institute of Molecular Medicine, University of Texas Health Science Center at Houston, Houston, TX 77030, USA; [22]Department of Internal Medicine, Section of Gerontology and Geriatrics, Leiden University Medical Center, Leiden 2300RC, the Netherlands; [23]Department of Epidemiology, University of Alabama at Birmingham, Birmingham, AL 35294, USA; [24]Department of Epidemiology, Harvard School of Public Health, Boston, MA 02115, USA; [25]Centre de Bioinformatique Biostatistique et Biologie Integrative (C3BI), Institut Pasteur, Paris 75015, France; [26]Cardiovascular Health Research Unit, Biostatistics

(Affiliations continued on next page)

Mathilde Boissel,[43] Jin Fang Chai,[44] Xu Chen,[45] Jasmin Divers,[28] Ilaria Gandin,[46] Chuan Gao,[47] Franco Giulianini,[48] Anuj Goel,[49,50] Sarah E. Harris,[51,52] Fernando Pires Hartwig,[53] Andrea R.V.R. Horimoto,[54] Fang-Chi Hsu,[28] Anne U. Jackson,[55] Mika Kähönen,[56,57] Anuradhani Kasturiratne,[58] Brigitte Kühnel,[59,60] Karin Leander,[61] Wen-Jane Lee,[62] Keng-Hung Lin,[63] Jian 'an Luan,[64] Colin A. McKenzie,[65] He Meian,[66] Christopher P. Nelson,[67,68] Rainer Rauramaa,[69] Nicole Schupf,[70] Robert A. Scott,[64] Wayne H.H. Sheu,[71,72,73,74] Alena Stančáková,[75] Fumihiko Takeuchi,[76] Peter J. van der Most,[42] Tibor V. Varga,[77] Heming Wang,[78] Yajuan Wang,[78] Erin B. Ware,[37,79] Stefan Weiss,[80,81] Wanqing Wen,[82] Lisa R. Yanek,[83] Weihua Zhang,[84,85] Jing Hua Zhao,[64] Saima Afaq,[84] Tamuno Alfred,[10] Najaf Amin,[15] Dan Arking,[86] Tin Aung,[11,12,13] R. Graham Barr,[87] Lawrence F. Bielak,[37] Eric Boerwinkle,[88,89] Erwin P. Bottinger,[10] Peter S. Braund,[67,68] Jennifer A. Brody,[90] Ulrich Broeckel,[91] Claudia P. Cabrera,[7,36] Brian Cade,[92] Yu Caizheng,[66] Archie Campbell,[93] Mickaël Canouil,[43] Aravinda Chakravarti,[94] The CHARGE Neurology Working Group, Ganesh Chauhan,[95] Kaare Christensen,[96] Massimiliano Cocca,[46] The COGENT-Kidney Consortium, Francis S. Collins,[97] John M. Connell,[98] Renée de Mutsert,[99] H. Janaka de Silva,[100] Stephanie Debette,[101,102,103] Marcus Dörr,[81,104] Qing Duan,[105] Charles B. Eaton,[106] Georg Ehret,[94,107] Evangelos Evangelou,[84,108] Jessica D. Faul,[109] Virginia A. Fisher,[38] Nita G. Forouhi,[64] Oscar H. Franco,[15] Yechiel Friedlander,[110] He Gao,[84,111] The GIANT Consortium,

(Author list continued on next page)

and Medicine, University of Washington, Seattle, WA 98101, USA; [27]Genome Institute of Singapore, Agency for Science Technology and Research, Singapore 138672, Singapore; [28]Division of Biostatistical Sciences, Department of Public Health Sciences, Wake Forest School of Medicine, Winston-Salem, NC 27157, USA; [29]Center for Human Genetics Research, Massachusetts General Hospital, Boston, MA 02114, USA; [30]Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA; [31]Human Genomics Laboratory, Pennington Biomedical Research Center, Baton Rouge, LA 70808, USA; [32]Icelandic Heart Association, Kopavogur 201, Iceland; [33]Faculty of Medicine, University of Iceland, Reykjavik 101, Iceland; [34]Health Disparities Research Section, Laboratory of Epidemiology and Population Sciences, National Institute on Aging, NIH, Baltimore, MD 21224, USA; [35]Department of Public Health Sciences, Loyola University Chicago, Maywood, IL 60153, USA; [36]NIHR Cardiovascular Biomedical Research Unit, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London EC1M 6BQ, UK; [37]School of Public Health, Department of Epidemiology, University of Michigan, Ann Arbor, MI 48109, USA; [38]Department of Biostatistics, Boston University School of Public Health, Boston, MA 02118, USA; [39]Laboratory for Statistical Analysis, Center for Integrative Medical Sciences, RIKEN, Yokohama 230-0045, Japan; [40]Department of Biostatistics, University of Washington, Seattle, WA 98105, USA; [41]Estonian Genome Center, University of Tartu, Tartu 51010, Estonia; [42]Department of Epidemiology, University of Groningen, University Medical Center Groningen, Groningen 9700 RB, the Netherlands; [43]CNRS UMR 8199, European Genomic Institute for Diabetes (EGID), Institut Pasteur de Lille, University of Lille, Lille 59000, France; [44]Saw Swee Hock School of Public Health, National University of Singapore, Singapore 117549, Singapore; [45]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm 17177, Sweden; [46]Department of Medical Sciences, University of Trieste, Trieste 34137, Italy; [47]Department of Molecular Genetics and Genomics Program, Wake Forest School of Medicine, Winston-Salem, NC 27157, USA; [48]Division of Preventive Medicine, Department of Medicine, Brigham and Women's Hospital, Boston, MA 02215, USA; [49]Division of Cardiovascular Medicine, Radcliffe Department of Medicine, University of Oxford, Oxford, Oxfordshire OX3 9DU, UK; [50]Wellcome Centre for Human Genetics, University of Oxford, Oxford OX3 7BN, UK; [51]Centre for Cognitive Ageing and Cognitive Epidemiology, The University of Edinburgh, Edinburgh EH8 9JZ, UK; [52]Medical Genetics Section, University of Edinburgh Centre for Genomic and Experimental Medicine and MRC Institute of Genetics and Molecular Medicine, The University of Edinburgh, Edinburgh EH4 2XU, UK; [53]Postgraduate Program in Epidemiology, Federal University of Pelotas, Pelotas, RS 96020220, Brazil; [54]Lab Genetics and Molecular Cardiology, Department of Cardiology, Heart Institute, University of Sao Paulo, Sao Paulo, Brazil; [55]Department of Biostatistics and Center for Statistical Genetics, University of Michigan, Ann Arbor, MI 48109, USA; [56]Department of Clinical Physiology, Faculty of Medicine and Life Sciences, University of Tampere, Tampere 33014, Finland; [57]Department of Clinical Physiology, Tampere University Hospital, Tampere, Finland; [58]Department of Public Health, University of Kelaniya, Ragama, Sri Lanka; [59]Research Unit of Molecular Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg 85764, Germany; [60]Institute of Epidemiology II, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg 85764, Germany; [61]Institute of Environmental Medicine, Karolinska Institutet, Stockholm 17177, Sweden; [62]Department of Medical Research, Taichung Veterans General Hospital, Department of Social Work, Tunghai University, Taichung 40705, Taiwan; [63]Department of Opthalmology, Taichung Veterans General Hospital, Taichung 40705, Taiwan; [64]MRC Epidemiology Unit, University of Cambridge, Cambridge CB2 0QQ, UK; [65]Tropical Metabolism Research Unit, Tropical Medicine Research Institute, University of the West Indies, Mona JMAAW15, Jamaica; [66]Department of Occupational and Environmental Health and State Key Laboratory of Environmental Health for Incubating, School of Public Health, Tongji Medical College Huazhong University of Science and Technology, Wuhan, China; [67]Department of Cardiovascular Sciences, University of Leicester, Leicester LE3 9QP, UK; [68]NIHR Leicester Cardiovascular Biomedical Research Unit, Glenfield Hospital, Leicester LE3 9QP, UK; [69]Foundation for Research in Health Exercise and Nutrition, Kuopio Research Institute of Exercise Medicine, Kuopio 70100, Finland; [70]Taub Institute for Research on Alzheimer disease and the Aging Brain, Department of Epidemiology, Columbia University Mailman School of Public Health, New York, NY 10032, USA; [71]Endocrinology and Metabolism, Department of Internal Medicine, Taichung Veterans General Hospital, Taichung 40705, Taiwan; [72]School of Medicine, National Yang-ming University, Taipei, Taiwan; [73]School of Medicine, National Defense Medical Center, Taipei, Taiwan; [74]Institute of Medical Technology, National Chung-Hsing University, Taichung 40705, Taiwan; [75]Institute of Clinical Medicine, Internal Medicine, University of Eastern Finland, Kuopio 70210, Finland; [76]Department of Gene Diagnostics and Therapeutics, Research Institute, National Center for Global Health and Medicine, Tokyo 1628655, Japan; [77]Genetic and Molecular Epidemiology Unit, Department of Clinical Sciences, Lund University, Malmö, Skåne 205 02, Sweden; [78]Department of Epidemiology and Biostatistics, Case Western Reserve University, Cleveland, OH 44106, USA; [79]Institute for Social Research, Research Center for Group Dynamics, University of Michigan, Ann Arbor, MI 48104, USA; [80]Interfaculty Institute for Genetics and Functional Genomics, University Medicine and Ernst-Moritz Arndt University Greifswald, Greifswald 17487, Germany; [81]DZHK (German Center for Cardiovascular Research), partner site Greifswald, Greifswald 17475, Germany; [82]Division of Epidemiology, Department of Medicine, Vanderbilt University School of Medicine, Nashville, TN 37203, USA; [83]General Internal Medicine, GeneSTAR Research Program, Department of Medicine, Johns Hopkins University School of Medicine, Baltimore, MD 21287, USA; [84]Department of Epidemiology and Biostatistics, Imperial College London, London W2 1PG, UK; [85]Department of Cardiology, Ealing Hospital, Middlesex UB1 3HW, UK; [86]McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA; [87]Departments of Medicine and Epidemiology, Columbia University Medical Center, New York, NY

(Affiliations continued on next page)

Bruna Gigante,[61] Misa Graff,[9] C. Charles Gu,[1] Dongfeng Gu,[112] Preeti Gupta,[11] Saskia P. Hagenaars,[51,113] Tamara B. Harris,[114] Jiang He,[115,116] Sami Heikkinen,[117] Chew-Kiat Heng,[118,119] Makoto Hirata,[120] Albert Hofman,[15] Barbara V. Howard,[121,122] Steven Hunt,[123,124] Marguerite R. Irvin,[23] Yucheng Jia,[8] Roby Joehanes,[125,126] Anne E. Justice,[9] Tomohiro Katsuya,[127,128] Joel Kaufman,[129] Nicola D. Kerrison,[64] Chiea Chuen Khor,[27,130] Woon-Puay Koh,[44,131] Heikki A. Koistinen,[132,133] Pirjo Komulainen,[69] Charles Kooperberg,[134] Jose E. Krieger,[54] Michiaki Kubo,[135] Johanna Kuusisto,[136] Carl D. Langefeld,[28] Claudia Langenberg,[64] Lenore J. Launer,[114] Benjamin Lehne,[84] Cora E. Lewis,[137] Yize Li,[1] Lifelines Cohort Study,[138] Sing Hui Lim,[11] Shiow Lin,[6] Ching-Ti Liu,[38] Jianjun Liu,[14,27] Jingmin Liu,[139] Kiang Liu,[140] Yeheng Liu,[8] Marie Loh,[84,141] Kurt K. Lohman,[28] Jirong Long,[82] Tin Louie,[40] Reedik Mägi,[41] Anubha Mahajan,[50] Thomas Meitinger,[142,143] Andres Metspalu,[41] Lili Milani,[41] Yukihide Momozawa,[144] Andrew P. Morris,[50,145] Thomas H. Mosley, Jr.,[146] Peter Munson,[147] Alison D. Murray,[148] Mike A. Nalls,[149,150] Ubaydah Nasri,[8] Jill M. Norris,[151] Kari North,[9] Adesola Ogunniyi,[152] Sandosh Padmanabhan,[153] Walter R. Palmas,[154] Nicholette D. Palmer,[155] James S. Pankow,[156] Nancy L. Pedersen,[45] Annette Peters,[60,157] Patricia A. Peyser,[37] Ozren Polasek,[158] Olli T. Raitakari,[159,160] Frida Renström,[77,161] Treva K. Rice,[1] Paul M. Ridker,[48] Antonietta Robino,[162] Jennifer G. Robinson,[163] Lynda M. Rose,[48] Igor Rudan,[164] Charumathi Sabanayagam,[11,12] Babatunde L. Salako,[152] Kevin Sandow,[8] Carsten O. Schmidt,[81,165] Pamela J. Schreiner,[156] William R. Scott,[84,166] Sudha Seshadri,[126,167] Peter Sever,[168] Colleen M. Sitlani,[90]

(Author list continued on next page)

10032, USA; [88]Human Genetics Center, The University of Texas School of Public Health, Houston, TX 77030, USA; [89]Institute of Molecular Medicine, The University of Texas Health Science Center, Houston, TX 77030, USA; [90]Cardiovascular Health Research Unit, Medicine, University of Washington, Seattle, WA 98101, USA; [91]Section of Genomic Pediatrics, Department of Pediatrics, Medicine and Physiology, Medical College of Wisconsin, Milwaukee, WI 53226, USA; [92]Sleep Medicine and Circadian Disorders, Brigham and Women's Hospital, Boston, MA 02115, USA; [93]Centre for Genomic & Experimental Medicine, Institute of Genetics & Molecular Medicine, University of Edinburgh, Edinburgh EH4 2XU, UK; [94]Center for Complex Disease Genomics, McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA; [95]Centre for Brain Research, Indian Institute of Schience, Bangalore 560012, India; [96]The Danish Aging Research Center, Institute of Public Health, University of Southern Denmark, Odense, Denmark; [97]Medical Genomics and Metabolic Genetics Branch, National Human Genome Research Institute, NIH, Bethesda, MD 20892, USA; [98]Ninewells Hospital & Medical School, University of Dundee, Dundee DD1 9SY, UK; [99]Department of Clinical Epidemiology, Leiden University Medical Center, Leiden 2300RC, the Netherlands; [100]Department of Medicine, University of Kelaniya, Ragama, Sri Lanka; [101]Inserm U1219 Neuroepidemiology, University of Bordeaux, Bordeaux, France; [102]Department of Neurology, University Hospital, Bordeaux, France; [103]Boston University School of Medicine, Boston, MA 02118, USA; [104]Department of Internal Medicine B, University Medicine Greifswald, Greifswald 17475, Germany; [105]Department of Genetics, University of North Carolina, Chapel Hill, NC 27514, USA; [106]Department of Family Medicine and Epidemiology, Alpert Medical School of Brown University, Providence, RI 02860, USA; [107]Division of Cardiology, Department of Specialties of Medicine, Geneva University Hospital, Geneva 1211, Switzerland; [108]Department of Hygiene and Epidemiology, University of Ioannina Medical School, Ioannina 45110, Greece; [109]Institute for Social Research, Survey Research Center, University of Michigan, Ann Arbor, MI 48104, USA; [110]Braun School of Public Health, Hebrew University-Hadassah Medical Center, Jerusalem 91120, Israel; [111]MRC-PHE Centre for Environment and Health, Department of Epidemiology & Biostatistics, School of Public Health, Imperial College London, London, UK; [112]Department of Epidemiology, State Key Laboratory of Cardiovascular Disease, Fuwai Hospital, National Center for Cardiovascular Diseases, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China; [113]Department of Psychology, The University of Edinburgh, Edinburgh EH8 9JZ, UK; [114]Laboratory of Epidemiology and Population Sciences, National Institute on Aging, NIH, Bethesda, MD 20892, USA; [115]Department of Epidemiology, Tulane University School of Public Health and Tropical Medicine, New Orleans, LA 70112, USA; [116]Department of Medicine, Tulane University School of Medicine, New Orleans, LA 70112, USA; [117]University of Eastern Finland, Institute of Biomedicine, Kuopio 70211, Finland; [118]Department of Paediatrics, Yong Loo Lin School of Medicine, National University of Singapore, Singapore 119228, Singapore; [119]Khoo Teck Puat – National University Children's Medical Institute, National University Health System, Singapore 119228, Singapore; [120]Laboratory of Genome Technology, Human Genome Center, Institute of Medical Science, The University of Tokyo, Minato-ku 108-8639, Japan; [121]MedStar Health Research Institute, Hyattsville, MD 20782, USA; [122]Center for Clinical and Translational Sciences and Department of Medicine, Georgetown-Howard Universities, Washington, DC 20057, USA; [123]Cardiovascular Genetics, Department of Internal Medicine, University of Utah, Salt Lake City, UT 84108, USA; [124]Weill Cornell Medicine in Qatar, Doha, Qatar; [125]Hebrew SeniorLife, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA 02131, USA; [126]Framingham Heart Study, National Heart, Lung, and Blood Institute, NIH, Bethesda, MD 20982, USA; [127]Department of Clinical Gene Therapy, Osaka University Graduate School of Medicine, Suita 5650871, Japan; [128]Department of Geriatric and General Medicine, Osaka University Graduate School of Medicine, Suita 5650871, Japan; [129]Epidemiology, Department of Occupational and Environmental Medicine Program, University of Washington, Seattle, WA 98105, USA; [130]Department of Biochemistry, National University of Singapore, Singapore 117596, Singapore; [131]Duke-NUS Medical School, Singapore 169857, Singapore; [132]Department of Health, National Institute for Health and Welfare, Helsinki 00271, Finland; [133]Department of Medicine and Abdominal Center: Endocrinology, University of Helsinki and Helsinki University Central Hospital, Helsinki 00029, Finland; [134]Fred Hutchinson Cancer Research Center, University of Washington School of Public Health, Seattle, WA 98109, USA; [135]Center for Integrative Medical Sciences, RIKEN, Yokohama 230-0045, Japan; [136]Institute of Clinical Medicine, Internal Medicine, University of Eastern Finland and Kuopio University Hospital, Kuopio 70210, Finland; [137]Department of Medicine, University of Alabama at Birmingham, Birmingham, AL 35205, USA; [138]Lifelines cohort study, University of Groningen, University Medical Center Groningen, Groningen 9700 RB, the Netherlands; [139]WHI CCC, Fred Hutchinson Cancer Research Center, Seattle, WA 98115, USA; [140]Preventive Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL 60611, USA; [141]Translational Laboratory in Genetic Medicine, Agency for Science, Technology and Research 138648, Singapore; [142]Institute of Human Genetics, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg 85764, Germany; [143]Institute of Human Genetics, Technische Universität München, Munich 80333, Germany; [144]Laboratory for Genotyping Development, Center for Integrative Medical Sciences, RIKEN, Yokohama 230-0045, Japan; [145]Department of Biostatistics, University of Liverpool, Liverpool L69 3GL, UK; [146]Geriatrics, Department of Medicine, University of Mississippi Medical Center, Jackson, MS 39216, USA; [147]Mathematical and Statistical Computing Laboratory, Center for Information Technology, NIH, Bethesda, MD 20892, USA; [148]The Institute of Medical Sciences, Aberdeen Biomedical Imaging Centre, University of Aberdeen, Aberdeen AB25 2ZD, UK; [149]Data Tecnica International, Glen Echo, MD 20812, USA; [150]Laboratory of Neurogenetics, National Institute on Aging, Bethesda, MD 20892, USA; [151]Department of Epidemiology, Colorado School of Public

(Affiliations continued on next page)

Jennifer A. Smith,[37] Harold Snieder,[42] John M. Starr,[51,169] Konstantin Strauch,[170,171] Hua Tang,[172] Kent D. Taylor,[8] Yik Ying Teo,[14,27,173,174,175] Yih Chung Tham,[11] André G. Uitterlinden,[176] Melanie Waldenberger,[59,60] Lihua Wang,[6] Ya X. Wang,[177] Wen Bin Wei,[178] Christine Williams,[6] Gregory Wilson,[179] Mary K. Wojczynski,[6] Jie Yao,[8] Jian-Min Yuan,[180,181] Alan B. Zonderman,[182] Diane M. Becker,[83] Michael Boehnke,[55] Donald W. Bowden,[155] John C. Chambers,[84,85] Yii-Der Ida Chen,[8] Ulf de Faire,[61] Ian J. Deary,[51,113] Tõnu Esko,[41,183] Martin Farrall,[49,50] Terrence Forrester,[65] Paul W. Franks,[77,184,185] Barry I. Freedman,[186] Philippe Froguel,[43,187] Paolo Gasparini,[46,188] Christian Gieger,[59,189] Bernardo Lessa Horta,[53] Yi-Jen Hung,[190] Jost B. Jonas,[178,191] Norihiro Kato,[76] Jaspal S. Kooner,[85,166] Markku Laakso,[136] Terho Lehtimäki,[192,193] Kae-Woei Liang,[72,194,195] Patrik K.E. Magnusson,[45] Anne B. Newman,[180] Albertine J. Oldehinkel,[196] Alexandre C. Pereira,[54,197] Susan Redline,[92] Rainer Rettig,[81,198] Nilesh J. Samani,[67,68] James Scott,[166] Xiao-Ou Shu,[82] Pim van der Harst,[199] Lynne E. Wagenknecht,[200] Nicholas J. Wareham,[64] Hugh Watkins,[49,50] David R. Weir,[109] Ananda R. Wickremasinghe,[58] Tangchun Wu,[66] Wei Zheng,[82] Yoichiro Kamatani,[39] Cathy C. Laurie,[40] Claude Bouchard,[31] Richard S. Cooper,[35] Michele K. Evans,[34] Vilmundur Gudnason,[32,33] Sharon L.R. Kardia,[37] Stephen B. Kritchevsky,[201] Daniel Levy,[126,202] Jeff R. O'Connell,[203,204] Bruce M. Psaty,[205,206] Rob M. van Dam,[44,207] Mario Sims,[17] Donna K. Arnett,[208] Dennis O. Mook-Kanamori,[99,209] Tanika N. Kelly,[115] Ervin R. Fox,[210] Caroline Hayward,[16] Myriam Fornage,[21] Charles N. Rotimi,[4] Michael A. Province,[6] Cornelia M. van Duijn,[15] E. Shyong Tai,[14,131,207]

(Author list continued on next page)

Health, Aurora, CO 80045, USA; [152]Department of Medicine, University of Ibadan, Ibadan, Nigeria; [153]Institute of Cardiovascular and Medical Sciences, University of Glasgow, Glasgow G12 8TA, UK; [154]Internal Medicine, Department of Medicine, Columbia University, New York, NY 10032, USA; [155]Department of Biochemistry, Wake Forest School of Medicine, Winston-Salem, NC 27157, USA; [156]Division of Epidemiology and Community Health, University of Minnesota School of Public Health, Minneapolis, MN 55454, USA; [157]DZHK (German Centre for Cardiovascular Research), partner site Munich Heart Alliance, Neuherberg 85764, Germany; [158]Faculty of Medicine, University of Split, Split, Croatia; [159]Department of Clinical Physiology and Nuclear Medicine, Turku University Hospital, Turku 20521, Finland; [160]Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, Turku 20520, Finland; [161]Department of Biobank Research, Umeå University, Umeå, Västerbotten 901 87, Sweden; [162]Medical Genetics, IRCCS Burlo, Garofolo 34137, Italy; [163]Department of Epidemiology and Medicine, University of Iowa, Iowa City, IA 52242, USA; [164]Usher Institute of Population Health Sciences and Informatics, University of Edinburgh, Edinburgh EH8 9AG, UK; [165]Institute for Community Medicine, University Medicine Greifswald, Greifswald 17475, Germany; [166]National Heart and Lung Institute, Imperial College London, London W12 0NN, UK; [167]Department of Neurology, Boston University School of Medicine, Boston, MA 02118, USA; [168]International Centre for Circulatory Health, Imperial College London, London W2 1PG, UK; [169]Alzheimer Scotland Dementia Research Centre, The University of Edinburgh, Edinburgh EH8 9JZ, UK; [170]Institute of Genetic Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg 85764, Germany; [171]Chair of Genetic Epidemiology, IBE, Faculty of Medicine, LMU Munich, Munich 81377, Germany; [172]Department of Genetics, Stanford University, Stanford, CA 94305, USA; [173]Life Sciences Institute, National University of Singapore, Singapore, Singapore 117456, Singapore; [174]NUS Graduate School for Integrative Science and Engineering, National University of Singapore, Singapore 117456, Singapore; [175]Department of Statistics and Applied Probability, National University of Singapore, Singapore 117546, Singapore; [176]Department of Internal Medicine, Erasmus University Medical Center, Rotterdam, the Netherlands; [177]Beijing Institute of Ophthalmology, Beijing Ophthalmology and Visual Science Key Lab, Beijing Tongren Eye Center, Capital Medical University, Beijing, China 100730, China; [178]Beijing Tongren Eye Center, Beijing Tongren Hospital, Capital Medical University, Beijing 100730, China; [179]Jackson Heart Study, Department of Public Health, Jackson State University, Jackson, MS 39213, USA; [180]Department of Epidemiology, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA 15261, USA; [181]Division of Cancer Control and Population Sciences, University of Pittsburgh Cancer Institute, Pittsburgh, PA 15232, USA; [182]Behavioral Epidemiology Section, Laboratory of Epidemiology and Population Sciences, National Institute on Aging, NIH, Baltimore, MD 21224, USA; [183]Broad Institute of the Massachusetts Institute of Technology and Harvard University, Boston, MA 02142, USA; [184]Harvard T.H. Chan School of Public Health, Department of Nutrition, Harvard University, Boston, MA 02115, USA; [185]Department of Public Health & Clinical Medicine, Umeå University, Umeå, Västerbotten 901 85, Sweden; [186]Division of Nephrology, Department of Internal Medicine, Wake Forest School of Medicine, Winston-Salem, NC 27157, USA; [187]Department of Genomics of Common Disease, Imperial College London, London W12 0NN, UK; [188]Division Experimental Genetics, Sidra, Doha 26999, Qatar; [189]German Center for Diabetes Research (DZD e.V.), Neuherberg 85764, Germany; [190]Endocrinology and Metabolism, Tri-Service General Hospital, National Defense Medical Center, Taipei City, Taipei 11490, Taiwan; [191]Department of Ophthalmology, Medical Faculty Mannheim, University Heidelberg, Mannheim 68167, Germany; [192]Department of Clinical Chemistry, Fimlab Laboratories, Tampere 33520, Finland; [193]Department of Clinical Chemistry, Finnish Cardiovascular Research Center - Tampere, Faculty of Medicine and Lifes Sciences, University of Tampere, Tampere 33014, Finland; [194]Cardiovascular Center, Taichung Veterans General Hospital, Taichung 40705, Taiwan; [195]Department of Medicine, China Medical University, Taichung 40705, Taiwan; [196]Department of Psychiatry, University of Groningen, University Medical Center Groningen, Groningen 9700 RB, the Netherlands; [197]Genetics, Harvard Medical School, Boston, MA 02115, USA; [198]Institute of Physiology, University Medicine Greifswald, Greifswald 17495, Germany; [199]Department of Cardiology, University of Groningen, University Medical Center Groningen, Groningen 9700 RB, the Netherlands; [200]Department of Public Health Sciences, Wake Forest School of Medicine, Winston-Salem, NC 27157, USA; [201]Sticht Center for Health Aging and Alzheimer's Prevention, Department of Internal Medicine, Wake Forest School of Medicine, Winston-Salem, NC 27157, USA; [202]Population Sciences Branch, Division of Intramural Research, National Heart, Lung, and Blood Institute, NIH, Bethesda, MD 20892, USA; [203]Division of Endocrinology, Diabetes, and Nutrition, University of Maryland School of Medicine, Baltimore, MD 21201, USA; [204]Program for Personalized and Genomic Medicine, University of Maryland School of Medicine, Baltimore, MD 21201, USA; [205]Cardiovascular Health Research Unit, Epidemiology, Medicine and Health Services, University of Washington, Seattle, WA 98101, USA; [206]Kaiser Permanente Washington, Health Research Institute, Seattle, WA 98101, USA; [207]Department of Medicine, Yong Loo Lin School of Medicine, National University of Singapore, Singapore 119228, Singapore; [208]Dean's Office, University of Kentucky College of Public Health, Lexington, KY 40536, USA; [209]Department of Public Health and Primary Care, Leiden University Medical Center, Leiden 2300RC, the Netherlands; [210]Cardiology, Department of Medicine, University of Mississippi Medical Center, Jackson, MS 39216, USA; [211]Icahn School of Medicine at Mount Sinai, The Mindich Child Health and Development Institute, New York, NY 10029, USA; [212]Genomic Outcomes, Department of Medicine, LABioMed at Harbor-UCLA Medical Center, Torrance, CA 90502, USA; [213]Department of Psychiatry, Washington University School of Medicine, St. Louis,

(Affiliations continued on next page)

Tien Yin Wong,[11,12,13] Ruth J.F. Loos,[10,211] Alex P. Reiner,[134] Jerome I. Rotter,[8,212] Xiaofeng Zhu,[78] Laura J. Bierut,[213] W. James Gauderman,[214] Mark J. Caulfield,[7,36,217] Paul Elliott,[84,111,217] Kenneth Rice,[40,217] Patricia B. Munroe,[7,215,217] Alanna C. Morrison,[5,217] L. Adrienne Cupples,[38,126,217] Dabeeru C. Rao,[1,217] and Daniel I. Chasman[48,217,*]

initially identified through three collaborative consortia for genome-wide BP genetics in people of European ancestry[1,3,4] explain less than 2.5% of the variance in systolic BP (SBP) or diastolic BP (DBP).[4] Recent reports based on larger sample sizes have increased the number of BP-associated variants which together explain about 3.5% of BP variance.[5–7] In contrast, only six BP loci have been identified by GWASs in African ancestry which explain less than 0.54% of BP variance.[8,9] A focus on main effects to the exclusion of interactions in these studies may have limited the discovery of a full complement of genetic influences on BP. In particular, incorporating interactions between genetic variants and environmental exposures (GxE) represents an additional route for discovery of genetic effects on complex traits,[10] including BP, and may more generally extend our knowledge of the genetic architecture of complex traits.[11]

Many lifestyle factors including physical activity, tobacco use, alcohol consumption, stress, and dietary factors influence BP.[12] These lifestyle exposures may also modify the effect of genetic variants on BP. Cigarette smoking is known to influence BP in both acute[13] and chronic[14,15] fashion, motivating genetic association studies accounting for potential gene-by-smoking interactions. This may help identify BP loci, and such BP loci driven by GxE interactions may reveal new biological insights and mechanisms that can be explored for treatment or prevention of hypertension.

The recently established Gene-Lifestyle Interactions Working Group within the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium has designed a series of multi-ancestry genome-wide interaction projects focused on assessing the impact of interactions with multiple lifestyle factors on the genetics of cardiovascular traits.[16] The primary goal of these investigations is to use interactions to identify trait loci that act synergistically with lifestyle factors. Large-scale interaction studies like this one represent "an important milestone on the path toward a far more complete understanding of the origins of cardiovascular disease and a better understanding of how to manage it."[17] Within this setting, we performed a genome-wide association meta-analysis incorporating gene-smoking interactions (overview shown in Figure 1) to identify SBP- and DBP-associated loci and understand the modulating role of cigarette smoking in the genetic architecture of BP. Here we report our findings based on a total of 610,091 individuals from five ancestry groups which provide adequate power for discovery.[16]

Material and Methods

Overview of Participating Studies

Men and women between the ages of 18 and 80 years from five self-reported ancestry groups are represented in this study: European (EUR), African (AFR), Asian (ASN), Hispanic (HIS), and Brazilian admixed (BRA). These participating studies are described in the Supplemental Note. Each study obtained informed consent from participants and approval from the appropriate institutional review boards. Although the participating studies are based on different study designs and populations, all of them have data on BP, smoking, and genotypes across the genome (data imputed using the 1000 Genomes reference panel in most cohorts). In total, this study involves two stages comprising 610,091 individuals.

A total of 48 cohorts participated in stage 1 and performed genome-wide interaction analyses (Table S1). This stage included 80,552 EUR, 27,118 AFR, 13,438 ASN, and 8,805 HIS for an overall total of 129,913 individuals. A total of 76 cohorts participated in stage 2 and performed analyses of 4,459 variants that were identified in stage 1 as either genome-wide significant ($p < 5 \times 10^{-8}$) or suggestive ($p < 10^{-6}$) for any of the BP-smoking combinations for either 1 df or 2 df tests (Table S2). This stage included 305,513 EUR, 7,786 AFR, 148,932 ASN, 13,533 HIS, and 4,414 Brazilian admixed (BRA) individuals to a total of 480,178 individuals in stage 2. Since discoveries to date are largely from EUR populations, we optimized the chances of discovery in non-EUR populations (especially in AFR) by recruiting most of the available non-EUR cohorts into stage 1.

Phenotypes and Lifestyle Variables

The two BP traits, resting SBP (mmHg) and DBP (mmHg), were analyzed separately. For individuals taking any anti-hypertensive (BP-lowering) medications, their SBP and DBP values were first adjusted for medication effects by adding 15 mmHg to SBP and adding 10 mmHg to DBP.[3] Summary statistics are shown in Table 1 (more details in Tables S3 and S4). These

MO 63110, USA; [214]Division of Biostatistics, Department of Preventive Medicine, University of Southern California, Los Angeles, CA 90032, USA; [215]NIHR Barts Cardiovascular Biomedical Research Unit, Queen Mary University of London, London EC1M 6BQ, UK
[216]These authors contributed equally to this work
[217]These authors contributed equally to this work
*Correspondence: yunju@wustl.edu (Y.J.S.), dchasman@research.bwh.harvard.edu (D.I.C.)
https://doi.org/10.1016/j.ajhg.2018.01.015.

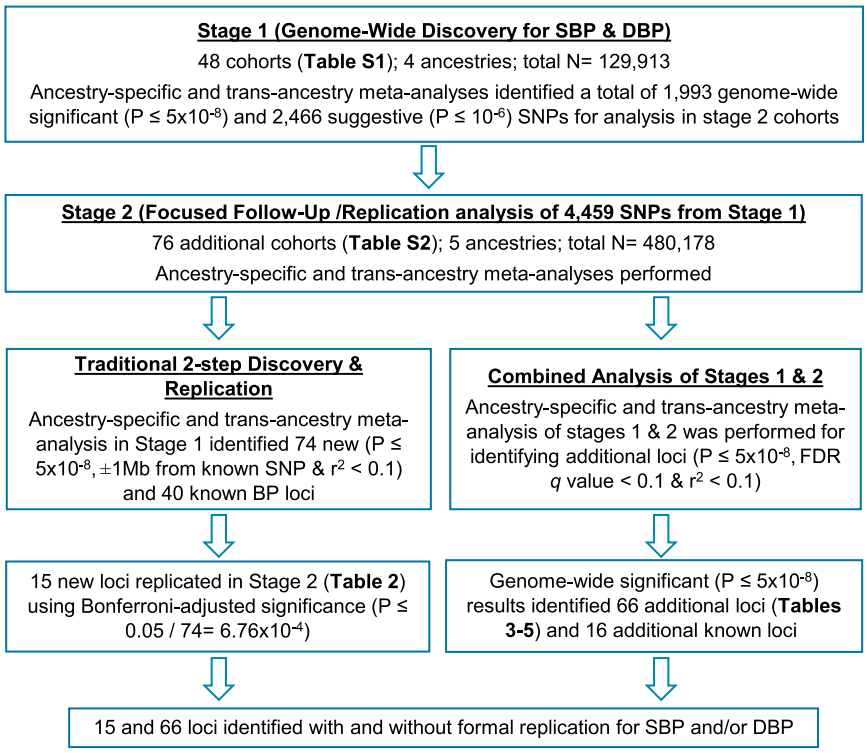


Figure 1. Study Design and Overall Workflow

Stage 1 analysis identified 74 significant novel loci, of which 15 were replicated in stage 2. Replication in stage 2 was hampered by limited sample sizes for African and Hispanic ancestries. Combined analysis leverages the full power of stages 1 and 2, identifying 66 additional BP loci missed by the 2-step approach which were validated by FDR. Association analyses were performed for each of SBP and DBP, accounting for two smoking exposure variables, "current smoking" status (CurSmk) and "ever smoking" status (EverSmk). For each ancestry, cohort-specific results were combined to perform the 1 degree of freedom (df) test of the interaction effect and the 2 df joint test of genetic main and interaction effects.

medication-adjusted BP variables were approximately normally distributed, as shown in Table S5 and Figure S1. In addition, to reduce the influence of possible outliers, winsorizing has been applied for each BP value that was more than six standard deviations away from the mean.

The participating cohorts have varying levels of information on smoking, some with a simple binary variable and others (such as UK Biobank) with more precise data. We considered two dichotomized smoking variables, "current smoking" status (CurSmk) and "ever smoking" status (EverSmk), as they were the most widely available information (Table 1). Current smoking status was coded as 1 if the subject smoked regularly in past year (and as 0 for non-current smokers, which includes both never and former smokers). Ever smoking status was coded as 1 if the subject smoked at least 100 cigarettes during his/her lifetime (and as 0 for the never-smokers). Smoking status was assessed at the time of the BP measurements. When subjects had multiple smoking measures that were inconsistent, they were excluded from analysis. Subjects with missing data for BP, the smoking variable, or any covariates were excluded from analysis.

Genotype Data

Genotyping was performed using Illumina or Affymetrix genotyping arrays. Each study performed imputation to impute genotypes for SNPs, short insertions and deletions (indels), and larger deletions that were not genotyped directly but are available from the 1000 Genomes Project.[18] Information on genotype and imputation for each study is presented in Tables S6 and S7. For imputation, most studies used the 1000 Genomes Project Phase I Integrated Release Version 3 Haplotypes (2010-11 data freeze, 2012-03-14 haplotypes), which contain haplotypes of 1,092 individuals of all ethnic backgrounds.

Cohort-Specific GWAS Analysis

For SBP and DBP separately, each study performed association analyses accounting for two smoking exposure variables, current smoking (CurSmk) and ever smoking (EverSmk). In stage 1, we considered two models to account for gene-smoking interactions. For the first "joint" model, a regression model including both genetic main and GxE interaction effects,

$$\mathbb{E}[Y \mid G, C] = \beta_0 + \beta_E Smk + \beta_G G + \beta_{GE} Smk * G + \boldsymbol{\beta_C C}$$

was applied to the entire sample. For the second "stratified" model, analyses of the genetic main-effect regression models

$$\mathbb{E}[Y \mid C,\ Smk = 0] = \gamma_0^{(0)} + \gamma_G^{(0)} G + \boldsymbol{\gamma_C^{(0)} C}$$

$$\mathbb{E}[Y \mid C,\ Smk = 1] = \gamma_0^{(1)} + \gamma_G^{(1)} G + \boldsymbol{\gamma_C^{(1)} C}$$

were applied separately to the Smk = 0 unexposed group and to the Smk = 1 exposed group (smokers). Y is the medication-adjusted BP value, Smk is the smoking variable (with 0/1 coding for the absence/presence of the smoking exposure), G is the dosage of the imputed genetic variant coded additively (from 0 to 2), and $\mathbf{C}$ is the vector of all other covariates, which include age, sex, field center (for multi-center studies), and principal component (PC) (to account for population stratification and admixture). No additional cohort-specific covariates were included. Our previous work showed that the two (joint and stratified) models provided highly similar inference.[19] Therefore, we considered only the first "joint" model in stage 2.

Each study in stage 1 performed GWAS analysis within each ancestry and provided (1) the estimated genetic main effect β_G, estimated interaction effect β_{GE}, and a robust estimate of the corresponding covariance matrix under the joint model; and (2) estimates of the stratum-specific effects $\gamma_G^{(0)}, \gamma_G^{(1)}$ and robust estimates of their standard errors (SE) under the stratified model. Each study in stage 2 provided estimates of the genetic main effect β_G, the interaction effect β_{GE}, and robust estimates of the corresponding covariance matrix under the joint model at 4,459 select variants. Robust estimates of covariance matrices and SEs were used to

Table 1. Basic Characteristics of Cohorts in Stages 1 and 2 in Each Ancestry

	Current Smoker		Former Smoker		Never Smoker					Age		SBP		DBP	
	N	%	N	%	N	%	% Male	% HT	% HT Meds	Mean	SD	Mean	SD	Mean	SD
Stage 1															
EUR	14,607	18.1	28,409	35.3	37,535	46.6	32.6	38.2	25.4	54.63	8.0	129.31	19.2	77.29	11.2
AFR	5,545	21.5	7,185	27.8	13,121	50.8	26.5	55.9	39.5	54.49	9.1	136.39	22.8	81.75	12.8
ASN	2,465	18.3	1,677	12.5	9,296	69.2	51.2	46.9	27.0	55.42	9.7	137.29	21.5	79.41	11.1
HIS	1,068	12.1	2,160	24.5	5,577	63.3	24.9	43.5	13.3	55.50	11.0	130.50	22.0	76.95	11.8
Stage 1 Total	23,685	18.4	39,431	30.7	65,529	50.9	32.8	43.1	27.7	54.74	8.6	131.69	20.4	78.42	11.6
Stage 2															
EUR	48,198	17.0	89,597	31.6	145,914	51.4	47.8	44.8	25.0	55.91	8.6	139.02	20.4	83.76	11.5
AFR	1,971	29.8	1,579	23.8	3,075	46.4	40.9	54.3	42.8	53.66	10.2	137.00	21.6	83.32	12.8
ASN	29,485	19.8	40,850	27.4	78,597	52.8	54.9	50.3	33.1	60.76	12.3	134.92	20.2	80.01	12.3
HIS	2,739	20.3	2,559	18.9	8,231	60.8	41.0	26.9	16.3	45.86	13.8	124.08	20.0	75.09	11.9
BRZ	998	22.6	514	11.6	2,902	65.8	48.0	15.5	6.3	27.78	3.2	119.91	16.0	74.68	11.5
Stage 2 Total	83,391	18.2	135,099	29.6	238,719	52.2	49.7	45.9	27.4	56.84	9.9	137.12	20.3	82.26	11.8
TOTAL	107,076	18.3	174,530	29.8	304,248	51.9	46.1	45.3	27.4	56.40	9.6	135.96	20.3	81.44	11.7

The cell entries for the covariates and BP traits correspond to sample-size weighted averages across all cohorts in each category.

safeguard against both mis-specification of the mean model and violation of the assumption of constant BP variance across smoking groups (heteroscedasticity).[20,21] Association analysis was performed using various software (Tables S6 and S7). To obtain robust estimates of covariance matrices and robust SEs, studies of unrelated subjects used either the R package sandwich[22] or ProbABEL.[23] To account for relatedness in families, family studies used either the generalized estimating equations (GEE) approach, treating each family as a cluster, or the linear mixed effect model approach with a random polygenic component (for which the covariance matrix depends on the kinship matrix).

Quality Control

Study investigators participating in this study have ample experience in main-effect-based GWASs for multiple phenotypes and are very familiar with validated approaches for quality control (QC) of phenotype, genotype, and imputed data. For example, cohort-level analyses used PCs as covariates to deal with population structure; family studies used suitable software packages to deal with relatedness (Table S6). Overlap among some of the participating cohorts is a potential possibility. However, when there was known overlap of samples across cohorts, one of the cohorts used a non-overlapping sub-sample for their analysis.

We performed extensive QC using the R package EasyQC[24] for all cohort-specific GWAS results. In stage 1, each cohort provided 12 GWAS result files (2 BPs × 2 smoking exposures × 3 analyses, 1 for model 1 and 2 for model 2) for each ancestry group. Each GWAS result file included approximately 8–15 million high-quality variants (depending on ancestry), as cohorts applied a preliminary filter on their imputed data excluding variants with minor allele frequency (MAF) $< 1\%$ or imputation quality measure < 0.1. We performed two QC levels: "study-level" and "meta-level." To identify problems with population substructures or relatedness, we have examined QQ plots and genomic control inflation factors (lambdas) on a study-by-study level (to identify study-specific issues) as well as on the meta-analysis result (to identify cross-study issues). Because GWASs were performed within each ancestry, the "study-level" QC also carefully checked the provided allele frequencies against the retrospective ancestry-specific 1000 Genomes reference panel. Finally, marker names were harmonized to ensure consistencies across cohorts. In addition, we contrasted results from the joint model and stratified models in stage 1 cohorts, as explained elsewhere.[19] The "meta-level" QC reviewed result files of a specific analysis (e.g., SBP-CurSmk-Model1) across all cohorts: this included (1) visually comparing summary statistics (mean, median, standard deviation, inter-quartile range, minimum, maximum) on all effect estimates standard errors (SEs) and p values and (2) examining SE-N and QQ plots to reveal issues with trait transformation[24] or other analytical problems. Any problems found during QC steps, including major differences from the ancestry-specific reference panel and any inflation of lambdas within studies, were communicated and resolved with the individual cohorts. Similar QC steps were applied to cohort-specific results in stage 2. More detailed information about the QC steps, including major QC problems encountered and how they were resolved, are described elsewhere.[16]

The most crucial filter during the meta-analysis was approximate df = min (MAC0, MAC1) * imputation quality measure; this is based on the minor allele count (MAC) in each stratum (MAC0 and MAC1) and imputation quality measure, where $MAC0 = 2 * MAF_{E0} * N_{E0}$ for the unexposed group (with MAF_{E0} and sample size N_{E0} for E = 0 stratum) and $MAC1 = 2 * MAF_{E1} * N_{E1}$ for the exposed group. In meta-analysis, to exclude unstable cohort-specific results that reflect small sample size, low MAF, or low imputation quality measures, variants were excluded if approximate df < 20. This filtering threshold was decided after considering various thresholds and examining the resulting QQ and Manhattan plots. More details are provided in the Supplemental Note. Variants were further excluded if imputation quality

measure < 0.5. This value of 0.5 was used regardless of the software used for imputation, because imputation quality measures are shown to be similar across imputation software.[25]

Meta-analysis

After conducting extensive quality control and selecting high-quality variants, approximately 18.8 million SNPs and small insertion and deletion (indels) variants were included in the meta-analysis (the number of variants varied across the ancestry groups). We performed meta-analysis using both models in stage 1 and using the joint model in stage 2. For both stages, we performed meta-analysis using the 1 degree of freedom (df) test of interaction effect and 2 df tests of testing both SNP main and interaction effects. Wald test statistics approximately follow either a chi-square distribution with 1 df under H_0: $\beta_{GE} = 0$ for the 1 df test or a chi-square distribution with 2 df under H_0: $\beta_G = \beta_{GE} = 0$, for the 2 df test. In the joint model, inverse-variance weighted meta-analysis was performed for the 1 df test and the joint meta-analysis of Manning et al.[26] for the 2 df test, both using METAL.[27] In the stratified model, we performed meta-analysis using the approach of Randall et al.[28] for the 1 df test and the approach of Aschard et al.[29] for the 2 df test. Both tests in the stratified model were computed using the R package EasyStrata.[30] More details are described elsewhere.[19]

Ancestry-specific meta-analyses using inverse-variance weighting were performed to combine cohort-specific results within each ancestry. The ancestry-specific results were then combined through meta-analysis to obtain evidence of "trans-ancestry" association. In stage 1, 80 separate genome-wide meta-analyses were performed: 2 BPs × 2 smoking exposures × 4 (2 tests in the joint model, 2 stratified groups in the stratified model) × 5 ancestries (4 ancestry-specific and 1 trans-ancestry to combine ancestry-specific results). In this stage, genomic control correction[31] was applied twice, first for cohort-specific GWAS results if their genomic control lambda value was greater than 1, and again after the meta-analysis results. Variants were excluded if they were represented by valid data in fewer than 5,000 samples and 3 cohorts. Variants that were genome-wide significant ($p < 5 \times 10^{-8}$) or suggestive ($p < 1 \times 10^{-6}$) in any of stage 1 analyses were pursued for stage 2 analysis. In stage 2, 48 separate meta-analyses were performed using the joint model: 2 BPs × 2 smoking exposures × 2 (2 tests; 1 df and 2 df tests) × 6 ancestries (5 ancestry-specific and 1 trans-ancestry to combine ancestry-specific results). Genomic control correction was not applied to the replication statistics as association analysis was performed only at select variants. Similarly, 48 separate meta-analyses were performed to combine stages 1 and 2 results.

Genome-wide Significant Variants

If a variant reached genome-wide significance ($p < 5 \times 10^{-8}$) through any of these 48 combined association meta-analyses (which are not independent), then the variant was considered as genome-wide significant. To identify a set of independent (index) variants through ancestry-specific and trans-ancestry analysis, we performed the linkage disequilibrium (LD)-based clumping procedure using PLINK[32] and EasyStrata.[30] A locus is defined through LD-based clumping that uses both physical distance (±1 Mb) and LD threshold of $r^2 > 0.1$. Since valid methods do not exist for conditional analysis involving interactions across multi-ancestry studies, we relied on a relatively more stringent LD threshold ($r^2 > 0.1$) for identifying "independent" loci. As LD reference, ancestry-specific 1000 Genomes Project data were used for ancestry-specific results and the entire cosmopolitan dataset was used for trans-ancestry results. False discovery rate (FDR) q-values were computed using the R function p.adjust using the step-up method by Benjamini and Hochberg.[33]

BP Variance Explained

Since variants weakly correlated with index variants ($0.1 \leq r^2 \leq 0.2$) can contribute to the percent variance, for the purposes of calculating percent variance, we carried out clumping using slightly less conservative LD threshold ($r^2 > 0.2$ instead of > 0.1). The percent of variance explained in SBP and DBP by all previously known (158) and newly identified (132 using LD threshold of > 0.2 for clumping) variants was evaluated in several studies from multiple ancestries (see Table S8). BP variants previously identified in any ancestry were considered as "known" variants. Similarly, we considered all index variants representing previously unreported loci as "novel" for this purpose regardless of which ancestry they were identified in; separate interaction terms were included for newly identified variants. Known and newly identified variants (combined from all ancestries) were used in assessing the percent variance.

Percent variance was calculated using standard regression models. Four nested models were considered. The first model included the smoking variables and standard covariates (age, sex, PCs, etc.); the second model included those covariates and all known variants; the third model contained all those previous variables and all newly identified variants (excluding any interaction terms); finally, the fourth model contained all those (covariates, known, and novel) plus the interaction terms. Each of SBP and DBP was regressed on the relevant predictors in each of the four models. The r^2 values obtained from the regressions were used as measures of the percent variance explained by the respective models. Through sequential subtraction of appropriate r^2 values, we determined the "additional" percent variance explained by a given set of variants. For studies with $N < 20{,}000$, we used a stepwise regression procedure with significance tests for inclusion of one variant at a time and for backward elimination of redundant variants.

Functional Inference

Variant Effect Predictor (VEP) from Ensembl was used to obtain the gene name for each locus. For the variants whose gene names were not identified by VEP, NCBI SNP database was used to obtain the closest gene. We applied several computational strategies to infer biological functions associated with our newly identified loci. We used HaploReg, RegulomeDB, and GTEx[34] to obtain annotations of the noncoding genome, chromatin state, and protein binding annotation from the Roadmap Epigenomics and ENCODE projects, sequence conservation across mammals, and the effect of SNPs on expression from eQTL studies. To further assess putative functionality for the new loci, we searched for *cis* associations between new variants and gene transcripts using previously published eQTL analyses, which includes the GTEx.[34]

Further eQTL evidence was queried using the eQTL database of Joehanes et al.[35] for transcripts associated in both *cis* and *trans* in more than 5,000 individuals from the Framingham Heart Study, with genome-wide false discovery rate (FDR) < 0.05. Two gene-set enrichment analysis (GSEA) queries were then performed on December 23, 2016 to determine the enrichment of biological

processes and disease pathways of the resulting transcripts. Prior to the queries, duplicated gene names and genes with provisional names (such as LOCXXX) were removed. Then, for each transcript probe associated with more than one gene name, only the first gene name was taken. This process yielded 127 gene names for the GSEA query. For querying biological processes, option C5:BP was selected on the GSEA website. For querying disease pathway, option C2:CP was selected. Both GSEA queries were set at FDR < 0.05 threshold to guard against multiple comparison errors.

Pathway and Gene Set Enrichment Analysis

We conducted four separate DEPICT analyses based on the following criteria that were applied to our combined association meta-analysis results. We utilized variants showing genome-wide significant joint effect association with (1) SBP in Europeans ($P_{EUR.SBP} < 5 \times 10^{-8}$), (2) DBP in Europeans ($P_{EUR.DBP} < 5 \times 10^{-8}$), (3) SBP in *trans*-ancestry analysis ($P_{Trans.SBP} < 5 \times 10^{-8}$), or (4) DBP in *trans*-ancestry analysis ($P_{Trans.DBP} < 5 \times 10^{-8}$). For each combination, DEPICT first performed the following steps to obtain the input of the prioritization and enrichment analyses: non-overlapping regions lists of independent variants were obtained using 500 kb flanking regions and LD $r^2 > 0.1$ using the 1000 Genomes data,[18] resulting variants were merged with overlapping genes ($r^2 > 0.5$ with a functional coding variant within the gene or *cis*-acting regulatory variant), and the major histocompatibility complex region on chromosome 6 (base position 25,000,000–35,000,000) was excluded.

DEPICT prioritized genes at the associated loci based on their functional similarity. Functional similarity of genes across associated loci was quantified by computing a gene score that was adjusted for bias through confounders such as gene length. Experiment-wide FDR for the gene prioritization was obtained by repeating the scoring step 50 times based on lead variants from 500 pre-compiled null GWASs. For the gene-set enrichment analyses, DEPICT utilized a total of 14,461 pre-compiled reconstituted gene sets comprising 737 Reactome database pathways, 2,473 phenotypic gene sets (derived from the Mouse Genetics Initiative), 184 Kyoto Encyclopedia of Genes and Genomes (KEGG) database pathways, 5,083 Gene Ontology database terms, and 5,984 protein molecular pathways (derived from protein-protein interactions). For the tissue and cell type enrichment analyses, DEPICT tested whether genes harboring associated loci are enriched for expression in any of the 209 MeSH annotations for 37,427 microarrays of the Affymetrix U133 Plus 2.0 Array platform.

To further identify connected gene sets and pathways implicated by our findings, we performed GeneGO analysis and text data mining using Literature Lab.[36] GeneGO (known also as MetaCore) evaluates p values for pathways by mapping a list of target genes to each pathway and comparing those that arise by chance using a hypergeometric distribution formula. GeneGO implements a correction of p values using a false discovery rate. Literature Lab of Acumenta evaluates co-occurrences in the publication records of a list of genes and biological and biochemical terms. The analysis compares the gene input set against the average of 1,000 randomly generated similar size sets, providing a spectrum of statistically significant associations. Our Literature Lab analysis included the use of 17,261,987 PubMed abstracts, out of which 10,091,778 abstracts include one or more human genes.

Results

Study Overview

We performed the traditional 2-step approach with discovery in stage 1 followed by formal replication in stage 2. Because this study was not optimally designed for replications in non-EUR (especially in AFR) ancestry, to identify additional loci, we performed combined analysis of stages 1 and 2 to maximize power for discovery[37] (Figure 1). For the 2-step approach, we performed ancestry-specific meta-analysis in each of five ancestries and trans-ancestry analysis in stage 2. We checked whether each of the genome-wide significant loci in stage 1 was replicated in stage 2 using Bonferroni-adjusted significance level (0.05/74, see details below). For the combined analysis, we performed ancestry-specific meta-analysis combining both stages 1 and 2 (discovery and follow-up) in each of 5 ancestries; these ancestry-specific meta-analyses results were then combined to perform trans-ancestry analysis at 4,459 variants using a total of up to 610,091 individuals.

Two-Step Approach of Discovery Followed by Replication

Of the 4,459 significant or suggestive variants selected from stage 1 meta-analyses, 3,222 were replicated in stage 2 with $p < 0.05/4,459$ (to an aggregate replication rate of 72.3%). Of the 1,993 variants that were genome-wide significant ($p < 5 \times 10^{-8}$) in stage 1 analysis, 1,836 were replicated in stage 2 with $p < 0.05/1,993$ to a replication rate of 92.1%. These 1,993 genome-wide significant variants in stage 1 belong to 114 independent loci. Of the 114 loci, 40 loci (consisting of 1,644 variants) contain previously published BP variants.[1,3–7] Of the remaining 74 newly identified loci (consisting of 349 variants), 15 loci were formally replicated in stage 2 using Bonferroni-adjusted significance level ($p < 0.05/74$) (Table 2); all 15 novel loci were replicated even when using the more conservative adjustment threshold $p < 0.05/349$. In addition, 25 more of the remaining 59 loci were nominally replicated ($p < 0.05$) in one or more of the analyses in stage 2 ($p < 0.05$), and 27 more showed the same direction of effect in stages 1 and 2. For 7 loci, no additional data were available in stage 2 and, therefore, it was not possible to check for replication. For the 15 formally replicated loci, estimates of the genetic main effects were all consistent between stages 1 and 2; estimates of SNP-smoking interaction effects were not statistically significant (forest plots; Figure S3). All of the 15 replicated loci were genome-wide significant in European ancestry. Furthermore, 10 loci also had supporting evidence from non-European ancestry, resulting in stronger statistical significance from trans-ancestry analysis (Figure S3, Table 2). Quantile-quantile (QQ) plots for the genome-wide stage 1 meta-analysis are shown in Figure S2.

Of the 15 formally replicated loci, six loci (indicated by f in Table 2) are least 1 Mb away from any previously

Table 2. Newly Identified Loci that Are Significant in Stage 1 and Formally Replicated in Stage 2

Locus[a]	Nearest Genes[b]	rsID	Chr:Pos[c]	EA	EAF	Ancestry and Trait	Stage	Genetic Main Effect Est[d]	Genetic Main Effect SE[d]	Interaction Effect Est[d]	Interaction Effect SE[d]	2 df Joint p Value[e]
1	*MTHFR;CLCN6*;NPPA*	rs202071545	1:11878161	d	0.945	ALL.SBP	1	1.24	0.28	−0.16	0.38	3.77×10^{-8}
							2	0.88	0.17	0.01	0.25	7.44×10^{-12}
							1+2	0.99	0.14	−0.04	0.20	$*9.39 \times 10^{-20}$
2	*CLCN6;NPPA;NPPB**	rs3753581	1:11920189	a	0.327	ALL.SBP	1	−0.63	0.09	0.16	0.21	4.34×10^{-12}
							2	−0.43	0.05	0.00	0.11	5.52×10^{-23}
							1+2	−0.48	0.04	0.04	0.10	$*1.31 \times 10^{-34}$
3	*NPPA;NPPB*	rs72640287	1:11965792	t	0.039	EUR.SBP	1	−2.05	0.43	−0.04	0.59	1.59×10^{-10}
							2	−0.86	0.19	−0.33	0.28	8.19×10^{-13}
							1+2	−1.06	0.18	−0.31	0.25	$*2.79 \times 10^{-21}$
4	*WNT2B**	rs351364	1:113045061	a	0.297	ALL.SBP	1	−0.60	0.10	0.53	0.22	1.67×10^{-8}
							2	−0.42	0.05	0.14	0.11	5.38×10^{-19}
							1+2	−0.45	0.04	0.22	0.10	$*1.20 \times 10^{-26}$
5[f]	*CEP170;SDCCAG8*;AKT3*	rs3897821	1:243420388	a	0.705	ALL.DBP	1	−0.35	0.06	0.20	0.13	2.49×10^{-9}
							2	−0.20	0.03	0.00	0.07	1.51×10^{-12}
							1+2	−0.23	0.03	0.05	0.06	$*1.67 \times 10^{-20}$
6[f]	*FER1L5**	rs7599598	2:97351840	a	0.564	EUR.DBP	1	−0.30	0.06	−0.15	0.14	5.93×10^{-8}
							2	−0.16	0.03	0.02	0.08	4.10×10^{-7}
							1+2	−0.19	0.03	−0.03	0.07	$*4.25 \times 10^{-13}$
7	*SLC4A7**	rs13063291	3:27446285	a	0.204	ALL.DBP	1	0.33	0.08	0.03	0.12	4.00×10^{-8}
							2	0.20	0.04	−0.14	0.06	3.75×10^{-6}
							1+2	0.23	0.04	−0.09	0.05	$*1.67 \times 10^{-11}$
8[f]	*PRAG1;MFHAS1*	rs7823056	8:8382705	a	0.397	EUR.SBP	1	−0.56	0.10	−0.02	0.22	1.54×10^{-8}
							2	−0.42	0.05	0.16	0.13	1.55×10^{-14}
							1+2	−0.45	0.05	0.10	0.11	$*3.01 \times 10^{-22}$
9[f]	*PPP1R3B;TNKS*	rs62493780	8:9151051	t	0.238	EUR.SBP	1	0.89	0.18	−0.19	0.25	3.47×10^{-8}
							2	0.46	0.09	−0.27	0.13	2.37×10^{-7}
							1+2	0.54	0.08	−0.24	0.12	$*2.95 \times 10^{-13}$

(*Continued on next page*)

Table 2. ***Continued***

Locus[a]	Nearest Genes[b]	rsID	Chr:Pos[c]	EA	EAF	Ancestry and Trait	Stage	Genetic Main Effect Est[d]	Genetic Main Effect SE[d]	Interaction Effect Est[d]	Interaction Effect SE[d]	2 df Joint p Value[e]
10[f]	*MIR124-1*;MSRA*	rs13271489	8:9803712	t	0.478	EUR.SBP	1	0.55	0.10	0.02	0.22	6.37×10^{-8}
							2	0.44	0.05	−0.13	0.14	9.35×10^{-16}
							1+2	0.46	0.05	−0.08	0.12	$*4.56 \times 10^{-23}$
11	*TNNI2;LSP1*;TNNT3*	rs7483477	11:1920255	t	0.75	ALL.SBP	1	−0.65	0.11	0.17	0.27	2.25×10^{-8}
							2	−0.36	0.05	0.08	0.12	1.77×10^{-13}
							1+2	−0.40	0.04	0.09	0.11	$*2.12 \times 10^{-20}$
12	*POC1B;ATP2B1*	rs7313874	12:89965049	t	0.325	ALL.SBP	1	−0.64	0.11	0.01	0.15	1.85×10^{-14}
							2	−0.48	0.06	−0.23	0.09	1.07×10^{-39}
							1+2	−0.52	0.05	−0.17	0.08	$*2.49 \times 10^{-54}$
13	*ATP2B1**	rs111337717	12:90037506	t	0.943	ALL.SBP	1	1.27	0.33	0.60	0.46	9.23×10^{-11}
							2	1.09	0.15	−0.26	0.22	2.86×10^{-18}
							1+2	1.13	0.13	−0.07	0.20	$*1.27 \times 10^{-27}$
14	*PTPN11*	rs7974266	12:113007602	t	0.513	ALL.DBP	1	0.19	0.13	0.57	0.18	3.58×10^{-8}
							2	0.23	0.06	0.19	0.09	6.50×10^{-12}
							1+2	0.22	0.06	0.28	0.08	$*5.91 \times 10^{-19}$
15[f]	*AKTIP;RPGRIP1L;FTO**	rs11642015	16:53802494	t	0.334	ALL.SBP	1	0.57	0.09	−0.19	0.21	2.78×10^{-9}
							2	0.29	0.05	0.08	0.11	6.74×10^{-13}
							1+2	0.35	0.04	0.03	0.10	$*9.91 \times 10^{-21}$

Each locus is genome-wide significant ($p < 5 \times 10^{-8}$) in stage 1 and formally replicated in stage 2 using Bonferroni-adjusted significance level ($p < 0.05/74$). Forest plots and LocusZoom plots are shown in Figures S3 and S4, respectively. Abbreviations: BP, blood pressure; SBP, systolic BP; DBP, diastolic BP; EA, effect allele; EAF, effect allele frequency; 2 df joint p, p value of the joint test with 2 degrees of freedom of genetic main and interaction effects; 1 df interaction p, p value of the interaction test with 1 degree of freedom; EUR, European ancestry; ALL, trans-ancestry (i.e., combining all ancestry groups through meta-analysis).

[a]Each locus was determined through LD-based clumping, using ± 1 Mb around index variants, followed by LD threshold of $r^2 > 0.1$; ancestry-specific LDs from 1000 Genomes Project were used when clumping within each ancestry and the entire cosmopolitan data were used for trans-ancestry clumping.

[b]Gene names were obtained using variant effect predictor (VEP) from Ensembl. Genes with intragenic index variants are indicated with an asterisk (*).

[c]Positions are based on build 37.

[d]Effect is in mmHg unit.

[e]The most significant p value (between 1 df interaction test and 2 df joint test) are indicated with an asterisk (*).

[f]These loci indicate "completely novel" loci, at least 1 Mb away from any of known BP loci.

published BP variants, and we term them "completely novel." Three of them (near *PRAG1*, *MIR124-1*, and *FTO*) show compelling biological relevance (see below) and eQTL evidence (Figure 2). The locus zoom plots of all newly identified loci identified in this paper are shown in Figure S4. The remaining 9 loci are novel signals (which meet our definition of a locus) near but not in LD ($r^2 < 0.1$) with known BP loci. For example, near the well-known BP locus *ATP2B1* on chromosome 12, there were two independent signals identified in European ($p = 4.1 \times 10^{-41}$), Asian ($p = 1.5 \times 10^{-13}$), and trans-ancestry ($p = 2.5 \times 10^{-54}$) analyses. Near another well-known BP locus, *MTHFR-NPPB-CLCN6*, we identified three additional independent signals (with p values as small as 4.3×10^{-34} at index variants, spanning 196 kb [from 11,827,796 to 12,023,500] on chromosome 1).

Combined Analysis of Stages 1 and 2

Combined meta-analysis of stages 1 and 2 identified a total of 82 additional independent loci ($p < 5 \times 10^{-8}$) not identified by the 2-step approach. Association statistics for all genome-wide significant variants in the combined meta-analysis are provided in Table S9. Manhattan plots of the combined meta-analysis for each BP trait using the 1 df interaction and 2 df joint tests are shown in Figures S5–S8. Summary Manhattan plots for SBP and DBP with the minimum p values across all analyses are shown in Figure S9. QQ plots are shown in Figure S10.

Of these 82 additional loci identified through combined analysis, 16 loci contain previously published BP variants.[1,3–7] All of the remaining 66 loci had a low false discovery rate (FDR q value < 0.1 for all 66 loci and < 0.01 for 60 of the loci, Table S10). Of these 66 loci, 18 and 13 loci were identified through trans-ancestry (Table 3) and European ancestry (Table 4), respectively. Except for one locus, they were suggestive ($p < 1 \times 10^{-6}$) in stage 1 analyses but became significant in the combined stages 1 and 2 meta-analysis (Tables 3, 4, and 5). The strength of the combined analysis was exemplified by a locus in *HOTTIP* on chromosome 7 (locus 4 in Table 3), which were suggestive in stage 1 analysis ($p = 9.4 \times 10^{-7}$) and identified through the combined analysis in European ($p = 6.0 \times 10^{-29}$), Asian ($p = 1.2 \times 10^{-10}$), and trans-ancestry ($p = 3.6 \times 10^{-41}$, see Figure S3). Genome-wide significant loci from trans-ancestry analysis did not show strong evidence of heterogeneity across ancestry groups.

Of the 66 identified loci, 35 were found through African-ancestry only (Table 5). These loci were mostly low frequency with MAF between 1% and 5% (Table 5). Of these 35 loci, 4 were genome-wide significant in stage 1 African ancestry and stayed significant in the combined analysis (although not formally replicated in stage 2). One such locus was near *BMP7* on chromosome 20 (with $p = 5.8 \times 10^{-10}$ in stage 1; $p = 0.03$ in stage 2; $p = 4.2 \times 10^{-12}$ in stages 1+2). Six loci were suggestive ($p < 1 \times 10^{-6}$) in stage 1 analyses but became significant in the combined stages 1 and 2 meta-analysis. One such locus was near *WSCD1* on chromosome 17 (with $p = 8.7 \times 10^{-7}$ in stage 1; $p = 0.00047$ in stage 2; $p = 1.8 \times 10^{-10}$ in stages 1+2). The remaining 25 loci were genome-wide significant in stage 1 African ancestry but not represented in stage 2 African ancestry due to limited sample sizes and low MAF. Furthermore, 15 loci were African-specific loci; they had MAF $< 1\%$ in the other ancestry groups and were filtered out by the individual studies (by design), and therefore results are unavailable for further analysis. In the non-AFR ancestry results, genome-wide significant variants at newly identified loci were mostly common (with MAF $\geq 5\%$) and had similar MAF distributions as those at known loci (Figure S10).

Known BP Loci

At most of the 56 known BP loci[1,3–7] identified in the two-step or combined analyses, the lead variant identified by our analyses was the same as the one previously published (Table S11); European, Asian, and trans-ancestry results identified 48, 14, and 50 of these variants, respectively. In the remaining loci, our results identified a variant in the same locus as the known BP variant. The most significant results were observed at well-known BP loci: *ATP2B1* (rs17249754 on chromosome 12, trans-ancestry $P_{SBP} = 4.8 \times 10^{-85}$; $P_{DBP} = 5.5 \times 10^{-57}$) and *SH2B3-ATXN2* (rs3184504 on chromosome 12, trans-ancestry $P_{SBP} = 3.2 \times 10^{-36}$; $P_{DBP} = 6.0 \times 10^{-67}$).

The Role of Interactions

Interaction effects contributed in varying degrees to the evidence of association for the 81 newly reported genome-wide significant loci (Tables 2, 3, 4, and 5). The genetic effects of these new index variants (each index variant representing a locus with the smallest p value) were different in smokers and non-smokers, thus highlighting the potentially important role of interactions (Figure 3). Among the 81 index variants, 10 variants showed genome-wide significant interactions with smoking exposure status (1 df interaction $p < 5 \times 10^{-8}$). All 10 of these variants, most of which were identified in African ancestry, show larger effects on BP in smokers (Figure 3). However, none of the interactions were replicated in stage 2. In addition, of the 158 previously reported BP variants, two (rs3752728 in *PDE3A* and rs3184504 in *SH2B3-ATXN2*) show significant evidence of interactions with smoking using Bonferroni correction (1 df interaction $p < 0.05/158$). 27 additional variants show nominal evidence of interaction (with $p < 0.05$).

To minimize spurious results, we winsorized extreme BP values and used robust standard errors in cohort-specific analyses. Moreover, since non-normality and unequal BP variances among smokers and non-smokers can lead to false positives, we examined these characteristics in three large studies (ARIC, UK Biobank, and WGHS). The distributions look very similar in exposed and unexposed groups (histograms in Figure S1). The variances across strata are also very similar (Table S5). Moreover, on average across

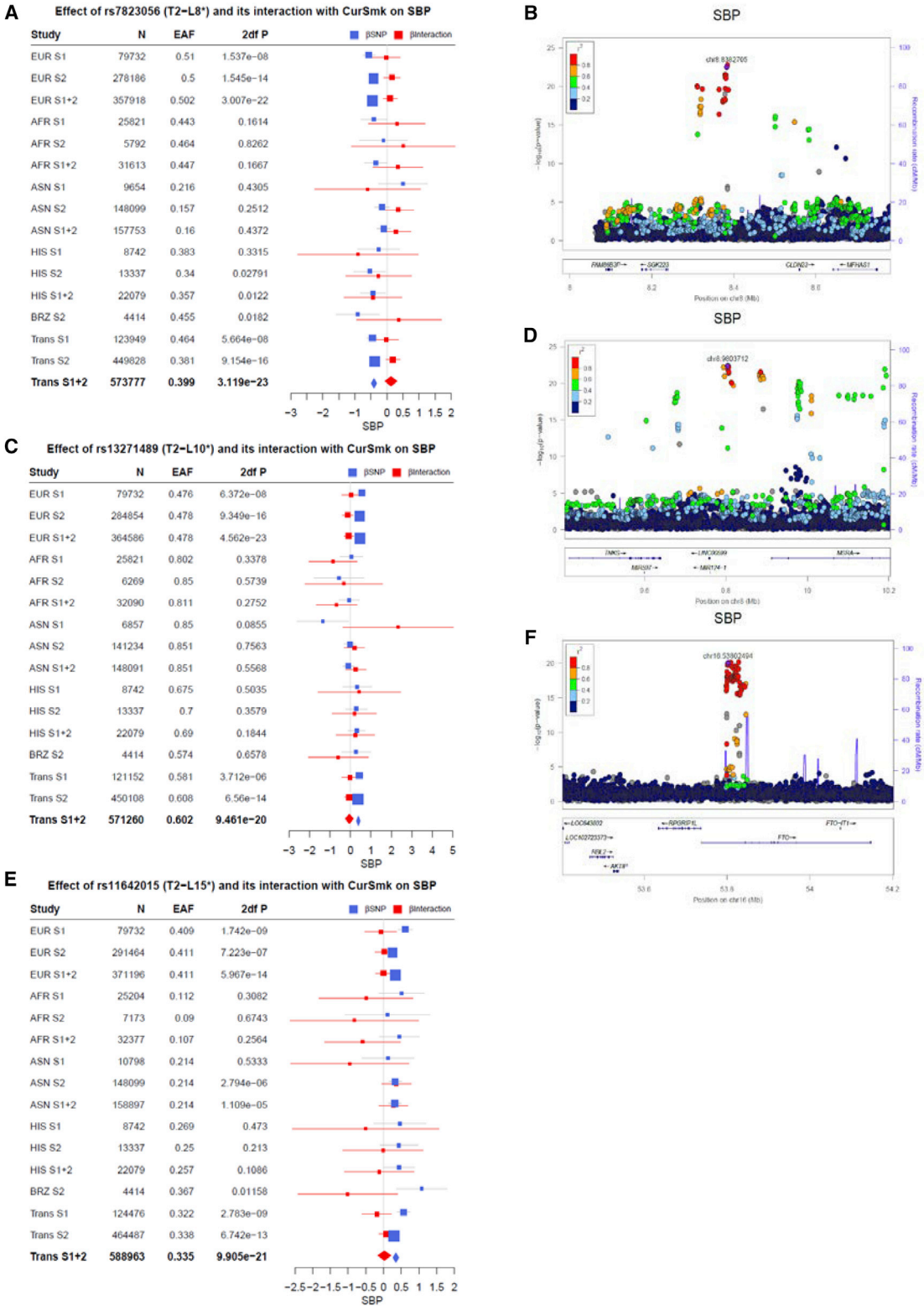


Figure 2. Forest Plots and LocusZoom Plots for Three Newly Identified Loci

(A and B) Variant rs7823056 and 10 additional variants on chromosome 8 are an eQTL for *PRAG1*, which is expressed in multiple tissues including the cerebellum and thyroid.

(C and D) Variant rs13271489 is a *cis*-eQTL for *MSRA* and predicted to modify enhancers in brain cells. *MSRA* has been shown to be associated with obesity-related traits and adipocyte function; it also promotes the survival and development of dopaminergic neurons.

(E and F) Variant rs11642015 is intronic to the well-known obesity/diabetes locus *FTO*. In addition, *AKTIP* in this locus has role in telomere maintenance.

Loci selected from Table 2.

Table 3. Additional Significant Loci from the Combined Trans-Ancestry Analyses of Stages 1 and 2

Locus[a]	Nearest Genes[b]	rsID	Chr:Pos[c]	EA	EAF	Effect[d]		p Value[e]		Trait
						Genetic Main	Interaction	1 df Interaction	2 df Joint	
1	*NPPA;NPPB*	rs12741980	1:11939593	a	0.943	0.68	0.02	0.852	$*3.04 \times 10^{-14}$	SBP
2[f]	*RSRC1**	rs201851995	3:157837508	d	0.648	−0.6	0.38	0.0016	$*4.65 \times 10^{-12}$	SBP
3[f]	*INPP4B;GAB1*	rs78763922	4:144054552	d	0.303	0.34	0.05	0.5067	$*4.03 \times 10^{-13}$	SBP
4	*HOTTIP**	rs2023843	7:27243221	t	0.837	0.7	−0.2	0.1634	$*3.69 \times 10^{-41}$	SBP
5[f]	*MFHAS1*;ERI1;PPP1R3B*	rs201133964	8:8607849	d	0.174	−0.52	−0.16	0.4366	$*1.24 \times 10^{-9}$	SBP
6[f]	*PPP1R3B;TNKS*	rs35904419	8:9376810	d	0.816	−0.19	−0.15	0.1761	$*1.34 \times 10^{-8}$	DBP
7	*FAM167A-AS1*;FAM167A;BLK*	rs4841531	8:11293390	t	0.161	−0.31	0.03	0.7825	$*1.32 \times 10^{-8}$	SBP
8[f]	*EBF2;LOC105379336*; PPP2R2A;DPYSL2;ADRA1A*	rs58429174	8:26011922	t	0.262	−0.12	−0.14	0.026	$*2.60 \times 10^{-9}$	DBP
9	*ADRB1*	rs180940	10:115722411	a	0.391	−0.19	0.06	0.1514	$*5.00 \times 10^{-12}$	DBP
10	*AP5B1;OVOL1*	rs201316070	11:65548558	d	0.061	−0.6	−0.23	0.462	$*1.54 \times 10^{-9}$	SBP
11[f]	*LRP6;GPR19;APOLD1*; GPRC5A*	rs72656645	12:12881055	a	0.7	0.36	−0.13	0.064	$*4.49 \times 10^{-15}$	SBP
12	*SLCO1C1;SLCO1B3; SLCO1B7; SLCO1B1*	rs73073686	12:20354507	a	0.231	−0.24	−0.07	0.2553	$*1.68 \times 10^{-18}$	DBP
13	*ATP2B1*	rs10858948	12:90478651	a	0.578	−0.18	0	0.6992	$*4.74 \times 10^{-15}$	DBP
14	*MED13L*	rs11067762	12:116198214	a	0.176	−0.24	−0.05	0.1951	$*5.30 \times 10^{-18}$	DBP
15	*CYP1A1-2;ULK3;SCAMP2*;MPI*	rs10628234	15:75211142	d	0.3	0.32	−0.22	0.0253	$*1.57 \times 10^{-24}$	DBP
16[f]	*LDHD;CFDP1*;TMEM231; TERF2IP*	rs4888411	16:75443183	a	0.56	0.26	0.12	0.0467	$*1.19 \times 10^{-18}$	SBP
17[f]	*SLC2A4;KCTD11;TNFSF12*; TNFSF13;ATP1B2*	rs9899183	17:7452977	t	0.742	−0.35	0.07	0.6683	$*1.24 \times 10^{-12}$	SBP
18[f]	*ACE**	rs4968782	17:61548476	a	0.616	−0.2	0.08	0.2179	$*3.30 \times 10^{-16}$	DBP

Each locus is genome-wide significant ($p < 5 \times 10^{-8}$) in the combined analyses of stages 1 and 2 and had FDR *q* value < 0.1. Forest plots and LocusZoom plots are shown in Figures S3 and S4, respectively. Abbreviations: BP, blood pressure; SBP, systolic BP; DBP, diastolic BP; EA, effect allele; EAF, effect allele frequency; 2 df joint p, p value of the joint test with 2 degrees of freedom of genetic main and interaction effects; 1 df interaction p, p value of the interaction test with 1 degree of freedom.

[a]Each locus was determined through LD-based clumping, using ± 1 Mb around index variants, followed by LD threshold of $r^2 > 0.1$; ancestry-specific LDs from 1000 Genomes Project were used when clumping within each ancestry and the entire cosmopolitan data were used for trans-ancestry clumping.

[b]Gene names were obtained using variant effect predictor (VEP) from Ensembl. Genes with intragenic index variants are indicated with an asterisk (*).

[c]Positions are based on build 37.

[d]Effect is in mmHg unit.

[e]The most significant p value (between 1 df interaction test and 2 df joint test) is indicated with an asterisk (*).

[f]These loci indicate "completely novel" loci, at least 1 Mb away from any of known BP loci.

all stage 1 cohorts, skewness is 0.64 for SBP and 0.36 for DBP; kurtosis is 3.52 for SBP and 3.32 for DBP (Table S3). There do not seem to be substantial deviations from normality although moderate deviations exist. Therefore, it is less likely that the interaction effects at these 10 newly identified loci are spurious.

BP Variance Explained

In several large cohorts, we calculated the percent of BP variance explained by various loci across four ancestries (Table S8). The variance explained by the 158 previously known loci ranges from 1.1% (in HIS) to 3.2% (in EUR) for SBP and ranges from 1.6% (in ASN and HIS) to 3.4% (in AFR) for DBP. The additional variance explained by the newly identified loci and their interactions ranges from 0.6% (in EUR) to 2.6% (in AFR) for SBP and ranges from 0.3% (in ASN) to 3.2% (in AFR) for DBP. The percent variance explained is ideally calculated in large individual studies which did not participate in our analysis in stage 1 or 2. However, having recruited most of the studies available to us into stage 1 or 2 (for maximizing power), we had to use some of the same studies for this purpose and therefore some of the variance estimates may be somewhat inflated. In an independent EUR study (Airwave study, N = 14,002) that did not participate in stage 1 or 2, known variants explained 1.6% of variance in SBP and DBP, and newly identified variants and their interactions explained 1.2% variance in SBP and 1.3% variance in DBP (Table S8). These variances are within the ranges noted, lending credibility to the results from other studies. Note that

Table 4. Additional Significant Loci from the Combined Analyses of Stages 1 and 2 in European Ancestry

Locus[a]	Nearest Genes[b]	rsID	Chr:Pos[c]	EA	EAF	Effect[d]		P value[e]		Trait
						Genetic Main	Interaction	1 df Interaction	2 df Joint	
1	*MTHFR*;CLCN6*	rs6541006	1:11857526	a	0.071	−0.85	0	0.6454	$*3.17 \times *10^{-19}$	SBP
2[f]	*KCNG3;DYNC2LI1*	rs73923009	2:43141074	a	0.099	−0.36	0.07	0.6165	$*1.21 \times 10^{-14}$	DBP
3	*SLC17A1-4;HFE*	rs7753826	6:26042239	a	0.189	0.36	−0.05	0.4371	$*1.72 \times 10^{-25}$	DBP
4	*SLC44A4;EHMT2*; STK19; CYP21A2;TNXB*	rs2243873	6:31863433	a	0.556	0.45	−0.19	0.0472	$*3.33 \times 10^{-14}$	SBP
5	*SLC44A4;EHMT2; HLA-DQB2*; STK19;CYP21A2;TNXB*	rs2071550	6:32730940	a	0.307	0.29	−0.22	0.0003	$*1.17 \times 10^{-9}$	DBP
6[f]	*TNKS;MSRA*	rs4841235	8:9683358	a	0.426	0.37	−0.1	0.7078	$*4.78 \times 10^{-15}$	SBP
7	*SOX7*;PINX1*	rs6995692	8:10587008	c	0.563	−0.44	0.31	0.0102	$*4.11 \times 10^{-19}$	SBP
8[f]	*ADARB2**	rs150155092	10:1769881	d	0.013	4.76	−18.32	$*7.43 \times 10^{-9}$	1.94×10^{-8}	SBP
9	*KAT5;RNASEH2C*	rs72941051	11:65478893	t	0.074	−0.39	0.07	0.3701	$*1.75 \times 10^{-11}$	DBP
10[f]	*FAM19A2*;AVPR1A*	rs17713040	12:62467714	t	0.977	0.24	0.31	0.7633	$*3.44 \times 10^{-8}$	DBP
11	*FAM109A;SH2B3*;ATXN2*	rs4375492	12:111835990	a	0.794	0.35	0.03	0.8187	$*1.03 \times 10^{-26}$	DBP
12	*MPI;COX5A;SCAMP5*	rs12050494	15:75260896	a	0.316	0.32	−0.06	0.525	$*3.01 \times 10^{-27}$	DBP
13[f]	*NAA38*;KCNAB3;VAMP2*	rs74439044	17:7781019	t	0.903	−0.36	−0.14	0.1507	$*2.43 \times 10^{-21}$	DBP

Each locus is genome-wide significant ($p < 5 \times 10^{-8}$) in the combined analyses of stages 1 and 2 and had FDR q value < 0.1. Forest plots and LocusZoom plots are shown in Figures S3 and S4, respectively. Abbreviations: BP, blood pressure; SBP, systolic BP; DBP, diastolic BP; EA, effect allele; EAF, effect allele frequency; 2 df joint p, p value of the joint test with 2 degrees of freedom of genetic main and interaction effects; 1 df interaction p, p value of the interaction test with 1 degree of freedom.

[a]Each locus was determined through LD-based clumping, using ± 1 Mb around index variants, followed by LD threshold of $r^2 > 0.1$; ancestry-specific LDs from 1000 Genomes Project were used when clumping within each ancestry and the entire cosmopolitan data were used for trans-ancestry clumping.

[b]Gene names were obtained using variant effect predictor (VEP) from Ensembl. Genes with intragenic index variants are indicated with an asterisk (*).

[c]Positions are based on build 37.

[d]Effect is in mmHg unit.

[e]The most significant p value (between 1 df interaction test and 2 df joint test) was set in bold.

[f]These loci indicate "completely novel" loci, at least 1 Mb away from any of known BP loci.

both known and newly identified variants (with their interactions) explain some of the BP variance across ancestry groups.

Functional Annotation and eQTL Evidence

For all 81 index variants representing the newly identified loci, we obtained functional annotations using HaploReg[38] and RegulomeDB.[39] There were 2 coding variants (1 missense and 1 synonymous). Of the remaining non-coding variants (29 intronic and 52 intergenic), 17 are located in promoter histone marks, 53 in enhancer histone marks, 29 in DNase I marks, and 10 altered the binding sites of regulatory proteins (Table S12). Conserved among vertebrates were 6 variants as identified via GERP[40] and 5 variants via SiPhy.[41] RegulomeDB assigned class 1f (strong evidence for enhancer function) for 2 variants (Table S12), each of which likely affects the binding of regulatory elements and is linked to expression of a gene target. Of these, rs12741980 (locus 2, Table 4) is near the well-known BP locus *MTHFR-NPPB-CLCN6* and a *cis*-acting expression quantitative trait locus (eQTL) for *NPPA-AS1*, which is expressed in multiple tissues, including thyroid and whole blood. Also, newly identified variant rs180940 (locus 10, Table 4), with RegulomeDB score of 1f, is a *cis*-eQTL for the known locus *ADRB1*, an adrenergic receptor that mediates effects of the hormone epinephrine and the neurotransmitter norepinephrine,[42] although it is about 80 kb upstream of this locus. Of note, our results identified this known BP locus (rs2782980, $p = 1.1 \times 10^{-21}$ and rs1801253, $p = 1.3 \times 10^{-22}$, in Table S11).

Among the 81 newly identified index variants, *cis*-eQTL evidence was available for 39 variants with varying degrees of association with expression probes (Table S12). In particular, 21 of them were identified by GTEx[34] as *cis*-eQTLs across various tissues (Table S13). However, most of them are for *cis*-eQTLs that differ from their nearest assigned genes. For example, an intronic variant in *WNT2B* (rs351364) is a *cis*-eQTL for *RHOC*, which serves as a microtubule-dependent signal that is required for the myosin contractile ring formation during cell cycle cytokinesis. Additionally, 11 variants (including rs7823056 in Figure 2) on chromosome 8 are *cis*-eQTLs for *PRAG1*, which is expressed in multiple tissues including the cerebellum and thyroid. The most abundant evidence of *cis*-eQTL association (with 44 eQTL hits from multiple studies) was observed for rs2243873, a intronic variant of *EHMT2*; it is predicted to regulate expression of many genes including *HLA-C*, *HLA-B*, and *HLA-DRB1* across multiple tissues.

Table 5. Additional Significant Loci from the Combined Analyses of Stages 1 and 2 in African Ancestry

Locus[a]	Nearest Genes[b]	rsID	Chr:Pos[c]	EA	EAF	Effect[d]		p Value[e]		Trait
						Genetic Main	Interaction	1 df Interaction	2 df Joint	
1[f]	*AJAP1**	rs12135881	1:4781922	c	0.988	−2.05	16.94	2.06×10^{-8}	$*3.09 \times 10^{-9}$	SBP
2[f]	*FABP3;SERINC2;TINAGL1*	rs11809589	1:31970118	a	0.012	−1.11	−18.04	1.54×10^{-7}	$*7.71 \times 10^{-10}$	SBP
3[f]	*LOC101928219*	rs182662555	1:96289336	t	0.988	6.15	−4.45	0.00201	$*1.79 \times 10^{-8}$	DBP
4[f]	*PXDN;MYT1L**	rs75247762	2:1893133	t	0.014	−2.37	−12.93	1.45×10^{-5}	$*1.17 \times 10^{-9}$	SBP
5[f]	*ASB3;ERLEC1;GPR75*	rs115234772	2:53650295	a	0.987	−0.1	8.5	2.13×10^{-9}	$*1.07 \times 10^{-11}$	DBP
6[f]	*SERTAD2;SLC1A4*	rs145162854	2:65104447	a	0.015	−3.17	−2.61	0.171	$*6.63 \times 10^{-9}$	SBP
7[f]	*ACOXL**	rs116008367	2:111807546	c	0.014	−0.86	−5.35	5.00×10^{-5}	$*3.09 \times 10^{-8}$	DBP
8[f]	*KCNE4;SCG2*	rs10166552	2:224036537	t	0.016	−0.15	−10.83	4.28×10^{-6}	$*1.52 \times 10^{-9}$	SBP
9[f]	*TPRA1*;MCM2*	rs139963642	3:127314188	t	0.013	−6.35	1.23	0.6742	$*1.55 \times 10^{-8}$	DBP
10[f]	*PCDH7*	rs11931572	4:30086104	a	0.968	−0.45	3.28	2.71×10^{-6}	$*2.91 \times 10^{-8}$	DBP
11[f]	*SPRY1;LINC01091**	rs62319742	4:124581262	a	0.014	1.98	−10.98	$*3.43 \times 10^{-8}$	4.09×10^{-8}	DBP
12[f]	*HSD17B4*	rs140543491	5:118923601	a	0.017	−3	−16.29	1.24×10^{-5}	$*5.34 \times 10^{-9}$	SBP
13[f]	*OFCC1*	rs148387718	6:9446000	t	0.014	0.59	−7.84	2.70×10^{-8}	$*1.77 \times 10^{-11}$	DBP
14[f]	*NEDD9;LOC105374928**	rs9348895	6:11496048	a	0.586	0.11	1.21	6.15×10^{-6}	$*1.71 \times 10^{-8}$	DBP
15[f]	*MYO6;IMPG1**	rs58806982	6:76688806	t	0.01	−11.24	14.92	1.47×10^{-5}	$*4.57 \times 10^{-8}$	SBP
16[f]	*TARID*;SLC2A12*	rs76987554	6:134080855	t	0.062	−1.57	0	0.6676	$*1.63 \times 10^{-8}$	SBP
17[f]	*ARID1B**	rs112140754	6:157245233	t	0.988	0.97	7.6	0.00104	$*2.44 \times 10^{-8}$	DBP
18[f]	*BZW2**	rs116196735	7:16710605	a	0.018	−2.88	−13.75	0.00037	$*6.98 \times 10^{-10}$	SBP
19[f]	*MED30;EXT1*	rs74701635	8:118758316	t	0.016	3.79	−19.2	2.38×10^{-9}	$*2.13 \times 10^{-9}$	SBP
20[f]	*ADAMTSL1*;MIR3152*	rs146250839	9:18189778	a	0.976	0.35	2.79	0.00029	$*4.36 \times 10^{-8}$	DBP
21[f]	*SPIN1;S1PR3;SHC3;CKS2*	rs192642798	9:91503987	a	0.012	−8.38	3.95	0.346	$*4.23 \times 10^{-9}$	SBP
22[f]	*FZD8*	rs76726877	10:36313497	t	0.015	−1.55	−9.14	4.17×10^{-6}	$*4.47 \times 10^{-10}$	DBP
23[f]	*SFRP5;CRTAC1**	rs11599481	10:99640463	t	0.058	−0.9	−3.33	1.38×10^{-5}	$*4.55 \times 10^{-11}$	SBP
24[f]	*TSPAN18;PRDM11;SYT13*	rs148772934	11:45005681	t	0.986	−0.57	11.66	1.00×10^{-8}	$*1.20 \times 10^{-9}$	DBP
25	*SLC15A3;CD6; LOC105369325*; CD5*	rs11601370	11:60834043	t	0.976	1.34	6.63	0.00867	$*3.01 \times 10^{-9}$	SBP
26[f]	*LOC101928944*	rs74601585	11:80140007	t	0.017	−3.93	−2.58	0.2715	$*8.06 \times 10^{-9}$	SBP
27[f]	*LOC105369408*	rs78103586	11:133893928	a	0.029	−1.65	−5.27	0.00163	$*2.26 \times 10^{-9}$	DBP

(*Continued on next page*)

Table 5. ***Continued***

Locus[a]	Nearest Genes[b]	rsID	Chr:Pos[c]	EA	EAF	Effect[d]		p Value[e]		Trait
						Genetic Main	Interaction	1 df Interaction	2 df Joint	
28[f]	*PLEKHG7;EEA1;LOC643339**	rs61935525	12:93645481	c	0.985	1.26	10.15	3.28×10^{-7}	$*3.28 \times 10^{-11}$	DBP
29[f]	*DICER1;CLMN*	rs187852559	14:95794914	a	0.013	−1.67	−4.93	0.0246	$*8.74 \times 10^{-10}$	DBP
30[f]	*SETD3;CCNK*	rs1257310	14:99810427	a	0.784	1.03	0.98	0.1335	$*1.67 \times 10^{-8}$	SBP
31	*GPR139;GP2;UMOD;PDILT*	rs148753653	16:20230175	a	0.981	5.25	−9.4	$*1.89 \times 10^{-8}$	6.30×10^{-8}	DBP
32[f]	*LOC339166*;WSCD1*	rs138973557	17:5699720	t	0.903	0.36	2.09	2.12×10^{-8}	$*1.81 \times 10^{-10}$	DBP
33[f]	*DYM;LIPG;ACAA2;MYO5B*	rs9965695	18:47261614	t	0.982	0.29	13.32	8.36×10^{-6}	$*1.63 \times 10^{-8}$	SBP
34[f]	*ZNF98**	rs10405764	19:22598479	t	0.017	0.91	−19.1	2.13×10^{-7}	$*4.30 \times 10^{-8}$	SBP
35[f]	*BMP7*	rs115893283	20:55404165	t	0.042	0.9	−9.05	2.53×10^{-8}	$*4.24 \times 10^{-12}$	SBP

Each locus is genome-wide significant ($p < 5 \times 10^{-8}$) in the combined analyses of stages 1 and 2 and had FDR q value < 0.1. Forest plots and LocusZoom plots are shown in Figures S3 and S4, respectively. Abbreviations: BP, blood pressure; SBP, systolic BP; DBP, diastolic BP; EA, effect allele; EAF, effect allele frequency; 2 df joint p, p value of the joint test with 2 degrees of freedom of genetic main and interaction effects; 1 df interaction p, p value of the interaction test with 1 degree of freedom.

[a]Each locus was determined through LD-based clumping, using ± 1 Mb around index variants, followed by LD threshold of $r^2 > 0.1$; ancestry-specific LDs from 1000 Genomes Project were used when clumping within each ancestry and the entire cosmopolitan data were used for trans-ancestry clumping.

[b]Gene names were obtained using variant effect predictor (VEP) from Ensembl. Genes with intragenic index variants are indicated with an asterisk (*).

[c]Positions are based on build 37.

[d]Effect is in mmHg unit.

[e]The most significant p value (between 1 df interaction test and 2 df joint test) is indicated with an asterisk (*).

[f]These loci indicate "completely novel" loci, at least 1 Mb away from any of known BP loci.

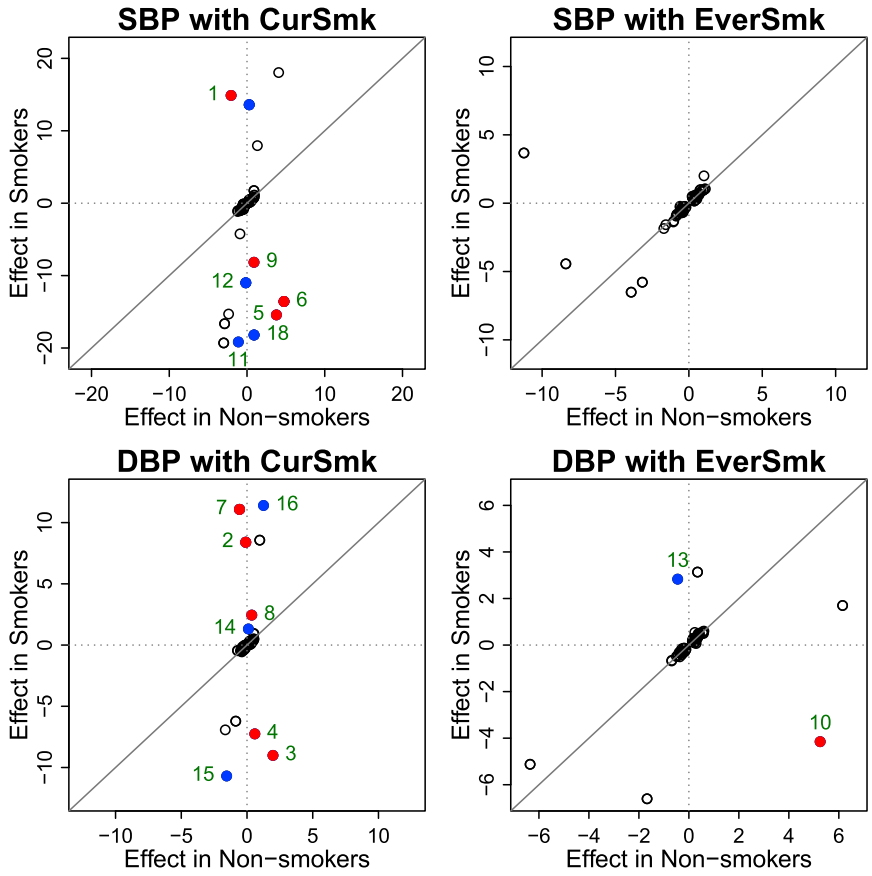


Figure 3. Scatterplots of Smoking-Specific Genetic Effect Sizes for BP Traits at the 15 Newly Identified and 66 Putative Index Variants Listed in Tables 2, 3, 4, and 5

The red points show variants with 1 df interaction $p < 5 \times 10^{-8}$ (1 = rs12135881; 2 = rs115234772; 3 = rs62319742; 4 = rs148387718; 5 = rs74701635; 6 = rs150155092; 7 = rs148772934; 8 = rs138973557; 9 = rs115893283; 10 = rs148753653). The blue points show variants with 1 df interaction $p < 1 \times 10^{-5}$ (11 = rs11809589; 12 = rs10166552; 13 = rs11931572; 14 = rs9348895; 15 = rs76726877; 16 = rs61935525; 17 = rs9965695; 18 = rs10405764).

The majority of the available data on tissue expression are derived from studies with a breadth of tissue types but with small sample sizes that limit the statistical power to detect association. A more in-depth but single-tissue functional annotation, reporting both *cis*- and *trans*-acting elements, was recently performed using microarray-based gene and exon expression levels in whole blood from more than 5,000 individuals of the Framingham Heart Study.[35] In this database, a total of 170 variant-transcript pairs (representing 36 variants) were significant at false discovery rate (FDR) < 0.05 (Table S14). There were 113 pairs for *cis*-eQTL, 3 pairs for *trans*-eQTL, and 54 pairs for long-range *cis*-eQTL where the variant is located more than 1 Mb away from the target transcript on the same chromosome. Among 36 variants, 9 variants were eQTLs for more than 5 gene transcripts. For example, the 4 SNPs with the most significant eQTL evidence were rs2243873 (described in the previous paragraph), rs2071550, rs7823056, and rs13271489 (locus 8 in Table 2 and Figure 2) associated with 29, 12, 11, and 10 transcripts, respectively.

Pathway and Gene Set Enrichment Analysis

In order to distinguish between functional properties of loci with SBP compared to DBP effects, as well as between European-specific and trans-ancestry mechanisms, we conducted gene prioritization, gene set enrichment, and tissue enrichment analyses using DEPICT[43] separately by the four combinations of ancestry (EUR versus trans-ancestry) and BP trait (DBP versus SBP; Material and Methods, Tables S15–S20). DEPICT significantly prioritized genes (FDR < 5%) at 12 European DBP loci, 26 European SBP loci, 34 trans-ancestry DBP loci, and 27 trans-ancestry SBP loci (Tables S15–S19). In 43 cases, the prioritized gene for a specific locus differed from the nearest gene of the lead variant. Our DEPICT gene-set enrichment analyses highlighted a role for the identified variants in the cardiovascular system—predominantly affecting blood vessel biology (FDR < 0.05 for a total of 134 gene-sets across the four analyses, Table S20).

To identify connected gene sets and pathways implicated by our findings, we performed GeneGO analysis and text data mining using Literature Lab.[36] The genes near our findings were enriched by GeneGO disease class "chronic kidney failure" ($p = 9.2 \times 10^{-6}$). These same genes were also included in the much larger network representing the GeneGO disease class "fibrosis" ($p = 3.39 \times 10^{-7}$), suggesting that genetic contribution of chronic kidney disease to BP is likely mediated by fibrosis. With Literature Lab, for the "diseases" medical subject heading (MeSH), hypertension was strongly enriched ($p = 0.0011$), with contributions from *ACE* (93.4%), *MTHFR* (2.12%), *ATP2B1* (1.18%), *NPPB* (0.54%), *SH2B3* (0.43%), and *SLC4A7* (0.13%). For the "physiology" MeSH, blood pressure and cardiovascular physiological phenomena were enriched. Blood pressure ($p = 0.0026$) had contributions from *ACE* (96.77%), *ATP2B1* (1.16%), *NPPB* (0.6%), *MTHFR* (0.46%), *SH2B3* (0.46%), and *FTO* (0.3%). Cardiovascular physiological phenomena ($p = 0.0056$) had contributions from *ACE* (97.89%), *NPPB* (1%), *ATP2B1* (0.37%), *MTHFR* (0.2%), *SH2B3* (0.16%), *TNFSF12* (0.09%), and *AP5B1* (0.05%).

Associations of BP Loci with Cardiometabolic Traits

To test association of all 81 newly identified BP-associated index variants with other cardiometabolic traits, we

obtained lookup results for coronary artery disease (CAD), stroke, and other cardiometabolic traits related to adiposity, diabetes, and renal function (Tables S21–S27). We found that several of our newly identified index variants corroborate those previously associated with these cardiometabolic traits. To quantify this, we counted the number of variants that show association with p value < 0.05 (highlighted in red). In the vast majority of cases (39 out of 47, $P_{Binomial} = 2.8 \times 10^{-6}$), the observed count is higher than that expected by chance alone (Table S27). For example, we observed 9 and 14 such associations with CAD and myocardial infarction, respectively, where the expected count is 2.6 for both traits. This is consistent with the known association of increased BP with CAD mortality, independent of other risk factors.[44] Likewise, overlapping signals with other cardiometabolic traits, including those related to adiposity, diabetes, and renal function, support the notion that these traits share a common pathophysiology. For many of the obesity-related trait associations found in the GIANT Consortium, the genetic effect was influenced by adjustment and/or stratification by smoking status[45] (Table S26).

We also found corroborating evidence for some well-known loci associated with the renin-angiotensin-aldosterone system (RAAS), including *NPPA*, *NPPB*, and *SLC17A1-4* (Tables 2, 3, and 4).[4] Variants in and near these loci have also been associated with CAD-related traits (*NPPA*/*NPPB*; Table S21), stroke (*NPPA*/*NPPB* and *SLC17A1-4*; Table S22), obesity-related traits (*NPPA*/*NPPB* and *SLC17A1-4*; Table S23), and diabetes-related traits (*SLC17A1-4*; Table S24) The confluence of these data provide further evidence of the biologic relevance of these loci to BP regulation and the shared pathophysiology among cardiometabolic traits.

Biological Relevance of Newly Identified Variants Associated with BP

Ciliopathies

Cilia are cellular protuberances found in several tissues including the kidney and brain that serve several purposes including cellular structure, growth, mobility, secretion, and environmental response. New BP candidate genes *SDCCAG8* (locus zoom plot in Figure 2), *RPGRIP1L*, and *TMEM231* encode products that play critical roles in the structure and function of primary cilia including microtubules, basal bodies, and centrosomes. Mutations in these genes can lead to nephronophthisis-related ciliopathy, a monogenic cause of end-stage renal disease. *DPYSL2*, which encodes a microtubule assembly protein, has also been implicated in polycystic kidney disease.[46] Cilia also contain actin fibers with motor proteins (dynein and kinesin) responsible for the transport of mitochondria and other cargo. *DYNC2LI1* is another dynein-associated protein associated with BP; dynein proteins co-localize in the kidney with the water channel aquaporin-2.[47]

Telomere Maintenance

Since telomere length shortens with successive cell divisions, it has been proposed as a reflection of biologic age.[48] Several genes with significant association with BP have roles in telomere maintenance including *TNKS*, *PINX1*, *AKTIP* (Tables 2, 3, and 4), and *TERF2IP*. *TNKS*, which is in a locus previously associated with stroke-, obesity-, and diabetes-related traits in other studies (Tables S22–S24), plays a role in the insulin-stimulated translocation of GLUT4 (glucose transporter) to the plasma membrane[49] and has additionally been associated with cardiovascular disease (CVD) risk and the inflammatory biomarker, C-reactive protein.[50] *PINX1* has been previously associated with CVD,[51] carotid artery intima-media thickness,[52] and serum triglyceride levels,[53] and has also been associated with obesity- and diabetes-related traits (Tables S23 and S24). *AKTIP* has been previously associated with stroke-related traits in other studies (Table S22). Of note, the association at *TNKS*, *PINX1*, and *AKTIP* with multiple adiposity traits in the GIANT Consortium were strengthened by adjustment for smoking status (Table S26). *TERF2IP* has also been associated with stroke risk[50] and coronary artery disease traits (Tables S21 and S22).

Central Dopaminergic Signaling

Dopaminergic signaling in the kidney is known to modulate the secretion of renin[54] and other key regulators of salt-water balance.[55] There is evidence that central dopamine signaling also modulates BP via mechanisms that are independent of changes in sodium excretion.[56] Early stages of Parkinson disease, a neurodegenerative disorder characterized by the loss of dopamine-secreting neurons, is characterized by autonomic dysfunction and BP dysregulation.[57] In the current study, genes involved in central dopamine signaling were associated with BP, including *MSRA* and *EBF2*, which promote the survival and development of dopaminergic neurons, and *GPR19*, a G-protein coupled receptor for the dopamine D_2 receptor. *MSRA* has been previously associated with body mass index after adjustment with smoking status in the GIANT Consortium (Table S26) and *GPR19* with renal function (Table S25) in the COGENT-Kidney Consortium.

Modulators of Vascular Structure and Function

CDKN1B, *BCAR1-CFDP1*, *PXDN*, and *EEA1* are involved in pathways that contribute to angiotensin II-induced vascular hypertrophy. Notably, the association of *PXDN* and *EEA1* with BP is limited to AFR. *CDKN1B* has been previously associated with renal function (Table S25). *BCAR1-CFDP1* has furthermore been identified as a genome-wide significant locus for carotid artery intima-media thickness and coronary artery disease risk (also Table S21);[58] a potential causal variant in a *BCAR1* regulatory domain has been identified.[59] *KCNG3* and *KCNE4* are subunit modifiers of voltage-gated potassium channels expressed in vascular smooth muscle cells; activation of these channels leads to vasodilation. *AVPR1A*, which was associated with BP in AFR only, is a receptor for the

vasoconstrictor vasopressin; murine knock-out models are hypotensive with impaired baroreceptor reflexes.[60]

Discussion

This is a large-scale multi-ancestry study to systematically use GxE interactions for identifying trait loci and for evaluating the role of GxE interactions in cardiovascular traits. In stage 1, we performed a genome-wide analysis of gene-smoking interactions in 129,913 individuals across four ancestry groups using 1000 Genomes-imputed data, with follow-up analysis in stage 2 of a small set of promising variants in 480,178 additional individuals across five ancestry groups. We identified 40 known BP loci at genome-wide significance level ($p < 5 \times 10^{-8}$) in stage 1 as well as 15 novel loci that are genome-wide significant in stage 1 and replicated in stage 2 using Bonferroni correction. A combined meta-analysis of stages 1 and 2 results yielded 16 additional known BP loci and 66 additional genome-wide significant loci ($p < 5 \times 10^{-8}$); 13, 35, and 18 loci were identified in European, African, and trans-ancestry, respectively. These 66 additional loci were validated with low false discovery rate (FDR q value < 0.1) (e.g., see Nelson et al.[61]).

Identification of novel loci in this GxE analysis demonstrates the importance of incorporating environmental exposures in association discovery. Our newly identified loci including interactions with smoking collectively explained up to 1.7% additional variance in BP (beyond that explained by known BP variants) in several European cohorts. Furthermore, it may be particularly striking that our analyses also identified *VAMP2*, a component of the renin-angiotensin-aldosterone system (RAAS), as a likely mediator of hypertension. *VAMP2* modulates cAMP-stimulated renin release by renal juxtaglomerular cells[62] but has not been previously identified, even though other components of RAAS including *NPPA*, *NPPB*, and *SLC17A1-4* have been found in previous GWASs and, indeed, among the 56 known BP loci identified in our study.[4,63–65]

Several of our newly identified BP loci show evidence for shared pathophysiology with cardiometabolic traits. This is encouraging as hypertension is a frequent comorbidity of a variety of cardiometabolic traits, including dyslipidemia, type 2 diabetes, obesity, and other disorders of substrate metabolism and storage. *XKR6-MIR598* and *MFHAS1* have been associated with serum triglyceride levels.[66] *LRP6*[67,68] and *PPP1R3B*[69] have been associated with serum low-density lipoprotein levels and the metabolic syndrome. *MSRA*[70] and *SERTAD2*[71] (associated in AFR) have been associated with obesity-related traits and adipocyte function, and *PPP1R3B* has been associated with steatohepatitis.[72] We also identified the well-known obesity/diabetes locus *FTO*[73,74] as a newly identified BP locus (Figure 2). In addition to a recent discovery of the effect of an *FTO* variant on *IRX3* and *IRX5*,[75] variants in intron 1 of *FTO* have been identified that regulate the expression of nearby *RPGRIP1L*,[74] shown to modulate leptin receptor trafficking and signaling in the hypothalamus.[76] Variants in and near *XKR6-MIR598*, *MFHAS1*, *MSRA*, and *FTO* have been associated with obesity- and diabetes-related traits in other studies (Tables S23 and S24). Among other variants in genes related to cardiometabolic traits, *VAMP2* plays a role in the trafficking of the GLUT4 glucose receptor to the adipocyte plasma membrane.[77] Finally, we identified a SNP (in AFR) in *FABP3*, a gene known to regulate mitochondrial β-oxidation.[78] Studies have shown that serum *FABP3* transcript and protein levels are elevated in animal models and humans with hypertension compared with normotensive controls.[79,80] Consistent with a recent paper,[6] our findings provide additional BP variants overlapping with metabolic trait loci.

Some of the newly identified BP loci have been previously reported as suggestive (but not genome-wide significant) for smoking and other addiction traits. Among our newly identified loci, *FTO*, *DPYSL2-ADRA1A*, *AJAP1*, and *SERINC2* have shown suggestive evidence of association with smoking-related traits,[81,82] illicit drug use,[83] and alcohol consumption and dependence.[84,85] In addition, dopaminergic signaling has been implicated in addictive behaviors.[86] Moreover, located in an intron of *TNFSF12* (tumor necrosis factor superfamily member), our newly identified variant rs9899183 has many compelling regulatory features supporting its candidacy (Table S12); it resides in a region characterized by promoter histone marks in 23 tissues, in enhancer histone marks in 7 tissues, and by DNase marks in 12 tissues. This variant is also identified as an eQTL for genes *TNFSF12*, *CHRNB1*, and *SAT2*; *CHRNB1* (1 nicotinic acetylcholine receptor subunit) may also contribute to nicotine dependence.[87]

BP regulation critically involves both central and peripheral regulation via neuroendocrine and hormonal regulation in a complex integrated system that includes the brain, kidneys, adrenal glands, and vasculature. In addition to validating loci known for their involvement in the RAAS system, natriuretic peptide signaling, solute channels, and adrenergic and cholinergic receptor signaling (among others), we identified variants in or near new biological candidates for BP regulation. For example, several of our newly identified loci identified genes that have been previously implicated in monogenic causes of ciliopathy (nephronophthisis-related ciliopathy), a cause of end-stage renal disease in children and young adults.[88,89] This condition is a genetically heterogeneous autosomal-recessive disease, and heterozygote siblings and other adults with incompletely penetrant versions of this disease may have variable degrees of hypertension, renal insufficiency, obesity, and diabetes.[90] Newly identified loci also include genes involved in dopaminergic signaling which may act both centrally and in the kidney to modulate BP regulation. Still other newly identified loci reside in or near genes involved in telomere maintenance.

Of the 81 newly identified loci, 10 show genome-wide significant interactions although none were replicated in stage 2. Nine were identified with current smoking status. The ever smoking status is more heterogeneous since the effect of (former) smoking on BP decays over time from cessation.[91] It is therefore not surprising that the analyses with the more homogeneous current smoking (CurSmk) status yielded larger (and more robust) effects on BP than did analyses using ever smoker (EverSmk) status. Although the joint 2 df test succeeded in identifying 71 of the 81 newly identified loci, the precise role of interaction is unclear. It is sobering to note that, although gene-smoking interactions may have helped identify a reasonably large number of the newly identified loci, the sample size we used here for genome-wide analysis in stage 1 appears inadequate for identifying a large number of interaction effects (should they exist) through the 1 df interaction test alone. This may be because, if the pathobiology of BP involves large numbers of interactions, the majority of the interaction effects are likely (relatively) small enough whose identification requires the 2 df joint test and/or require much larger sample sizes for identifying them through the 1 df interaction test. Moreover, smoking is only one of many lifestyle attributes that may have interaction effects on BP.[12] It is possible that some interactions we report here are driven by other lifestyle factors that may be correlated with smoking. A follow-up study (such as Young et al.[92] and Tyrrell et al.[93]) that jointly examines multiple lifestyle factors can shed light on further understanding of interaction effects on BP.

Several large consortia-based BP GWAS papers have been published in recent years, dramatically increasing the number of BP loci. We treated 158 as known BP loci, which included the 71 loci that were reported by three recent papers.[5–7] Of the 56 known BP loci we identified, 8 overlap with these newly identified 71 loci. Hoffmann et al.[94] reported 75 novel loci (and 241 additional loci not validated) based on >300,000 individuals. The use of repeated measurements, beside the large sample size, appears to be responsible for the large number of novel loci discovered. Their study demonstrates the power of large sample sizes and repeated measurements. Warren et al.[95] reported 107 validated loci. As shown in Table S28 in detail, nine of our newly identified loci include variants reported by these two papers.[94,95] Based on African ancestry, Liang et al. reported three validated BP loci,[96] one of which overlaps with our newly identified loci.

35 loci were identified in African ancestry meta-analyses. As previous discoveries of BP loci were mostly in European ancestry, some using very large sample sizes, it may be harder to detect newly identified signals in European ancestry in our study. There are also more opportunities to identify lower frequency variants in African ancestry meta-analysis because there are more of these variants in this genetically more diverse population. However, because of the highly limited sample sizes available for African ancestry in stage 2, genome-wide significant loci in stage 1 African ancestry could not be formally replicated in stage 2. Nevertheless, there is evidence supporting the validity of many of the African-specific newly identified loci: African-specific QQ plots were very similar with and without the known BP loci (Figures S10 and S12). Genomic control values are all close to 1, and the top signals are away from the expected null line in the QQ plots, suggesting that these may be real associations. Forest plots at the African-specific loci (Figure S13) were not heterogeneous across cohorts. For most loci, there exists at least one non-African ancestry showing effects in the same direction as those in African ancestry. They may also relate at least in part to unique smoking behaviors or BP regulation or both in African ancestry. However, these African-specific loci require further validation.

There are several limitations in this large-scale multi-ancestry genome-wide investigation incorporating gene-smoking interactions. First, main effect only analysis without regard to smoking was not performed, and this limits our ability to resolve whether any of our loci newly identified through the 2 df joint test could be found without smoking or gene-smoking interaction in the model. Second, although the strategy of clumping with a stringent LD threshold ($r^2 > 0.1$) in addition to large physical distance threshold (± 1 Mb) is reasonable for inferring independent loci, conditional analysis of summary statistics from interaction analysis (similar to GCTA) would be more rigorous; however, such methods do not exist currently. Third, the relatively smaller stage 2 sample sizes available in African and Hispanic ancestries limit our ability to formally replicate the loci that were newly identified in stage 1 in those ancestries (including the 10 interactions). Fourth, power for discovery using interactions may be limited even in this reasonably large sample size. Fifth, if there is a G-C correlation, a potential confounding of GxE with interaction between covariate and smoking exposure (CxE) may exist, which can inflate type I error of the GxE interaction test;[97,98] using a stratified model may help overcome such confounding. Sixth, our use of the fixed effect meta-analysis for trans-ancestry analysis may have limited the power in the presence of heterogeneous effects across ancestries; however, specialized trans-ancestry methods for GxE interactions do not exist. Seventh, subjects were grouped into each ancestry based on self-reported information instead of genetically computed ancestry. Finally, the use of multiple hypothesis tests, multiple phenotypes and exposures, and multiple ancestries may contribute to inflation at some level. Striking a balance between false positives and false negatives, especially in the context of interactions, remains a challenge.

In summary, our study identified a total of 137 genome-wide significant loci; 56 known loci, 15 new loci identified in stage 1 and formally replicated in stage 2, and 66 additional BP loci identified through the combined analysis of stages 1 and 2 and validated through low FDR. Our ability to identify this many loci is likely due to four

factors: focus on gene-smoking interactions, consideration of multiple ancestries, the large aggregate sample sizes available, and the densely imputed data using the recent 1000 Genomes Project reference panel in stage 1 analysis. The 10 newly identified loci with significant interactions showed larger effects on BP in smokers. 35 loci were identified only in African ancestry, highlighting the importance of pursuing genetic studies in diverse populations. In addition to evidence for shared pathophysiology with cardiometabolic traits, smoking, and other addiction traits, our results provide compelling evidence for biological candidates for BP regulation such as modulators of vascular structure and function, ciliopathies, telomere maintenance, and central dopaminergic signaling. Our findings demonstrate how the interplay between genes and environment can help identify loci, open up new avenues for investigation about BP homeostasis, and highlight the promise of gene-lifestyle interactions for more in-depth genetic and environmental dissection of BP and other complex traits.

Supplemental Data

Supplemental Data include Supplemental Notes, 17 figures, and 28 tables and can be found with this article online at https://doi.org/10.1016/j.ajhg.2018.01.015.

Conflicts of Interest

The authors declare no competing financial interests except for the following. B.M.P. serves on the DSMB of a clinical trial funded by the manufacturer (Zoll LifeCor) and on the Steering Committee of the Yale Open Data Access Project funded by Johnson & Johnson; O.H.F. received grants from Metagenics (on women's health and epigenetics) and from Nestle (on child health); L.J.B. is listed as an inventor on Issued U.S. Patent 8,080,371,"Markers for Addiction" covering the use of certain SNPs in determining the diagnosis, prognosis, and treatment of addiction; P.S. has received research awards from Pfizer Inc., is a consultant for Mundipharma Co. (Cambridge, UK), is a patent holder with Biocompatibles UK Ltd. (Franham, Surrey, UK) (Title: Treatment of eye diseases using encapsulated cells encoding and secreting neuroprotective factor and/or anti-angiogenic factor; Patent number: 20120263794), and has a patent application with University of Heidelberg (Heidelberg, Germany) (Title: Agents for use in the therapeutic or prophylactic treatment of myopia or hyperopia; Europäische Patentanmeldung 15 000 771.4); P.W.F. has been a paid consultant for Eli Lilly and Sanofi Aventis and has received research support from several pharmaceutical companies as part of a European Union Innovative Medicines Initiative (IMI) project; M.A.N.'s participation is supported by a consulting contract between Data Tecnica Internation and the National Institute on Aging (NIH, Bethesda, MD, USA), and he also consults for Illumina, Inc., the Michael J. Fox Foundation, and University of California Healthcare among others; and M.J.C. is Chief Scientist for Genomics England, a UK government company.

Acknowledgments

The various Gene-Lifestyle Interaction projects, including this one, are largely supported by a grant from the U.S. National Heart, Lung, and Blood Institute (NHLBI), the National Institutes of Health, R01HL118305. A Career Development Award (K25HL121091), also from the NHLBI, enabled Y.J.S. to lead this project. Full set of study-specific funding sources and acknowledgments appear in the Supplemental Note.

Received: October 10, 2017
Accepted: January 18, 2018
Published: February 15, 2018

Web Resources

dbSNP, https://www.ncbi.nlm.nih.gov/projects/SNP/
DEPICT, https://data.broadinstitute.org/mpg/depict
GeneGo, https://clarivate.com/product-category/life-sciences/
GSEA, http://software.broadinstitute.org/gsea/msigdb/annotate.jsp
GTEx Portal, https://www.gtexportal.org/home/
HaploReg, http://www.broadinstitute.org/mammals/haploreg/haploreg.php
Literature Lab, http://www.acumenta.com
LocusZoom, http://locuszoom.sph.umich.edu/locuszoom/
METAL, http://genome.sph.umich.edu/wiki/METAL_Documentation
National Human Genome Research Institute (NHGRI) GWAS catalog, https://www.genome.gov/gwastudies/
NCBI Gene, http://www.ncbi.nlm.nih.gov/gene
RegulomeDB, http://RegulomeDB.org/
Roadmap, http://www.roadmapepigenomics.org/

References

1. Levy, D., Ehret, G.B., Rice, K., Verwoert, G.C., Launer, L.J., Dehghan, A., Glazer, N.L., Morrison, A.C., Johnson, A.D., Aspelund, T., et al. (2009). Genome-wide association study of blood pressure and hypertension. Nat. Genet. *41*, 677–687.
2. Manolio, T.A., Collins, F.S., Cox, N.J., Goldstein, D.B., Hindorff, L.A., Hunter, D.J., McCarthy, M.I., Ramos, E.M., Cardon, L.R., Chakravarti, A., et al. (2009). Finding the missing heritability of complex diseases. Nature *461*, 747–753.
3. Newton-Cheh, C., Johnson, T., Gateva, V., Tobin, M.D., Bochud, M., Coin, L., Najjar, S.S., Zhao, J.H., Heath, S.C., Eyheramendy, S., et al.; Wellcome Trust Case Control Consortium (2009). Genome-wide association study identifies eight loci associated with blood pressure. Nat. Genet. *41*, 666–676.
4. Ehret, G.B., Munroe, P.B., Rice, K.M., Bochud, M., Johnson, A.D., Chasman, D.I., Smith, A.V., Tobin, M.D., Verwoert, G.C., Hwang, S.J., et al.; International Consortium for Blood Pressure Genome-Wide Association Studies; CARDIoGRAM consortium; CKDGen Consortium; KidneyGen Consortium; EchoGen consortium; and CHARGE-HF consortium (2011). Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. Nature *478*, 103–109.
5. Ehret, G.B., Ferreira, T., Chasman, D.I., Jackson, A.U., Schmidt, E.M., Johnson, T., Thorleifsson, G., Luan, J., Donnelly, L.A., Kanoni, S., et al.; CHARGE-EchoGen consortium; CHARGE-HF consortium; and Wellcome Trust Case Control Consortium (2016). The genetics of blood pressure regulation and its target organs from association studies in 342,415 individuals. Nat. Genet. *48*, 1171–1184.

6. Liu, C., Kraja, A.T., Smith, J.A., Brody, J.A., Franceschini, N., Bis, J.C., Rice, K., Morrison, A.C., Lu, Y., Weiss, S., et al.; CHD Exome+ Consortium; ExomeBP Consortium; GoT2DGenes Consortium; T2D-GENES Consortium; Myocardial Infarction Genetics and CARDIoGRAM Exome Consortia; and CKDGen Consortium (2016). Meta-analysis identifies common and rare variants influencing blood pressure and overlapping with metabolic trait loci. Nat. Genet. *48*, 1162–1170.
7. Surendran, P., Drenos, F., Young, R., Warren, H., Cook, J.P., Manning, A.K., Grarup, N., Sim, X., Barnes, D.R., Witkowska, K., et al.; CHARGE-Heart Failure Consortium; EchoGen Consortium; METASTROKE Consortium; GIANT Consortium; EPIC-InterAct Consortium; Lifelines Cohort Study; Wellcome Trust Case Control Consortium; Understanding Society Scientific Group; EPIC-CVD Consortium; CHARGE+ Exome Chip Blood Pressure Consortium; T2D-GENES Consortium; GoT2DGenes Consortium; ExomeBP Consortium; and CHD Exome+ Consortium (2016). Trans-ancestry meta-analyses identify rare and common variants associated with blood pressure and hypertension. Nat. Genet. *48*, 1151–1161.
8. Franceschini, N., Fox, E., Zhang, Z., Edwards, T.L., Nalls, M.A., Sung, Y.J., Tayo, B.O., Sun, Y.V., Gottesman, O., Adeyemo, A., et al.; Asian Genetic Epidemiology Network Consortium (2013). Genome-wide association analysis of blood-pressure traits in African-ancestry individuals reveals common associated genes in African and non-African populations. Am. J. Hum. Genet. *93*, 545–554.
9. Zhu, X., Feng, T., Tayo, B.O., Liang, J., Young, J.H., Franceschini, N., Smith, J.A., Yanek, L.R., Sun, Y.V., Edwards, T.L., et al.; COGENT BP Consortium (2015). Meta-analysis of correlated traits via summary statistics from GWASs with an application in hypertension. Am. J. Hum. Genet. *96*, 21–36.
10. Manning, A.K., Hivert, M.F., Scott, R.A., Grimsby, J.L., Bouatia-Naji, N., Chen, H., Rybin, D., Liu, C.T., Bielak, L.F., Prokopenko, I., et al.; DIAbetes Genetics Replication And Meta-analysis (DIAGRAM) Consortium; and Multiple Tissue Human Expression Resource (MUTHER) Consortium (2012). A genome-wide approach accounting for body mass index identifies genetic variants influencing fasting glycemic traits and insulin resistance. Nat. Genet. *44*, 659–669.
11. Hunter, D.J. (2005). Gene-environment interactions in human diseases. Nat. Rev. Genet. *6*, 287–298.
12. Go, A.S., Mozaffarian, D., Roger, V.L., Benjamin, E.J., Berry, J.D., Blaha, M.J., Dai, S., Ford, E.S., Fox, C.S., Franco, S., et al.; American Heart Association Statistics Committee and Stroke Statistics Subcommittee (2014). Heart disease and stroke statistics–2014 update: a report from the American Heart Association. Circulation *129*, e28–e292.
13. Mann, S.J., James, G.D., Wang, R.S., and Pickering, T.G. (1991). Elevation of ambulatory systolic blood pressure in hypertensive smokers. A case-control study. JAMA *265*, 2226–2228.
14. Primatesta, P., Falaschetti, E., Gupta, S., Marmot, M.G., and Poulter, N.R. (2001). Association between smoking and blood pressure: evidence from the health survey for England. Hypertension *37*, 187–193.
15. Green, M.S., Jucha, E., and Luz, Y. (1986). Blood pressure in smokers and nonsmokers: epidemiologic findings. Am. Heart J. *111*, 932–940.
16. Rao, D.C., Sung, Y.J., Winkler, T.W., Schwander, K., Borecki, I., Cupples, L.A., Gauderman, W.J., Rice, K., Munroe, P.B., Psaty, B.M.; and CHARGE Gene-Lifestyle Interactions Working Group* (2017). Multi-ancestry study of gene-lifestyle interactions for cardiovascular traits in 610,475 individuals from 124 cohorts: Design and rationale. Circ Cardiovasc Genet *10*, e001649.
17. Kirk, E.P. (2017). Genes, environment, and the heart: putting the pieces together. Circ Cardiovasc Genet *10*, 10.
18. Abecasis, G.R., Auton, A., Brooks, L.D., DePristo, M.A., Durbin, R.M., Handsaker, R.E., Kang, H.M., Marth, G.T., McVean, G.A.; and 1000 Genomes Project Consortium (2012). An integrated map of genetic variation from 1,092 human genomes. Nature *491*, 56–65.
19. Sung, Y.J., Winkler, T.W., Manning, A.K., Aschard, H., Gudnason, V., Harris, T.B., Smith, A.V., Boerwinkle, E., Brown, M.R., Morrison, A.C., et al. (2016). An empirical comparison of joint and stratified frameworks for studying G × E interactions: systolic blood pressure and smoking in the CHARGE Gene-Lifestyle Interactions Working Group. Genet. Epidemiol. *40*, 404–415.
20. Tchetgen Tchetgen, E.J., and Kraft, P. (2011). On the robustness of tests of genetic associations incorporating gene-environment interaction when the environmental exposure is misspecified. Epidemiology *22*, 257–261.
21. Voorman, A., Lumley, T., McKnight, B., and Rice, K. (2011). Behavior of QQ-plots and genomic control in studies of gene-environment interaction. PLoS ONE *6*, e19416.
22. Zeileis, A. (2006). Object-oriented computation of sandwich estimators. J. Stat. Softw. *16*, 1–16.
23. Aulchenko, Y.S., Struchalin, M.V., and van Duijn, C.M. (2010). ProbABEL package for genome-wide association analysis of imputed data. BMC Bioinformatics *11*, 134.
24. Winkler, T.W., Day, F.R., Croteau-Chonka, D.C., Wood, A.R., Locke, A.E., Mägi, R., Ferreira, T., Fall, T., Graff, M., Justice, A.E., et al.; Genetic Investigation of Anthropometric Traits (GIANT) Consortium (2014). Quality control and conduct of genome-wide association meta-analyses. Nat. Protoc. *9*, 1192–1212.
25. Marchini, J., and Howie, B. (2010). Genotype imputation for genome-wide association studies. Nat. Rev. Genet. *11*, 499–511.
26. Manning, A.K., LaValley, M., Liu, C.T., Rice, K., An, P., Liu, Y., Miljkovic, I., Rasmussen-Torvik, L., Harris, T.B., Province, M.A., et al. (2011). Meta-analysis of gene-environment interaction: joint estimation of SNP and SNP × environment regression coefficients. Genet. Epidemiol. *35*, 11–18.
27. Willer, C.J., Li, Y., and Abecasis, G.R. (2010). METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics *26*, 2190–2191.
28. Randall, J.C., Winkler, T.W., Kutalik, Z., Berndt, S.I., Jackson, A.U., Monda, K.L., Kilpeläinen, T.O., Esko, T., Mägi, R., Li, S., et al.; DIAGRAM Consortium; and MAGIC Investigators (2013). Sex-stratified genome-wide association studies including 270,000 individuals show sexual dimorphism in genetic loci for anthropometric traits. PLoS Genet. *9*, e1003500.
29. Aschard, H., Hancock, D.B., London, S.J., and Kraft, P. (2010). Genome-wide meta-analysis of joint tests for genetic and gene-environment interaction effects. Hum. Hered. *70*, 292–300.
30. Winkler, T.W., Kutalik, Z., Gorski, M., Lottaz, C., Kronenberg, F., and Heid, I.M. (2015). EasyStrata: evaluation and visualization of stratified genome-wide association meta-analysis data. Bioinformatics *31*, 259–261.

31. Devlin, B., and Roeder, K. (1999). Genomic control for association studies. Biometrics *55*, 997–1004.
32. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., de Bakker, P.I., Daly, M.J., and Sham, P.C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. *81*, 559–575.
33. Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate - a practical and powerful approach to multiple testing. J Roy Stat Soc B Met *57*, 289–300.
34. GTEx Consortium (2015). Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. Science *348*, 648–660.
35. Joehanes, R., Zhang, X., Huan, T., Yao, C., Ying, S.X., Nguyen, Q.T., Demirkale, C.Y., Feolo, M.L., Sharopova, N.R., Sturcke, A., et al. (2017). Integrated genome-wide analysis of expression quantitative trait loci aids interpretation of genomic association studies. Genome Biol. *18*, 16.
36. Febbo, P.G., Mulligan, M.G., Slonina, D.A., Stegmaier, K., Di Vizio, D., Martinez, P.R., Loda, M., and Taylor, S.C. (2007). Literature Lab: a method of automated literature interrogation to infer biology from microarray analysis. BMC Genomics *8*, 461.
37. Skol, A.D., Scott, L.J., Abecasis, G.R., and Boehnke, M. (2006). Joint analysis is more efficient than replication-based analysis for two-stage genome-wide association studies. Nat. Genet. *38*, 209–213.
38. Ward, L.D., and Kellis, M. (2012). HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. Nucleic Acids Res. *40*, D930–D934.
39. Boyle, A.P., Hong, E.L., Hariharan, M., Cheng, Y., Schaub, M.A., Kasowski, M., Karczewski, K.J., Park, J., Hitz, B.C., Weng, S., et al. (2012). Annotation of functional variation in personal genomes using RegulomeDB. Genome Res. *22*, 1790–1797.
40. Davydov, E.V., Goode, D.L., Sirota, M., Cooper, G.M., Sidow, A., and Batzoglou, S. (2010). Identifying a high fraction of the human genome to be under selective constraint using GERP++. PLoS Comput. Biol. *6*, e1001025.
41. Garber, M., Guttman, M., Clamp, M., Zody, M.C., Friedman, N., and Xie, X. (2009). Identifying novel constrained elements by exploiting biased substitution patterns. Bioinformatics *25*, i54–i62.
42. Pickrell, J.K., Marioni, J.C., Pai, A.A., Degner, J.F., Engelhardt, B.E., Nkadori, E., Veyrieras, J.B., Stephens, M., Gilad, Y., and Pritchard, J.K. (2010). Understanding mechanisms underlying human gene expression variation with RNA sequencing. Nature *464*, 768–772.
43. Pers, T.H., Karjalainen, J.M., Chan, Y., Westra, H.J., Wood, A.R., Yang, J., Lui, J.C., Vedantam, S., Gustafsson, S., Esko, T., et al.; Genetic Investigation of ANthropometric Traits (GIANT) Consortium (2015). Biological interpretation of genome-wide association studies using predicted gene functions. Nat. Commun. *6*, 5890.
44. Lewington, S., Clarke, R., Qizilbash, N., Peto, R., Collins, R.; and Prospective Studies Collaboration (2002). Age-specific relevance of usual blood pressure to vascular mortality: a meta-analysis of individual data for one million adults in 61 prospective studies. Lancet *360*, 1903–1913.
45. Justice, A.E., Winkler, T.W., Feitosa, M.F., Graff, M., Fisher, V.A., Young, K., Barata, L., Deng, X., Czajkowski, J., Hadley, D., et al. (2017). Genome-wide meta-analysis of 241,258 adults accounting for smoking behaviour identifies novel loci for obesity traits. Nat. Commun. *8*, 14977.
46. Husson, H., Moreno, S., Smith, L.A., Smith, M.M., Russo, R.J., Pitstick, R., Sergeev, M., Ledbetter, S.R., Bukanov, N.O., Lane, M., et al. (2016). Reduction of ciliary length through pharmacologic or genetic inhibition of CDK5 attenuates polycystic kidney disease in a model of nephronophthisis. Hum. Mol. Genet. *25*, 2245–2255.
47. Marples, D., Schroer, T.A., Ahrens, N., Taylor, A., Knepper, M.A., and Nielsen, S. (1998). Dynein and dynactin colocalize with AQP2 water channels in intracellular vesicles from kidney collecting duct. Am. J. Physiol. *274*, F384–F394.
48. Butt, H.Z., Atturu, G., London, N.J., Sayers, R.D., and Bown, M.J. (2010). Telomere length dynamics in vascular disease: a review. Eur. J. Vasc. Endovasc. Surg. *40*, 17–26.
49. Guo, H.L., Zhang, C., Liu, Q., Li, Q., Lian, G., Wu, D., Li, X., Zhang, W., Shen, Y., Ye, Z., et al. (2012). The Axin/TNKS complex interacts with KIF3A and is required for insulin-stimulated GLUT4 translocation. Cell Res. *22*, 1246–1257.
50. Zee, R.Y., Ridker, P.M., and Chasman, D.I. (2011). Genetic variants in eleven telomere-associated genes and the risk of incident cardio/cerebrovascular disease: The Women's Genome Health Study. Clin. Chim. Acta *412*, 199–202.
51. Hemerich, D., van der Laan, S.W., Tragante, V., den Ruijter, H.M., de Borst, G.J., Pasterkamp, G., de Bakker, P.I., and Asselbergs, F.W. (2015). Impact of carotid atherosclerosis loci on cardiovascular events. Atherosclerosis *243*, 466–468.
52. Bis, J.C., Kavousi, M., Franceschini, N., Isaacs, A., Abecasis, G.R., Schminke, U., Post, W.S., Smith, A.V., Cupples, L.A., Markus, H.S., et al.; CARDIoGRAM Consortium (2011). Meta-analysis of genome-wide association studies from the CHARGE consortium identifies common variants associated with carotid intima media thickness and plaque. Nat. Genet. *43*, 940–947.
53. Teslovich, T.M., Musunuru, K., Smith, A.V., Edmondson, A.C., Stylianou, I.M., Koseki, M., Pirruccello, J.P., Ripatti, S., Chasman, D.I., Willer, C.J., et al. (2010). Biological, clinical and population relevance of 95 loci for blood lipids. Nature *466*, 707–713.
54. Imbs, J.L., Schmidt, M., and Schwartz, J. (1975). Effect of dopamine on renin secretion in the anesthetized dog. Eur. J. Pharmacol. *33*, 151–157.
55. Boone, M., and Deen, P.M. (2008). Physiology and pathophysiology of the vasopressin-regulated renal water reabsorption. Pflugers Arch. *456*, 1005–1024.
56. Li, X.X., Bek, M., Asico, L.D., Yang, Z., Grandy, D.K., Goldstein, D.S., Rubinstein, M., Eisner, G.M., and Jose, P.A. (2001). Adrenergic and endothelin B receptor-dependent hypertension in dopamine receptor type-2 knockout mice. Hypertension *38*, 303–308.
57. Sakata, M., Sei, H., Toida, K., Fujihara, H., Urushihara, R., and Morita, Y. (2002). Mesolimbic dopaminergic system is involved in diurnal blood pressure regulation. Brain Res. *928*, 194–201.
58. Gertow, K., Sennblad, B., Strawbridge, R.J., Ohrvik, J., Zabaneh, D., Shah, S., Veglia, F., Fava, C., Kavousi, M., McLachlan, S., et al. (2012). Identification of the BCAR1-CFDP1-TMEM170A locus as a determinant of carotid intima-media thickness and coronary artery disease risk. Circ Cardiovasc Genet *5*, 656–665.

59. Boardman-Pretty, F., Smith, A.J., Cooper, J., Palmen, J., Folkersen, L., Hamsten, A., Catapano, A.L., Melander, O., Price, J.F., Kumari, M., et al. (2015). Functional analysis of a carotid intima-media thickness locus implicates BCAR1 and suggests a causal variant. Circ Cardiovasc Genet *8*, 696–706.
60. Koshimizu, T.A., Nasa, Y., Tanoue, A., Oikawa, R., Kawahara, Y., Kiyono, Y., Adachi, T., Tanaka, T., Kuwaki, T., Mori, T., et al. (2006). V1a vasopressin receptors maintain normal blood pressure by regulating circulating blood volume and baroreflex sensitivity. Proc. Natl. Acad. Sci. USA *103*, 7807–7812.
61. Nelson, C.P., Goel, A., Butterworth, A.S., Kanoni, S., Webb, T.R., Marouli, E., Zeng, L., Ntalla, I., Lai, F.Y., Hopewell, J.C., et al.; EPIC-CVD Consortium; CARDIoGRAMplusC4D; and UK Biobank CardioMetabolic Consortium CHD working group (2017). Association analyses based on false discovery rate implicate new loci for coronary artery disease. Nat. Genet. *49*, 1385–1391.
62. Yu, K.Y., Wang, Y.P., Wang, L.H., Jian, Y., Zhao, X.D., Chen, J.W., Murao, K., Zhu, W., Dong, L., Wang, G.Q., and Zhang, G.X. (2014). Mitochondrial KATP channel involvement in angiotensin II-induced autophagy in vascular smooth muscle cells. Basic Res. Cardiol. *109*, 416.
63. Del Greco M, F., Pattaro, C., Luchner, A., Pichler, I., Winkler, T., Hicks, A.A., Fuchsberger, C., Franke, A., Melville, S.A., Peters, A., et al. (2011). Genome-wide association analysis and fine mapping of NT-proBNP level provide novel insight into the role of the MTHFR-CLCN6-NPPA-NPPB gene cluster. Hum. Mol. Genet. *20*, 1660–1671.
64. Ganesh, S.K., Tragante, V., Guo, W., Guo, Y., Lanktree, M.B., Smith, E.N., Johnson, T., Castillo, B.A., Barnard, J., Baumert, J., et al.; CARDIOGRAM, METASTROKE; and LifeLines Cohort Study (2013). Loci influencing blood pressure identified using a cardiovascular gene-centric array. Hum. Mol. Genet. *22*, 1663–1678.
65. Johnson, T., Gaunt, T.R., Newhouse, S.J., Padmanabhan, S., Tomaszewski, M., Kumari, M., Morris, R.W., Tzoulaki, I., O'Brien, E.T., Poulter, N.R., et al.; Cardiogenics Consortium; and Global BPgen Consortium (2011). Blood pressure loci identified with a gene-centric array. Am. J. Hum. Genet. *89*, 688–700.
66. Kathiresan, S., Willer, C.J., Peloso, G.M., Demissie, S., Musunuru, K., Schadt, E.E., Kaplan, L., Bennett, D., Li, Y., Tanaka, T., et al. (2009). Common variants at 30 loci contribute to polygenic dyslipidemia. Nat. Genet. *41*, 56–65.
67. Tomaszewski, M., Charchar, F.J., Barnes, T., Gawron-Kiszka, M., Sedkowska, A., Podolecka, E., Kowalczyk, J., Rathbone, W., Kalarus, Z., Grzeszczak, W., et al. (2009). A common variant in low-density lipoprotein receptor-related protein 6 gene (LRP6) is associated with LDL-cholesterol. Arterioscler. Thromb. Vasc. Biol. *29*, 1316–1321.
68. Singh, R., Smith, E., Fathzadeh, M., Liu, W., Go, G.W., Subrahmanyan, L., Faramarzi, S., McKenna, W., and Mani, A. (2013). Rare nonconservative LRP6 mutations are associated with metabolic syndrome. Hum. Mutat. *34*, 1221–1225.
69. Waterworth, D.M., Ricketts, S.L., Song, K., Chen, L., Zhao, J.H., Ripatti, S., Aulchenko, Y.S., Zhang, W., Yuan, X., Lim, N., et al.; Wellcome Trust Case Control Consortium (2010). Genetic variants influencing circulating lipid levels and risk of coronary artery disease. Arterioscler. Thromb. Vasc. Biol. *30*, 2264–2276.
70. Reilly, D., Hao, K., Jensen, M.K., Girman, C.J., and Rimm, E.B. (2013). Use of systems biology approaches to analysis of genome-wide association studies of myocardial infarction and blood cholesterol in the nurses' health study and health professionals' follow-up study. PLoS ONE *8*, e85369.
71. Liew, C.W., Boucher, J., Cheong, J.K., Vernochet, C., Koh, H.J., Mallol, C., Townsend, K., Langin, D., Kawamori, D., Hu, J., et al. (2013). Ablation of TRIP-Br2, a regulator of fat lipolysis, thermogenesis and oxidative metabolism, prevents diet-induced obesity and insulin resistance. Nat. Med. *19*, 217–226.
72. Speliotes, E.K., Yerges-Armstrong, L.M., Wu, J., Hernaez, R., Kim, L.J., Palmer, C.D., Gudnason, V., Eiriksdottir, G., Garcia, M.E., Launer, L.J., et al.; NASH CRN; GIANT Consortium; MAGIC Investigators; and GOLD Consortium (2011). Genome-wide association analysis identifies variants associated with nonalcoholic fatty liver disease that have distinct effects on metabolic traits. PLoS Genet. *7*, e1001324.
73. Frayling, T.M., Timpson, N.J., Weedon, M.N., Zeggini, E., Freathy, R.M., Lindgren, C.M., Perry, J.R., Elliott, K.S., Lango, H., Rayner, N.W., et al. (2007). A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. Science *316*, 889–894.
74. Stratigopoulos, G., Martin Carli, J.F., O'Day, D.R., Wang, L., Leduc, C.A., Lanzano, P., Chung, W.K., Rosenbaum, M., Egli, D., Doherty, D.A., and Leibel, R.L. (2014). Hypomorphism for RPGRIP1L, a ciliary gene vicinal to the FTO locus, causes increased adiposity in mice. Cell Metab. *19*, 767–779.
75. Claussnitzer, M., Dankel, S.N., Kim, K.H., Quon, G., Meuleman, W., Haugen, C., Glunk, V., Sousa, I.S., Beaudry, J.L., Puviindran, V., et al. (2015). FTO obesity variant circuitry and adipocyte browning in humans. N. Engl. J. Med. *373*, 895–907.
76. Stratigopoulos, G., LeDuc, C.A., Cremona, M.L., Chung, W.K., and Leibel, R.L. (2011). Cut-like homeobox 1 (CUX1) regulates expression of the fat mass and obesity-associated and retinitis pigmentosa GTPase regulator-interacting protein-1-like (RPGRIP1L) genes and coordinates leptin receptor signaling. J. Biol. Chem. *286*, 2155–2170.
77. Kawaguchi, T., Tamori, Y., Kanda, H., Yoshikawa, M., Tateya, S., Nishino, N., and Kasuga, M. (2010). The t-SNAREs syntaxin4 and SNAP23 but not v-SNARE VAMP2 are indispensable to tether GLUT4 vesicles at the plasma membrane in adipocyte. Biochem. Biophys. Res. Commun. *391*, 1336–1341.
78. Pelsers, M.M., Hermens, W.T., and Glatz, J.F. (2005). Fatty acid-binding proteins as plasma markers of tissue injury. Clin. Chim. Acta *352*, 15–35.
79. Stoynev, N., Dimova, I., Rukova, B., Hadjidekova, S., Nikolova, D., Toncheva, D., and Tankova, T. (2014). Gene expression in peripheral blood of patients with hypertension and patients with type 2 diabetes. J. Cardiovasc. Med. (Hagerstown) *15*, 702–709.
80. Setsuta, K., Seino, Y., and Mizuno, K. (2014). Heart-type fatty acid-binding protein is a novel prognostic marker in patients with essential hypertension. Int. J. Cardiol. *176*, 1323–1325.
81. Swan, G.E., Hops, H., Wilhelmsen, K.C., Lessov-Schlaggar, C.N., Cheng, L.S., Hudmon, K.S., Amos, C.I., Feiler, H.S., Ring, H.Z., Andrews, J.A., et al. (2006). A genome-wide screen for nicotine dependence susceptibility loci. Am. J. Med. Genet. B. Neuropsychiatr. Genet. *141B*, 354–360.
82. Bierut, L.J., Madden, P.A., Breslau, N., Johnson, E.O., Hatsukami, D., Pomerleau, O.F., Swan, G.E., Rutter, J., Bertelsen,

S., Fox, L., et al. (2007). Novel genes identified in a high-density genome wide association study for nicotine dependence. Hum. Mol. Genet. *16*, 24–35.
83. Levran, O., Peles, E., Randesi, M., Correa da Rosa, J., Ott, J., Rotrosen, J., Adelson, M., and Kreek, M.J. (2015). Susceptibility loci for heroin and cocaine addiction in the serotonergic and adrenergic pathways in populations of different ancestry. Pharmacogenomics *16*, 1329–1342.
84. Taylor, A., and Wang, K.S. (2014). Association between DPYSL2 gene polymorphisms and alcohol dependence in Caucasian samples. J. Neural Transm. (Vienna) *121*, 105–111.
85. Edwards, A.C., Aliev, F., Bierut, L.J., Bucholz, K.K., Edenberg, H., Hesselbrock, V., Kramer, J., Kuperman, S., Nurnberger, J.I., Jr., Schuckit, M.A., et al. (2012). Genome-wide association study of comorbid depressive syndrome and alcohol dependence. Psychiatr. Genet. *22*, 31–41.
86. Nutt, D.J., Lingford-Hughes, A., Erritzoe, D., and Stokes, P.R. (2015). The dopamine theory of addiction: 40 years of highs and lows. Nat. Rev. Neurosci. *16*, 305–312.
87. Saccone, N.L., Schwantes-An, T.H., Wang, J.C., Grucza, R.A., Breslau, N., Hatsukami, D., Johnson, E.O., Rice, J.P., Goate, A.M., and Bierut, L.J. (2010). Multiple cholinergic nicotinic receptor genes affect nicotine dependence risk in African and European Americans. Genes Brain Behav. *9*, 741–750.
88. Simms, R.J., Hynes, A.M., Eley, L., and Sayer, J.A. (2011). Nephronophthisis: a genetically diverse ciliopathy. Int. J. Nephrol. *2011*, 527137.
89. Schueler, M., Halbritter, J., Phelps, I.G., Braun, D.A., Otto, E.A., Porath, J.D., Gee, H.Y., Shendure, J., O'Roak, B.J., Lawson, J.A., et al. (2016). Large-scale targeted sequencing comparison highlights extreme genetic heterogeneity in nephronophthisis-related ciliopathies. J. Med. Genet. *53*, 208–214.
90. Croft, J.B., and Swift, M. (1990). Obesity, hypertension, and renal disease in relatives of Bardet-Biedl syndrome sibs. Am. J. Med. Genet. *36*, 37–42.
91. Ambatipudi, S., Cuenin, C., Hernandez-Vargas, H., Ghantous, A., Le Calvez-Kelm, F., Kaaks, R., Barrdahl, M., Boeing, H., Aleksandrova, K., Trichopoulou, A., et al. (2016). Tobacco smoking-associated genome-wide DNA methylation changes in the EPIC study. Epigenomics *8*, 599–618.
92. Young, A.I., Wauthier, F., and Donnelly, P. (2016). Multiple novel gene-by-environment interactions modify the effect of FTO variants on body mass index. Nat. Commun. *7*, 12724.
93. Tyrrell, J., Wood, A.R., Ames, R.M., Yaghootkar, H., Beaumont, R.N., Jones, S.E., Tuke, M.A., Ruth, K.S., Freathy, R.M., Davey Smith, G., et al. (2017). Gene-obesogenic environment interactions in the UK Biobank study. Int. J. Epidemiol. *46*, 559–575.
94. Hoffmann, T.J., Ehret, G.B., Nandakumar, P., Ranatunga, D., Schaefer, C., Kwok, P.Y., Iribarren, C., Chakravarti, A., and Risch, N. (2017). Genome-wide association analyses using electronic health records identify new loci influencing blood pressure variation. Nat. Genet. *49*, 54–64.
95. Warren, H.R., Evangelou, E., Cabrera, C.P., Gao, H., Ren, M., Mifsud, B., Ntalla, I., Surendran, P., Liu, C., Cook, J.P., et al.; International Consortium of Blood Pressure (ICBP) 1000G Analyses; BIOS Consortium; Lifelines Cohort Study; Understanding Society Scientific group; CHD Exome+ Consortium; ExomeBP Consortium; T2D-GENES Consortium; GoT2DGenes Consortium; Cohorts for Heart and Ageing Research in Genome Epidemiology (CHARGE) BP Exome Consortium; International Genomics of Blood Pressure (iGEN-BP) Consortium; and UK Biobank CardioMetabolic Consortium BP working group (2017). Genome-wide association analysis identifies novel blood pressure loci and offers biological insights into cardiovascular risk. Nat. Genet. *49*, 403–415.
96. Liang, J., Le, T.H., Edwards, D.R.V., Tayo, B.O., Gaulton, K.J., Smith, J.A., Lu, Y., Jensen, R.A., Chen, G., Yanek, L.R., et al. (2017). Single-trait and multi-trait genome-wide association analyses identify novel loci for blood pressure in African-ancestry populations. PLoS Genet. *13*, e1006728.
97. Keller, M.C. (2014). Gene × environment interaction studies have not properly controlled for potential confounders: the problem and the (simple) solution. Biol. Psychiatry *75*, 18–24.
98. Shi, G., and Nehorai, A. (2017). Robustness of meta-analyses in finding gene × environment interactions. PLoS ONE *12*, e0171446.

The American Journal of Human Genetics, Volume *102*

Supplemental Data

A Large-Scale Multi-ancestry Genome-wide Study Accounting for Smoking Behavior Identifies Multiple Significant Loci for Blood Pressure

Yun J. Sung, Thomas W. Winkler, Lisa de las Fuentes, Amy R. Bentley, Michael R. Brown, Aldi T. Kraja, Karen Schwander, Ioanna Ntalla, Xiuqing Guo, Nora Franceschini, Yingchang Lu, Ching-Yu Cheng, Xueling Sim, Dina Vojinovic, Jonathan Marten, Solomon K. Musani, Changwei Li, Mary F. Feitosa, Tuomas O. Kilpeläinen, Melissa A. Richard, Raymond Noordam, Stella Aslibekyan, Hugues Aschard, Traci M. Bartz, Rajkumar Dorajoo, Yongmei Liu, Alisa K. Manning, Tuomo Rankinen, Albert Vernon Smith, Salman M. Tajuddin, Bamidele O. Tayo, Helen R. Warren, Wei Zhao, Yanhua Zhou, Nana Matoba, Tamar Sofer, Maris Alver, Marzyeh Amini, Mathilde Boissel, Jin Fang Chai, Xu Chen, Jasmin Divers, Ilaria Gandin, Chuan Gao, Franco Giulianini, Anuj Goel, Sarah E. Harris, Fernando Pires Hartwig, Andrea R.V.R. Horimoto, Fang-Chi Hsu, Anne U. Jackson, Mika Kähönen, Anuradhani Kasturiratne, Brigitte Kühnel, Karin Leander, Wen-Jane Lee, Keng-Hung Lin, Jian 'an Luan, Colin A. McKenzie, He Meian, Christopher P. Nelson, Rainer Rauramaa, Nicole Schupf, Robert A. Scott, Wayne H.H. Sheu, Alena Stančáková, Fumihiko Takeuchi, Peter J. van der Most, Tibor V. Varga, Heming Wang, Yajuan Wang, Erin B. Ware, Stefan Weiss, Wanqing Wen, Lisa R. Yanek, Weihua Zhang, Jing Hua Zhao, Saima Afaq, Tamuno Alfred, Najaf Amin, Dan Arking, Tin Aung, R. Graham Barr, Lawrence F. Bielak, Eric Boerwinkle, Erwin P. Bottinger, Peter S. Braund, Jennifer A. Brody, Ulrich Broeckel, Claudia P. Cabrera, Brian Cade, Yu Caizheng, Archie Campbell, Mickaël Canouil, Aravinda Chakravarti, The CHARGE Neurology Working Group, Ganesh Chauhan, Kaare Christensen, Massimiliano Cocca, The COGENT-Kidney Consortium, Francis S. Collins, John M. Connell, Renée de Mutsert, H. Janaka de Silva, Stephanie Debette, Marcus Dörr, Qing Duan, Charles B. Eaton, Georg Ehret, Evangelos Evangelou, Jessica D. Faul, Virginia A. Fisher, Nita G. Forouhi, Oscar H. Franco, Yechiel Friedlander, He Gao, The GIANT Consortium, Bruna Gigante, Misa Graff, C. Charles Gu, Dongfeng Gu, Preeti Gupta, Saskia P. Hagenaars, Tamara B. Harris, Jiang He, Sami Heikkinen, Chew-Kiat Heng, Makoto Hirata, Albert Hofman, Barbara V. Howard, Steven Hunt, Marguerite R. Irvin, Yucheng Jia, Roby Joehanes, Anne E. Justice, Tomohiro Katsuya, Joel Kaufman, Nicola D. Kerrison, Chiea Chuen Khor, Woon-Puay Koh, Heikki A. Koistinen, Pirjo Komulainen, Charles Kooperberg, Jose E. Krieger, Michiaki Kubo, Johanna Kuusisto, Carl D. Langefeld, Claudia Langenberg, Lenore J. Launer, Benjamin Lehne, Cora E. Lewis, Yize Li, Lifelines Cohort Study, Sing Hui Lim, Shiow Lin, Ching-Ti Liu, Jianjun Liu, Jingmin Liu, Kiang Liu, Yeheng Liu, Marie Loh, Kurt K. Lohman, Jirong Long, Tin Louie, Reedik Mägi, Anubha Mahajan, Thomas Meitinger, Andres Metspalu, Lili Milani, Yukihide Momozawa, Andrew P. Morris, Thomas H. Mosley Jr., Peter Munson, Alison D. Murray, Mike A. Nalls, Ubaydah Nasri, Jill M. Norris, Kari North, Adesola

Table of Contents

Table S7: Genotyping, imputation, and statistical analysis for Stage 2 (focused discovery/replication) analysis.
Table S8: Percent (%) BP variance explained by known and newly-identified index variants across multiple studies/ancestries.

Table S9: Effect sizes and p-values at 3,391 genome-wide significant variants identified from combined Stage 1 and 2 analyses.

Table S10: False discovery rate (FDR) q values at the 66 index variants of loci additionally identified through the combined Stage 1 and 2 meta-analysis.

Table S11: Known BP variants identified by our combined results.

Table S12: Regulation information of the 81 index variants representing novel loci obtained from HaploReg and RegulomeDB.

Table S13: eQTL evidence for the 21 index variants obtained from GTEx.

Table S14: eQTL evidence in the Framingham Heart Study.

Table S15: Results from the DEPICT gene prioritization for the genome-wide significant European SBP loci.

Table S16: Results from the DEPICT gene prioritization for the genome-wide significant European DBP loci.

Table S17: Results from the DEPICT gene prioritization for the genome-wide significant trans-ancestry SBP loci.

Table S18: Results from the DEPICT gene prioritization for the genome-wide significant trans-ancestry DBP loci.

Table S19: Results from the four DEPICT gene set enrichment analyses (European SBP and DBP, trans-ancestry SBP and DBP). The table is limited to gene sets that showed significant enrichment (FDR < 0.05) in at least one of the four separate analyses.

Table S20: Results from the four DEPICT cell type and tissue enrichment analyses on European SBP and DBP and trans-ancestry SBP and DBP.

Table S21: Look-up information for coronary artery disease traits at the 81 index variants.

Table S22: Look-up information for stroke traits at the 81 index variants.

Table S23: Look-up information for adiposity traits (by itself and stratified by gender) at the 81 index variants.

Table S24: Look-up information for diabetes traits at the 81 index variants.

Table S25: Look-up information for kidney traits at the 81 index variants.

Table S26: Look-up information for adiposity traits stratified by smoking at the 81 index variants

Table S27: Summary of lookup information (Tables S21-S26). Boldfaced red are those for which the number of variants with observed nominal significant association ($P < 0.05$) is higher than that expected by chance alone ($P_{Binomial} < 0.05/11$, corrected for 11 traits used in the lookups: coronary artery disease, myocardial infarction, stroke, body mass index, height, waist circumferrence, waist-hip ratio, fasting glucose, fasting insulin, T2D, eGFR).

Table S28: Overlap of our novel loci with those identified by the two recent BP papers (Hoffmann et al, Warren et al)

Supplemental Notes

More Details on the Quality Control (QC)

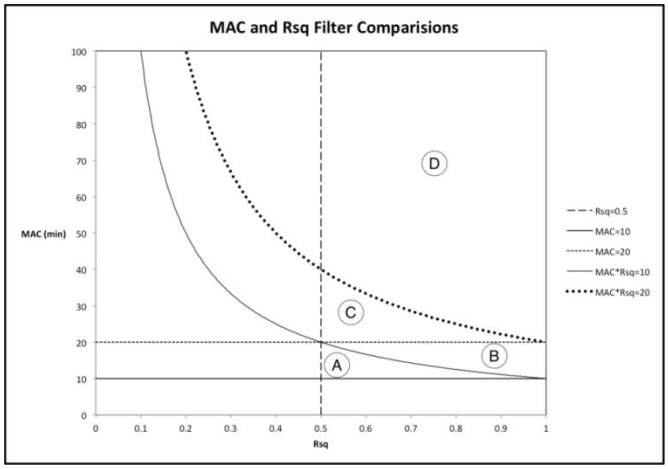
Cohorts participating in this study have ample experience in main-effect based GWAS for multiple phenotypes and are very familiar with validated approaches for quality control (QC) of phenotype, genotype, and imputed data. However, because of the use of interaction models and imputed data using the 1000 Genome Project, we were particularly thorough on QC steps. We relied heavily on the package EasyQC (Winkler et al, Nature Protocols 2014), which was extended for interaction analysis with the 1000 Genomes-based imputed data by the developer. In addition, we contrasted results from the joint model and stratified models in Stage 1 cohorts, as explained more in Sung et al (Genetic Epidemiology 2016). Any unusual findings or patterns were resolved together with the study analyst; in some cases, cohorts were asked to repeat the analysis. The Supplemental Material in Rao et al (Circulation Cardiovascular Genetics 2017; pages 21-23) covers these QC steps in more detail.

One of the latter QC steps involved determining which filter was most appropriate for excluding unstable cohort-specific results that reflect small sample size, low MAF, or low imputation quality measures. Among the various filters considered, we finally used

DF = min (MAC0, MAC1) * imputation quality measure,

where MAC is the minor allele count in each stratum: MAC0 = 2 * MAF_{E0} * N_{E0} and MAC1 = 2 * MAF_{E1} * N_{E1}, for the unexposed (E0) and exposed group, respectively. In addition to imputation quality measure ≥ 0.5, we considered the following four filtering thresholds. (E1)

	Filters	Area (in Figure right)
1	min(MAC0, MAC1) ≥10	A+B+C+D
2	DF ≥ 10	B+C+D
3	min(MAC0, MAC1) ≥20	C+D
4	DF ≥ 20	D



After reviewing the QQ plots for each section (A, B, C, D) separately (Figures S14-S17), we decided to use DF ≥ 20. We could clearly see that the QQ plots for section D (5th column in QQ plots) were much better behaved.

Stage 1 (Genome-wide Discovery) Study Descriptions

Brief descriptions are provided below for each of the discovery studies some of which are based outside the United States.:

AGES (Age Gene/Environment Susceptibility Reykjavik Study): The AGES Reykjavik study originally comprised a random sample of 30,795 men and women born in 1907-1935 and living in Reykjavik in 1967. A total of 19,381 people attended, resulting in a 71% recruitment rate. The study sample was divided into six groups by birth year and birth date within month. One group was designated for longitudinal follow up and was examined in all stages; another was designated as a control group and was not included in examinations until 1991. Other groups were invited to participate in specific stages of the study. Between 2002 and 2006, the AGESReykjavik study re-examined 5,764 survivors of the original cohort who had participated before in the Reykjavik Study. The midlife data blood pressure measurement was taken from stage 3 of the Reykjavik Study (1974-1979), if available. Half of the cohort attended during this period. Otherwise an observation was selected closest in time to the stage 3 visit. The supine blood pressure was measured twice by a nurse using a mercury sphygmomanometer after 5 minutes rest following World Health Organization recommendations.

ARIC (Atherosclerosis Risk in Communities): The ARIC study is a population-based prospective cohort study of cardiovascular disease sponsored by the National Heart, Lung, and Blood Institute (NHLBI). ARIC included 15,792 individuals, predominantly European American and African American,aged 45-64 years at baseline (1987-89), chosen by probability sampling from four US communities. Cohort members completed three additional triennial follow-up examinations and a fifth exam in 2011-2013. The ARIC study has been described in detail previously (The ARIC Investigators.The Atherosclerosis Risk in Communities (ARIC) study: Design and objectives. Am J Epidemiol. 1989;129:687-702).. Blood pressure was measured using a standardized Hawksley random-zero mercury column sphygmomanometer with participants in a sitting position after a resting period of 5 minutes. The size of the cuff was chosen according to the arm circumference. Three sequential recordings for systolic and diastolic blood pressure were obtained; the mean of the last two measurements was used in this analysis, discarding the first reading. Blood pressure lowering medication use was recorded from the medication history.

Baependi Heart Study (Brazil): The Baependi Heart Study, is an ongoing family-based cohort conducted in a rural town of the state of Minas Gerais. The study has enrolled approximate 2,200 individuals (over 10% of the town's adult population) and 10-year follow up period of longitudinal data. Briefly, probands were selected at random across 11 out of the 12 census districts in Baependi. After enrolment, the proband's first-degree (parents, siblings, and offspring), second-degree (half-siblings, grandparents/grandchildren, uncles/aunts, nephews/nieces, and double cousins), and third-degree (first cousins, great uncles/aunts, and great nephews/nieces) relatives, and his/her respective spouse's relatives resident both within Baependi (municipal and rural area) and surrounding towns were invited to participate. Only individuals age 18 and older were eligible to participate in the study. The study is conducted from a clinic/office in an easily accessible sector of the town, where the questionnaires were completed. A broad range of phenotypes ranging from cardiovascular, neurocognitive, psychiatric, imaging, physiologic and several layers of endophenotypes like metabolomics and lipidomics have been collected throughout the years Details about follow-up visits and available data can be found in the cohort profile paper (PMID: 18430212). DNA samples were genotyped using the Affymetrix 6.0 genechip. After quality control, the data were prephased using SHAPEIT and imputed using IMPUTE2 based on 1000 Genomes haplotypes.

Bio*Me* Biobank (Bio*Me* Biobank of Institute for Personalized Medicine at Mount Sinai): The Bio*Me* Biobank, founded in September 2007, is an ongoing, consented electronic medical record (EMR)-linked bio- and data repository that enrolls participants non-selectively from the Mount Sinai

Medical Center patient population. The Bio*Me* Biobank currently (Winter 2015) comprises over 31,000 participants from diverse ancestries characterized by a broad spectrum of (longitudinal) biomedical traits. On average 400 new participants are consented each month. Bio*Me* participants represent the broad ancestral, ethnic and socioeconomic diversity with a distinct and population-specific disease burden, characteristic of Northern Manhattan communities served by Mount Sinai Hospital. Enrolled participants consent to be followed throughout their clinical care (past, present, and future) at Mount Sinai in real-time, integrating their genomic information with their electronic health record for discovery research and clinical care implementation. Bio*Me* participants are predominantly of African, Hispanic/Latino, and European ancestry. Participants who self-identify as Hispanic/Latino further report to be of Puerto Rican (39%), Dominican (23%), Central/South American (17%), Mexican (5%) or other Hispanic (16%) ancestry. More than 40% of European ancestry participants are genetically determined to be of Ashkenazi Jewish ancestry.

The IRB-approved Bio*Me* Biobank consent permits use of samples and de-identified linkable past, present and future clinical information from EMRs; re-contacting participants for enrollment in future research; unlimited duration of storage, and access to clinical information from the entire medical records, as well as local and external sharing of specimens and data.

The Bio*Me* Biobank has a longitudinal design as participants consent to make any EMR data from past (dating back as far as 2003), present and future inpatient or outpatient encounters available for research. The median number of clinical encounters per participant is 21, reflecting predominant enrollment of participants with common chronic conditions from primary care facilities. Mount Sinai's system-wide Epic EMR implementation captures a full spectrum of biomedical phenotypes, including clinical outcomes, covariate and exposure data. This clinical information is complemented by detailed information on ancestry, residence history, familial medical history, education, socio-economic status, physical activity, smoking, alcohol use, and weight history being collected in a systematic manner by interview-based questionnaire at time of enrollment. Phenotype harmonization and validation is critical to facilitate consortium-wide analyses. By applying advanced medical informatics and data mining tools, high-quality and validated phenotype data can be culled from Mount Sinai's Epic EMR. Fully-implemented phenotype algorithms include; T2D, CKD, CAD, lipid disorders, peripheral artery disease, resistant hypertension, blood cell traits, abdominal aortic aneurism, venous thromboembolism among others (see also Phenotype KnowledgeBase (PheKB) of the eMERGE Network (http://emerge.mc.vanderbilt.edu/emerge-network).

A total of 14,017 participants have been genotyped for both GWAS (11,150 Illumina OmniExpress BeadChip, 2,867 Affymetrix Human SNP Array 6.0) *and* ExomeChip (Illumina HumanExome v1.0 BeadChip) arrays funded by institutional sources. An additional 16,000 Bio*Me* participants are scheduled for genotyping using the Illumina MEGA Chip (by April 2015), funded by NHGRI through our PAGEII grant (U01HG007417) (n=12,500) and through institutional funds (n=3,500).

CARDIA (Coronary Artery Risk Development in Young Adults): CARDIA is a prospective multicenter study with 5,115 adults Caucasian and African American participants of the age group 18-30 years, recruited from four centers at the baseline examination in 1985-1986. The recruitment was done from the total community in Birmingham, AL, from selected census tracts in Chicago, IL and Minneapolis, MN; and from the Kaiser Permanente health plan membership in Oakland, CA. The details of the study design for the CARDIA study have been previously published. Eight examinations have been completed since initiation of the study, respectively in the years 0, 2, 5, 7, 10, 15, 20 and 25. Written informed consent was obtained from participants at each examination and all study protocols were approved by the institutional review boards of the participating institutions. Systolic and diastolic blood pressure was measured in triplicate on the right arm using a random-zero sphygmomanometer with the participant seated and following a 5-min. rest. The average of the second and third measurements was taken as the blood pressure value. Blood pressure medication use was obtained by questionnaire.

CHS (Cardiovascular Health Study): CHS is a population-based cohort study of risk factors for cardiovascular disease in adults 65 years of age or older conducted across four field centers [PMID: 1669507]. The original predominantly European ancestry cohort of 5,201 persons was recruited in 1989-1990 from random samples of the Medicare eligibility lists and an additional predominately African-American cohort of 687 persons was enrolled in 1992-93 for a total sample of 5,888. Research staff with central training in blood pressure measurement assessed repeated right-arm seated systolic and diastolic blood pressure levels at baseline with a Hawksley random-zero sphygmomanometer. Blood samples were drawn from all participants at their baseline examination and DNA was subsequently extracted from available samples. European ancestry participants were excluded from the GWAS study sample due to prevalent coronary heart disease, congestive heart failure, peripheral vascular disease, valvular heart disease, stroke, or transient ischemic attack at baseline. After QC, genotyping was successful for 3271 European ancestry and 823 African-American participants. CHS was approved by institutional review committees at each site and individuals in the present analysis gave informed consent including consent to use of genetic information for the study of cardiovascular disease.

ERF (Erasmus Rucphen Family study): Erasmus Rucphen Family is a family based study that includes inhabitants of a genetically isolated community in the South-West of the Netherlands, studied as part of the Genetic Research in Isolated Population (GRIP) program. The goal of the study is to identify the risk factors in the development of complex disorders. Study population includes approximately 3,000 individuals who are living descendants of 22 couples who lived in the isolate between 1850 and 1900 and had at least six children baptized in the community church. All data were collected between 2002 and 2005. All participants gave informed consent, and the Medical Ethics Committee of the Erasmus University Medical Centre approved the study.

FamHS (Family Heart Study): The NHLBI FamHS study design, collection of phenotypes and covariates as well as clinical examination have been previously described (https://dsgweb.wustl.edu/fhscc/; PMID: 8651220). In brief, the FamHS recruited 1,200 families (approximately 6,000 individuals), half randomly sampled, and half selected because of an excess of coronary heart disease (CHD) or risk factor abnormalities as compared with age- and sex-specific population rates. The participants were sampled from four population-based parent studies: the Framingham Heart Study, the Utah Family Tree Study, and two centers for the Atherosclerosis Risk in Communities study (ARIC: Minneapolis, and Forsyth County, NC). These individuals attended a clinic exam (1994-1996) and a broad range of phenotypes were assessed in the general domains of CHD, atherosclerosis, cardiac and vascular function, inflammation and hemostasis, lipids and lipoproteins, blood pressure, diabetes and insulin resistance, pulmonary function, diet, education, socioeconomic status, habitual behavior, physical activity, anthropometry, medical history and medication use. Approximately 8 years later, study participants belonging to the largest pedigrees were invited for a second clinical exam (2002-04). The most important CHD risk factors were measured again, including lipids, parameters of glucose metabolism, blood pressure, anthropometry, and several biochemical and hematologic markers. In addition, a computed tomography examination provided measures of coronary and aortic calcification, and abdominal and liver fat burden. Medical history and medication use was updated. A total of 2,756 European ancestry subjects in 510 extended random and high CHD risk families were studied. Also, 633 African ancestry subjects were recruited at ARIC field center at the University of Alabama in Birmingham. Informed consent was obtained from all participants.

FHS (Framingham Heart Study): FHS began in 1948 with the recruitment of an original cohort of 5,209 men and women (mean age 44 years; 55 percent women). In 1971 a second generation of study participants was enrolled; this cohort (mean age 37 years; 52% women) consisted of 5,124 children and spouses of children of the original cohort. A third generation cohort of 4,095 children of offspring cohort participants (mean age 40 years; 53 percent women) was enrolled in 2002-2005 and are seen every 4 to 8 years. Details of study designs for the three cohorts are summarized elsewhere. At each clinic visit, a medical history was obtained with a focus on cardiovascular content, and participants

underwent a physical examination including measurement of height and weight from which BMI was calculated. Systolic and diastolic blood pressures were measured twice by a physician on the left arm of the resting and seated participant using a mercury column sphygmomanometer. Blood pressures were recorded to the nearest even number. The means of two separate systolic and diastolic blood pressure readings at each clinic examination were used for statistical analyses.

GENOA (Genetic Epidemiology Network of Arteriopathy): GENOA is one of four networks in the NHLBI Family-Blood Pressure Program (FBPP).[The FBPP Investigators. Multi-center genetic study of hypertension: The Family Blood Pressure Program (FBPP). Hypertension 2002;39:3-9.; Daniels PR, Kardia SL, Hanis CL, Brown CA, Hutchinson R, Boerwinkle E, Turner ST; Genetic Epidemiology Network of Arteriopathy study. Familial aggregation of hypertension treatment and control in the Genetic Epidemiology Network of Arteriopathy (GENOA) study. Am J Med. 2004 May 15;116(10):676-81. PubMed PMID: 15121494.] GENOA's long-term objective is to elucidate the genetics of target organ complications of hypertension, including both atherosclerotic and arteriolosclerotic complications involving the heart, brain, kidneys, and peripheral arteries. The longitudinal GENOA Study recruited European-American and African-American sibships with at least 2 individuals with clinically diagnosed essential hypertension before age 60 years. All other members of the sibship were invited to participate regardless of their hypertension status. Participants were diagnosed with hypertension if they had either 1) a previous clinical diagnosis of hypertension by a physician with current anti-hypertensive treatment, or 2) an average systolic blood pressure ≥ 140 mm Hg or diastolic blood pressure ≥ 90 mm Hg based on the second and third readings at the time of their clinic visit. Exclusion criteria were secondary hypertension, alcoholism or drug abuse, pregnancy, insulin-dependent diabetes mellitus, or active malignancy. During the first exam (1995-2000), 1,583 European Americans from Rochester, MN and 1,854 African Americans from Jackson, MS were examined. Between 2000 and 2005, 1,241 of the European Americans and 1,482 of the African Americans returned for a second examination. Because African-American probands for GENOA were recruited through the Atherosclerosis Risk in Communities (ARIC) Jackson field center participants, we excluded ARIC participants from analyses.

GenSalt (Genetic Epidemiology Network of Salt Sensitivity): GenSalt is a multi-center, family based study designed to identify, through dietary sodium and potassium intervention, salt-sensitivitivity susceptibility genes which may underlie essential hypertension in rural Han Chinese families. Approximately 629 families with at least one ‘proband’ with high blood pressure were recruited and tested for a wide variety of physiological, metabolic and biochemical measures at baseline and at multiple times during the 3-week intervention. The intervention consisted of one week on a low sodium diet, followed by one week on a high sodium diet, and finally one week on a high sodium diet with a potassium supplement.

GOLDN (Genetics of Diet and Lipid Lowering Network): GOLDN is a multi-center family pharmacogenetic study that is investigating gene- environment interactions on lipid profiles. 1,200 subjects in extended pedigrees were measured before and after two environmental exposures: 1) a dietary fat challenge to assess genetic regulators of fat uptake and clearance and 2) a 3 week clinical trial of fenofibrate to assess pharmacogenetic influences on response to treatment. The goals of the study are to identify and characterize genetic loci that predict the lipid profile treatment responses. https://dsgweb.wustl.edu/PROJECTS/MP5.html

HANDLS (Healthy Aging in Neighborhoods of Diversity across the Life Span): HANDLS is a community-based, longitudinal epidemiologic study examining the influences of race and socioeconomic status (SES) on the development of age-related health disparities among a sample of socioeconomically diverse African Americans and whites. This unique study will assess over a 20-year period physical parameters and also evaluate genetic, biologic, demographic, and psychosocial, parameters of African American and white participants in higher and lower SES to understand the driving factors behind persistent black-white health disparities in overall longevity, cardiovascular disease, and cognitive decline. The study recruited 3,722 participants from Baltimore, MD with a mean

age of 47.7 years, 2,200 African Americans and 1,522 whites, with 41% reporting household incomes below the 125% poverty delimiter.

Genotyping was done on a subset of self-reporting African American participants by the Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health (NIH). A larger genotyping effort included a small subset of self-reporting European ancestry samples. This research was supported by the Intramural Research Program of the NIH, NIA and the National Center on Minority Health and Health Disparities.

Health ABC (Health, Aging, and Body Composition): Cohort description: The Health ABC study is a prospective cohort study investigating the associations between body composition, weight-related health conditions, and incident functional limitation in older adults. Health ABC enrolled well-functioning, community-dwelling black (n=1281) and white (n=1794) men and women aged 70-79 years between April 1997 and June 1998. Participants were recruited from a random sample of white and all black Medicare eligible residents in the Pittsburgh, PA, and Memphis, TN, metropolitan areas. Participants have undergone annual exams and semi-annual phone interviews. The current study sample consists of 1559 white participants who attended the second exam in 1998-1999 with available genotyping data.

Genotyping: Genotyping was performed by the Center for Inherited Disease Research (CIDR) using the Illumina Human1M-Duo BeadChip system. Samples were excluded from the dataset for the reasons of sample failure, genotypic sex mismatch, and first-degree relative of an included individual based on genotype data. Genotyping was successful in 1663 Caucasians. Analysis was restricted to SNPs with minor allele frequency ≥ 1%, call rate ≥97% and HWE p≥10-6. Genotypes were available on 914,263 high quality SNPs for imputation based on the HapMap CEU (release 22, build 36) using the MACH software (version 1.0.16). A total of 2,543,888 imputed SNPs were analyzed for association with vitamin D levels.

Association analysis: Linear regression models were used to generate cohort-specific residuals of naturally log transformed vitamin D levels adjusted for age, sex, BMI and season defined as summer (June-August), fall (September-November), winter (December to February) and spring (March to May) standardized to have mean 0 and variance of 1. Association between the additively coded SNP genotypes and the vitamin D residuals standardized was assessed using linear regression models. For imputed SNPs, expected number of minor alleles (i.e. dosage) was used in assessing association with the vitamin D residuals.

HERITAGE (Health, Risk Factors, Exercise Training and Genetics): The HERITAGE is the only known family-based study of exercise intervention to evaluate the role of genes and sequence variants involved in the response to a physically active lifestyle. The current study is based on the data collected at baseline of the study from 99 White families (244 males, 255 females). All subjects were required to be sedentary and free of chronic diseases at baseline. There are over 18 trait domains (e.g. dietary, lipids and lipoproteins, glucose and insulin metabolism [fasting and IVGTT], steroids, body composition and body fat distribution, cardiorespiratory fitness), for a grand total of over one thousand variables. Moreover, most of the outcome traits were measured twice on two separate days both at baseline and after exercise training was completed. Marker data include a genome-wide linkage scan and GWAS, in addition to a large number of candidate genes.

HUFS (Howard University Family Study): HUFS followed a population-based selection strategy designed to be representative of African American families living in the Washington, DC metropolitan area. The major objectives of the HUFS were to study the genetic and environmental basis of common complex diseases including hypertension, obesity and associated phenotypes. Participants were sought through door-to-door canvassing, advertisements in local print media and at health fairs and other community gatherings. In order to maximize the utility of this cohort for the study of multiple

common traits, families were not ascertained based on any phenotype. During a clinical examination, demographic information was collected by interview.

HyperGEN (Hypertension Genetic Epidemiology Network): HyperGEN is a family-based study that looks at the genetic causes of hypertension and related conditions in EA and AA subjects. HyperGEN recruited hypertensive sibships, along with their normotensive adult offspring, and an age-matched random sample. HyperGEN has collected data on 2,471 Caucasian-American subjects and 2,300 African-American subjects, from five field centers in Alabama, Massachusetts, Minnesota, North Carolina, and Utah.

IGMM (Institute of Genetics and Molecular Medicine): IGMM oversees three participating studies: CROATIA-Korcula; CROATIA-Vis; GS:SFHS (Generation Scotland: Scottish Family Health Study. **CROATIA-Korcula:** The CROATIA-Korcula study is a family-based, cross-sectional study in the isolated island of Korcula that included 965 examinees aged 18-95. Blood samples were collected in 2007 along with many clinical and biochemical measures and lifestyle and health questionnaires. **CROATIA-Vis:** The CROATIA-Vis study is a family-based, cross-sectional study in the isolated island of Vis that included 1,056 examinees aged 8-93. Blood samples were collected in 2003 and 2004 along with many clinical and biochemical measures and lifestyle and health questionnaires. **GS:SFHS:** The Generation Scotland (www.generationscotland.org) Scottish Family Health Study (GS:SFHS) is a family-based genetic epidemiology cohort with DNA, other biological samples (serum, urine and cryopreserved whole blood) and socio-demographic and clinical data from approximately 24,000 volunteers, aged 18-98 years, in ~7,000 family groups. An important feature of GS:SFHS is the breadth of phenotype information, including detailed data on cognitive function, personality traits and mental health. Although data collection was cross-sectional, GS:SFHS becomes a longitudinal cohort as a result of the ability to link to routine NHS data, using the community health index (CHI) number.

JHS (Jackson Heart Study): The Jackson Heart Study is a longitudinal, community-based observational cohort study investigating the role of environmental and genetic factors in the development of cardiovascular disease in African Americans. Between 2000 and 2004, a total of 5301 participants were recruited from a tri-county area (Hinds, Madison, and Rankin Counties) that encompasses Jackson, MS. Details of the design and recruitment for the Jackson Heart Study cohort has been previously published.1-3 Briefly, approximately 30% of participants were former members of the Atherosclerosis Risk in Communities (ARIC) study. The remainder were recruited by either 1) random selection from the Accudata list, 2) commercial listing, 3) a constrained volunteer sample, in which recruitment was distributed among defined demographic cells in proportions designed to mirror those in the overall population, or through the Jackson Heart Study Family Study.

1. Wyatt SB, Diekelmann N, Henderson F, Andrew ME, Billingsley G, Felder SH et al. A community-driven model of research participation: the Jackson Heart Study Participant Recruitment and Retention Study. Ethn Dis 2003; 13(4):438-455.

2. Taylor HA, Jr., Wilson JG, Jones DW, et al. Toward resolution of cardiovascular health disparities in African Americans: design and methods of the Jackson Heart Study. Ethn Dis 2005; 15:S6-17.

3. Fuqua SR, Wyatt SB, Andrew ME, et al. Recruiting African-American research participation in the Jackson Heart Study: methods, response rates, and sample description. Ethn Dis 2005; 15:S6-29.

Maywood-Loyola Study: Participants were self-identified African Americans from a working class suburb of Chicago, Illinois, USA who were enrolled in studies of BP at the Loyola University Medical Center in Maywood, Illinois, USA as part of the International Collaborative Study on Hypertension in Blacks (ICSHIB) which is described in detail elsewhere **(PMID: 9103091)**. Briefly, nuclear families were

identified through middle-aged probands who were not ascertained based on any phenotype. Thereafter all available first-degree relatives 18 years old and above were enrolled into the study cohort of families. A screening exam was completed by trained and certified research staff using a standardized protocol **(PMID: 9103091 & 10234089)**. Information was obtained on medical history, age, body weight and height. Protocols were reviewed and approved by the IRB at the Loyola University Chicago Stritch School of Medicine prior to recruitment activities. This present study included unrelated adults sampled and for whom information on anthropometrics, BP and use of antihypertensive medication was available. BP measurements were obtained using an oscillometric device, previously evaluated in our field settings **(PMID: 10234089)**. Three measurements were taken three minutes apart and the average of the final two was used in the analysis. Individuals with SBP ≥140 mmHg, DBP ≥90 mmHg or on anti-hypertensive medication at time of exam were defined as hypertensive. Participants with hypertension were offered treatment after detection at the screening exam.

Maywood-Nigeria Study: The sampling frame for the Nigeria cohort was also provided by the International Collaborative Study on Hypertension in Blacks (ICSHIB) as described in detail elsewhere **(PMID: 9103091)**. Study participants were recruited from Igbo-Ora and Ibadan in southwest Nigeria as part of a long-term study on the environmental and genetic factors underlying hypertension. The base cohort consists of over 15,000 participants with information available on anthropometrics, BP and use of antihypertensive medication. BP measurements followed the same protocol described in the Loyola-Maywood study. This present study included unrelated adults samples from the cohort and some hypertensive participants who were recruited as controls in the Africa-America Diabetes Mellitus (AADM) Study recruited from Ibadan in similar neighborhoods **(PMID: 11164120)**. Both projects were reviewed and approved by the sponsoring US institutions (Loyola University Chicago and Howard University) and the University of Ibadan. All participants signed informed consent administered in either English or Yoruba. BP measurements were obtained using an oscillometric device, previously evaluated in our field settings **(PMID: 10234089)**. Three measurements were taken three minutes apart and the average of the final two was used in the analysis. Individuals with SBP ≥140 mmHg, DBP ≥90 mmHg or on anti-hypertensive medication at time of exam were defined as hypertensive. Participants with hypertension were offered treatment after detection at the screening exam.

MESA (Multi-Ethnic Study of Atherosclerosis): The Multi-Ethnic Study of Atherosclerosis (MESA) is a study of the characteristics of subclinical cardiovascular disease and the risk factors that predict progression to clinically overt cardiovascular disease or progression of the subclinical disease. MESA consisted of a diverse, population-based sample of an initial 6,814 asymptomatic men and women aged 45-84. 38 percent of the recruited participants were white, 28 percent African American, 22 percent Hispanic, and 12 percent Asian, predominantly of Chinese descent. Participants were recruited from six field centers across the United States: Wake Forest University, Columbia University, Johns Hopkins University, University of Minnesota, Northwestern University and University of California - Los Angeles. Participants are being followed for identification and characterization of cardiovascular disease events, including acute myocardial infarction and other forms of coronary heart disease (CHD), stroke, and congestive heart failure; for cardiovascular disease interventions; and for mortality. The first examination took place over two years, from July 2000 - July 2002. It was followed by four examination periods that were 17-20 months in length. Participants have been contacted every 9 to 12 months throughout the study to assess clinical morbidity and mortality.

Bild DE, Bluemke DA, Burke GL, Detrano R, Diez Roux AV, Folsom AR, Greenland P, Jacob DR Jr, Kronmal R, Liu K, Nelson JC, O'Leary D, Saad MF, Shea S, Szklo M, Tracy RP. Multi-ethnic study of atherosclerosis: objectives and design. Am J Epidemiol. 2002 Nov 1;156(9):871-81. PubMed PMID: 12397006.

NEO (The Netherlands Epidemiology of Obesity study): The NEO was designed for extensive phenotyping to investigate pathways that lead to obesity-related diseases. The NEO study is a population-based, prospective cohort study that includes 6,671 individuals aged 45–65 years, with an

oversampling of individuals with overweight or obesity. At baseline, information on demography, lifestyle, and medical history have been collected by questionnaires. In addition, samples of 24-h urine, fasting and postprandial blood plasma and serum, and DNA were collected. Genotyping was performed using the Illumina HumanCoreExome chip, which was subsequently imputed to the 1000 genome reference panel. Participants underwent an extensive physical examination, including anthropometry, electrocardiography, spirometry, and measurement of the carotid artery intima-media thickness by ultrasonography. In random subsamples of participants, magnetic resonance imaging of abdominal fat, pulse wave velocity of the aorta, heart, and brain, magnetic resonance spectroscopy of the liver, indirect calorimetry, dual energy X-ray absorptiometry, or accelerometry measurements were performed. The collection of data started in September 2008 and completed at the end of September 2012. Participants are currently being followed for the incidence of obesity-related diseases and mortality.

Pelotas Birth Cohort Study (The 1982 Pelotas Birth Cohort Study, Brazil): The maternity hospitals in Pelotas, a southern Brazilian city (current population ~330,000), were visited daily in the year of 1982. The 5,914 liveborns whose families lived in the urban area were examined and their mothers interviewed. Information was obtained for more than 99% of the livebirths. These subjects have been followed-up at the following mean ages: 11.3 months (all children born from January to Abril 1982; n=1457), 19.4 months (entire cohort; n=4934), 43.1 months (entire cohort; n=4742), 13.1 years (random subsample; n=715), 14.7 years (systematic subsample; n=1076); 18.2 (male cohorts attending to compulsory Army recruitment examination; n=2250), 18.9 (systematic subsample; n=1031), 22.8 years (entire cohort; n=4297) and 30.2 years (entire cohort; n=3701). Details about follow-up visits and available data can be found in the two Cohort Profile papers (PMID: 16373375 and 25733577). DNA samples (collected at the mean age of 22.8 years) were genotyped for ~2.5 million of SNPs using the Illumina HumanOmni2.5-8v1 array (which includes autosomal, X and Y chromosomes, and mitochondrial variants). After quality control, the data were prephased using SHAPEIT and imputed using IMPUTE2 based on 1000 Genomes haplotypes.

RS (Rotterdam Study): The Rotterdam Study is a prospective, population-based cohort study among individuals living in the well-defined Ommoord district in the city of Rotterdam in The Netherlands. The aim of the study is to determine the occurrence of cardiovascular, neurological, ophthalmic, endocrine, hepatic, respiratory, and psychiatric diseases in elderly people. The cohort was initially defined in 1990 among approximately 7,900 persons, aged 55 years and older, who underwent a home interview and extensive physical examination at the baseline and during follow-up rounds every 3-4 years (RS-I). Cohort was extended in 2000/2001 (RS-II, 3,011 individuals aged 55 years and older) and 2006/2008 (RS-III, 3,932 subjects, aged 45 and older). Written informed consent was obtained from all participants and the Medical Ethics Committee of the Erasmus Medical Center, Rotterdam, approved the study.

SCHS-CHD (Singapore Chinese Health Study - Coronary Heart Disease): SCHS-CHD is a case-control study of coronary heart disease that was nested within the Singapore Chinese Health Study (SCHS), a prospective cohort study of 63,257 Singaporean Chinese men and women aged 45-74 years living in Singapore. We selected cases and controls from participants that provided blood samples and were free of coronary heart disease and stroke at the time of blood collection (N=24,454). Cases (N=760) had acute myocardial infarction (AMI) or died of coronary heart disease. AMI was identified through the Singapore Myocardial Infarction Registry or through the nationwide hospital discharge database followed by confirmation of AMI by cardiologists' review of medical records using the Multi-Ethnic Study of Atherosclerosis criteria (available at: http://www.mesa-nhlbi.org/manuals.aspx). Coronary heart disease deaths were identified through the Singapore Registry of Births and Deaths (ICD9 410-414 as first stated cause of death). Matched controls (N=1,491) were selected using a risk-set sampling strategy. Controls were participants who were alive and free of coronary heart disease at the time of the diagnosis or death of the index cases and were matched for age, sex, dialect group, year of recruitment and date of blood collection. In-person interviews and phlebotomy were conducted

before the onset of disease and non-fasting venous blood was stored at -80^{0}C for extraction of DNA and blood biochemistry.

Singapore: SCES (Singapore Chinese Eye Study): SCES is a population-based, cross-sectional study of Chinese adults aged 40–80+ years residing in the South-Western part of Singapore, which is part of the Singapore Epidemiology of Eye Disease (SEED). Age stratified random sampling was used to select 6,350 eligible participants, of which 3,300 participated in the study (73% response rate). Detailed methodology has been published. Two readings of blood pressure were taken from participants after 5 minutes of rest, seated, using an automated blood pressure monitor (Dinamap Pro100V2; Criticon, Norderstedt, Germany) by trained observers. One of two cuff sizes (regular, large) was chosen on the basis of the circumference of the participant's arm. A third reading was performed if the difference between two readings of either the systolic blood pressure was greater than 10mmHg or the diastolic blood pressure was greater than 5mmHg. The mean values of the closest two readings were calculated. **SiMES (Singapore Malay Eye Study):** SiMES is a population-based cross-sectional epidemiological study of 3,280 individuals from one of the three major ethnic groups residing in Singapore. SiMES is part of the Singapore Epidemiology of Eye Disease (SEED) study. In summary, 5,600 individuals have been selected by an age-stratified sampling strategy. Among these 4,168 individuals are eligible for this study. 3,280 individuals finally participated in the study. All subjects were Malay and aged 40-79 years. Two readings of blood pressure were taken from participants after 5 minutes of rest, seated, using an automated blood pressure monitor (Dinamap Pro100V2; Criticon, Norderstedt, Germany) by trained observers. One of two cuff sizes (regular, large) was chosen on the basis of the circumference of the participant's arm. A third reading was performed if the difference between two readings of either the systolic blood pressure was greater than 10mmHg or the diastolic blood pressure was greater than 5mmHg. The mean values of the closest two readings were calculated. **SINDI (Singapore Indian Eye Study):** is a population-based, cross-sectional study of Asian Indian adults aged 40–80+ years residing in the South-Western part of Singapore, which is part of the Singapore Epidemiology of Eye Disease (SEED). Age stratified random sampling was used to select 6,350 eligible participants, of which 3,400 participated in the study (75.6% response rate). Detailed methodology has been published. Two readings of blood pressure were taken from participants after 5 minutes of rest, seated, using an automated blood pressure monitor (Dinamap Pro100V2; Criticon, Norderstedt, Germany) by trained observers. One of two cuff sizes (regular, large) was chosen on the basis of the circumference of the participant's arm. A third reading was performed if the difference between two readings of either the systolic blood pressure was greater than 10mmHg or the diastolic blood pressure was greater than 5mmHg. The mean values of the closest two readings were calculated. **SP2 (Singapore 2):** The SP2 is a population-based study of diabetes and cardiovascular disease in Singapore. It first surveyed subjects (Chinese, Malay and Indian) from four cross-sectional studies that were conducted in Singapore between 1982 and 1998. Subjects were between the ages of 24-95 years and represented a random sample of the Singapore population. Subjects were re-visited between 2003 and 2007. Among the 10,747 individuals who were eligible, 5,157 subjects completed a questionnaire and the subsequent clinical examinations. Data from this re-visit were utilized for this study. Two readings of blood pressure were taken from participants after 5 min of rest, seated, using an automated blood pressure monitor (Dinamap Pro100V2; Criticon, Norderstedt, Germany) by trained observers. One of two cuff sizes (regular, large) was chosen on the basis of the circumference of the participant's arm. A third reading was performed if the difference between two readings of either the systolic blood pressure was greater than 10mmHg or the diastolic blood pressure was greater than 5mmHg. The mean values of the closest two readings were calculated.

WGHS (Women's Genome Health Study): WGHS is a prospective cohort of female North American health care professionals representing participants in the Women's Health Study (WHS) trial who provided a blood sample at baseline and consent for blood-based analyses. Participants in the WHS were 45 years or older at enrollment and free of cardiovascular disease, cancer or other major chronic illness. The current data are derived from 23,294 WGHS participants for whom whole genome

genotype information was available at the time of analysis and for whom self-reported European ancestry could be confirmed by multidimensional scaling analysis of 1,443 ancestry informative markers in PLINK v. 1.06. At baseline, BP and lifestyle habits related to smoking, consumption of alcohol, and physical activity as well as other general clinical information were ascertained by a self-reported questionnaire, an approach which has been validated in the WGHS demographic, namely female health care professionals. Questionnaires recorded systolic BP in 9 categories (<110, 110-119, 120-129, 130-139, 140-149, 150-159, 160-169, 170-179, ≥180 mmHg), and diastolic BP in 7 categories (<65, 65-74, 75-84, 85-89, 90-94, 95-104, ≥105 mmHg). All analyses treated these BP responses as quantitative variables representing each category with its midpoint value. Hypertension was defined as one or more of reported physician diagnosis, systolic BP ≥140 mmHg, or diastolic BP ≥90 mmHg.

WHI (Women's Health Initiative): WHI is a long-term national health study that focuses on strategies for preventing common diseases such as heart disease, cancer and fracture in postmenopausal women. A total of 161,838 women aged 50–79 years old were recruited from 40 clinical centers in the US between 1993 and 1998. WHI consists of an observational study, two clinical trials of postmenopausal hormone therapy (HT, estrogen alone or estrogen plus progestin), a calcium and vitamin D supplement trial, and a dietary modification trial [1]. Study recruitment and exclusion criteria have been described previously [1]. Recruitment was done through mass mailing to age-eligible women obtained from voter registration, driver's license and Health Care Financing Administration or other insurance list, with emphasis on recruitment of minorities and older women [2]. Exclusions included participation in other randomized trials, predicted survival < 3 years, alcoholism, drug dependency, mental illness and dementia. For the CT, women were ineligible if they had a systolic BP > 200 mm Hg or diastolic BP > 105 mm Hg, a history of hypertriglyceridemia or breast cancer. Study protocols and consent forms were approved by the IRB at all participating institutions. Socio-demographic characteristics, lifestyle, medical history and self-reported medications were collected using standardized questionnaires at the screening visit. Physical measures of height, weight and blood pressure were measured at a baseline clinical visit [2]. BP was measured by certified staff using standardized procedures and instruments [3]. Two BP measures were recorded after 5 minutes rest using a mercury sphygmomanometer. Appropriate cuff bladder size was determined at each visit based on arm circumference. Diastolic BP was taken from the phase V Korotkoff measures. The average of the two measurements, obtained 30 seconds apart, was used in analyses. The genome wide association study (GWAS) non-overlapping samples are composed of a case-control study (WHI Genomics and Randomized Trials Network – GARNET, which included all coronary heart disease, stroke, venous thromboembolic events and selected diabetes cases that happened during the active intervention phase in the WHI HT clinical trials and aged matched controls), women selected to be "representative" of the HT trial (mostly younger white HT subjects that were also enrolled in the WHI memory study - WHIMS) and the WHI SNP Health Association Resource (WHI SHARe), a randomly selected sample of 8,515 African American and 3,642 Hispanic women from WHI. GWAS was performed using Affymetrix 6.0 (WHI-SHARe), HumanOmniExpressExome-8v1_B (WHIMS), Illumina HumanOmni1-Quad v1-0 B (GARNET). Extensive quality control (QC) of the GWAS data included alignment ("flipping") to the same reference panel, imputation to the 1000G data (using the recent reference panel - v3.20101123), identification of genetically related individuals, and computations of principal components (PCs) using methods developed by Price et al. (using EIGENSOFT software 53), and finally the comparison with self-reported ethnicity. After QC and exclusions from analysis protocol, the number of women included in analysis is 4,423 whites for GARNET, 5,202 white for WHIMS, 7,919 for SHARe African American and 3,377 for SHARe Hispanics.

1. Hays J, Hunt JR, Hubbell FA, Anderson GL, Limacher M, Allen C, Rossouw JE. The women's health initiative recruitment methods and results. Ann Epidemiol. 2003;13:S18-77

2. Design of the women's health initiative clinical trial and observational study. The women's health initiative study group. Control Clin Trials. 1998;19:61-109

3. Hsia J, Margolis KL, Eaton CB, Wenger NK, Allison M, Wu L, LaCroix AZ, Black HR. Prehypertension and cardiovascular disease risk in the women's health initiative. Circulation. 2007;115:855-860

Stage 2 (Focused Discovery/replication) Study Descriptions

Brief descriptions are provided below for each of the replication studies/cohorts:

AA-DHS (African American Diabetes Heart Study): AA-DHS objectives are to improve understanding of ethnic differences in CAC and CP in populations of African and European ancestry. The AA-DHS consists of self-reported African Americans with T2D recruited from two Wake Forest School of Medicine (WFSM) studies: the family-based Diabetes Heart Study (DHS) and unrelated individuals in the AA-DHS. DHS is a cross-sectional study of European American and African American families with siblings concordant for T2D. AA-DHS started after DHS and enrolled unrelated African Americans. The AA-DHS GWAS utilized the Illumina 5M chip with imputation to 1,000 Genomes.

Airwave (The Airwave Health Monitoring Study): The Airwave Health Monitoring Study (22) was established to evaluate possible health risks associated with use of TETRA, a digital communication system used by police forces and other emergency services in Great Britain since 2001. The study has been broadened to investigate more generally the health of the work force. From 2004, participants from each force who agreed to participate were enrolled either with an enrolment questionnaire or a comprehensive health screening performed locally. This includes questionnaire, 7-day food diaries, anthropometry, measurements of cardiovascular and cognitive function, blood chemistry, coagulation and hematology. Systolic and diastolic blood pressures were measured as three consecutive readings using a digital blood pressure monitor (Omron HEM 705-CP digital BP monitor). By March 2015, the study had recruited 53,606 participants, of whom 45,433 had attended the health screening, and 14,002 have genotype data (1000G imputed).

Ref: Elliott, P. et al. The Airwave Health Monitoring Study of police officers and staff in Great Britain: rationale, design and methods. Environ Res 134, 280-5 (2014).

ASCOT (Anglo-Scandinavian Cardiac Outcomes Trial): ASCOT is a randomised control clinical trial investigating the cardiac outcomes of blood pressure lowering and lipid lowering treatments. Of 19,342 hypertensive patients (40–79 years of age with at least three other cardiovascular risk factors) who were randomized to one of two antihypertensive regimens in ASCOT (atenolol, Beta-Blocker vs amlodipine, Calcium-Channel-Blocker), 10,305 patients with non-fasting total cholesterol concentrations of 6.5 mmol/l or less (measured at the non-fasting screening visit) had been randomly assigned additional atorvastatin 10 mg or placebo. Only a proportion of United Kingdom, Irish, Sweden, Norway, Finland and Denmark consented to contribute DNA and participate in genetic studies. PMID 11685901

BBJ (Biobank Japan Project): The Biobank Japan (BBJ) Project was established in 2003 with the aim of the implementation of personalized medicine as a leading project of Ministry of Education, Culture, Sports, Science and Technology (MEXT). In collaboration with twelve cooperating institutes, the BBJ has recruited a total of 200,000 people, suffering from at least one of the 47 target common diseases, in the first phase (5-year period). BBJ has collected biospecimens including DNA and serum as well as various clinical and lifestyle information through interview or medical records by using standardized questionnaire. All participants gave written informed consent to this project and this study was approved by ethical committees of RIKEN and participating institutes.

BES (Beijing Eye Study): Beijing Eye Study is a population-based study that assess the associated and risk factors of ocular and general diseases in China population. The study was initialized in 2001, collected data from 4439 subjects aged ≥ 40 years from seven communities in Beijing area, where three of the communities were located in rural districts and four were located in urban districts. BES was followed-up in 2006, with 3251 of the original subjects participated, and in 2011, with 2695 subjects returned for the follow-up examination. At the examinations in 2006 and 2011, trained research staffs asked the subjects questions from a standard questionnaire providing information on family status, level

of education, income, quality of life, psychic depression, physical activity, and known major systemic diseases. Fasting blood samples were taken for measurement of blood lipids, glucose, and glycosylated hemoglobin. Individuals were classified as self-reported non-smokers or self-reported current smokers. Alcohol consumption habits based on number of drinks per day were collected. All variables used in analyses were taken from examinations in 2006 or in 2011. The BES subjects were genotyped on two arrays, Illumina Human610-Quad (N = 832) and Illumina OmniExpress (N = 814).

BRIGHT (British Genetics of Hypertension): Participants of the BRIGHT Study are recruited from the Medical Research Council General Practice Framework and other primary care practices in the UK. Each case had a history of hypertension diagnosed prior to 60 years of age with confirmed blood pressure recordings corresponding to seated levels >150/100mmHg (1 reading) or mean of 3 readings >145/95 mmHg. BRIGHT is focused on recruitment of hypertensive individuals with BMI<30. Sample selection for GWAS was based on DNA availability and quantity. PMID 12826435

CAGE-Amagasaki (Cardio-metabolic Genome Epidemiology Network, Amagasaki Study): The Amagasaki Study (CAGE-Amagasaki) is an ongoing population-based cohort study of 5,743 individuals (3,435 males and 2,310 females), aged >18 years and recruited for a baseline examination between September 2002 to August 2003. Participants were interviewed by trained personnel to obtain information on medical and lifestyle variables, and consented to provide DNA for genotyping of molecular variants to investigate genetic susceptibility for so-called lifestyle-related diseases such as hypertension and cardiovascular disorder.

CFS (Cleveland Family Study): The Cleveland Family Study (CFS) is a family-based, longitudinal study designed to characterize the genetic and non-genetic risk factors for sleep apnea. In total, 2534 individuals (46% African American) from 352 families were studied on up to 4 occasions over a period of 16 years (1990-2006). The initial aim of the study was to quantify the familial aggregation of sleep apnea. 632 African Americans were genotyped on the Affymetrix array 6.0 platform through the CARe Consortium with suitable genotying quality control. A further 122 African-Americans had genotyping based on the Illumina OmniExpress + Exome platform. Genomes were imputed separately for each chip based on a 1000 Genomes Project Phase 3 Version 5 cosmopolitan template using SHAPEIT and IMPUTE2. Participants had three supine BP measurements each performed after lying quietly for 10 minutes, before bed (10:00 P.M.) and upon awakening (7:00 A.M.), and another three sitting at 11 am, following standardized guidelines using a calibrated sphygmomanometer. Cuff size was determined by the circumference of the upper arm and the appropriate bladder size from a standard chart. BP phenotypes were determined from the average of the nine measurements.

Colaus (Cohorte Lausannoise): The cohort is a random population sample of the city of Lausanne aged 35-75 years. Recruitment began in June 2003 and ended in May 2006, and the first follow-up was conducted between April 2009 and September 2012. The CoLaus study was approved by the Institutional Ethics Committee of the University of Lausanne and informed consent was appropriately obtained by all participants. Both at baseline and follow-up, all participants attended the outpatient clinic of the University Hospital of Lausanne in the morning after an overnight fast. Data were collected by trained field interviewers in a single visit lasting about 60 min.

DESIR (Data from an Epidemiological Study on the Insulin Resistance): The DESIR cohort study aims to: describe and understand the relations between the abnormalities of the syndrome, their evolution, according to age and sex; search for risk factors of insulin resistance, in particular factors associated with the environment, lifestyle and genetic markers; quantify the links between the syndrome and both cardiovascular disease and diabetes; evaluate the frequency of the syndrome in terms of its consequences on public health.

DFTJ (Dongfeng-Tongji Cohort Study): The DFTJ-cohort study includes 27,009 retired employees from a state-owned automobile enterprise in China. This study was launched in 2008 and will be

followed up every 5 years. In 2013 we conducted the first follow-up. By using semi-structural questionnaire and health examination, those having cancer or severe diseases were excluded. Fasting blood samples and detailed epidemiology data were collected. The main goal of the cohort was to identify the environmental and genetic risk factors and the gene-environment interactions on chronic diseases, and to find novel biomarkers for chronic disease and mortality prediction. Finally, 1,461 included in the present study with GWAS data. All of the participants wrote informed consent and the ethical committees in the Tongji Medical College approved this research project. Detailed information has been described in elsewhere(1).

QC criteria and imputation methods:
We did the GWAS scan on the DFTJ-cohort with Affymetrix Genome-Wide Human SNP Array 6.0 chips. In total, we genotyped 906,703 SNPs among 1,461 subjects. After stringent QC filtering, SNPs with MAF < 0.01, Hardy-Weinberg Equilibrium (HWE) < 0.0001, and SNP call rate < 95% were excluded. Individuals with call rates < 95% were also not included for further analysis. In total, we retained 1,452 subjects with 658,288 autosomal SNPs for statistical analyses, with an overall call rate of 99.68%. We used MACH 1.0 software to impute untyped SNPs using the LD information from the HapMap phase II database (CHB+JPT as a reference set (2007-08_rel22, released 2007-03-02). Imputed SNPs with high genotype information content (Rsq > 0.3 for MACH) were kept for the further association analysis.

Reference
1) Wang, F., Zhu, J., Yao, P., Li, X., He, M., Liu, Y., Yuan, J., Chen, W., Zhou, L., Min, X. et al. (2012) Cohort profile: The Dongfeng-Tongji cohort study of retired workers. International journal of epidemiology.

DHS (Diabetes Heart Study): The Diabetes Heart Study (DHS) is an ongoing family-based cohort study investigating the epidemiology and genetics of cardiovascular disease (CVD) in a population-based sample. The DHS recruited T2D-affected siblings without advanced renal insufficiency from 1998 through 2005 in western North Carolina. DHS has collected genetic data on 1,220 self-described European American (EA) individuals from 475 families. Genotyping was completed using an Affymetrix Genome-Wide Human SNP Array 5.0 with imputation of 1,000 Genomes project SNPs from this array using IMPUTE2 and the Phase I v2, cosmopolitan (integrated) reference panel, build 37.

DR's EXTRA (Dose Responses to Exercise Training): The Dose-Responses to Exercise Training (DR's EXTRA) Study is a 4-year RCT on the effects of regular physical exercise and healthy diet on endothelial function, atherosclerosis and cognition in a randomly selected population sample (n=3000) of Eastern Finnish men and women, identified from the national population register, aged 55-74 years. Of the eligible sample, 1410 individuals were randomized into one of the 6 groups: aerobic exercise, resistance exercise, diet, combined aerobic exercise and diet, combined resistance exercise and diet, or reference group following baseline assessments. During the four year intervention the drop-out rate was 15%.

EGCUT (Estonian Genome Center - University of Tartu (Estonian Biobank)): The Estonian Biobank is the population-based biobank of the Estonian Genome Center at the University of Tartu (www.biobank.ee; EGCUT). The entire project is conducted according to the Estonian Gene Research Act and all of the participants have signed the broad informed consent. The cohort size is up to 51535 individuals from 18 years of age and up, which closely reflects the age, sex and geographical distribution of the Estonian population. All of the subjects are recruited randomly by general practitioners and physicians in hospitals. A Computer Assisted Personal interview is filled within 1-2 hours at a doctor's office, which includes personal, genealogical, educational, occupational history and lifestyle data. Anthropometric measurements, blood pressure and resting heart rate are measured and venous blood taken during the visit. Medical history and current health status is recorded according to ICD-10 codes.

EPIC (European Prospective Investigation into Cancer and Nutrition): The European Prospective Investigation of Cancer (EPIC) began as a large multi-centre cohort study primarily looking at the connection between diet, lifestyle factors and cancer, although the study was broadened from the outset to include other conditions. The EPIC-Norfolk participants are men and women who were aged between 40 and 79 when they joined the study and who lived in Norwich and the surrounding towns and rural areas. They have been contributing information about their diet, lifestyle and health through questionnaires and health checks over two decades. The Norwich Local Research Ethics Committee granted ethical approval for the study. All participants gave written informed consent.

FENLAND (The Fenland Study): The Fenland study is a population-based cohort study that uses objective measures of disease exposure to investigate the influence of diet, lifestyle and genetic factors on the development of diabetes and obesity. The volunteers are recruited from general practice lists in and around Cambridgeshire (Cambridge, Ely, and Wisbech) in the United Kingdom from birth cohorts from 1950–1975.

FUSION (Finland-United States Investigation of NIDDM Genetics): The Finland-United States Investigation of NIDDM Genetics (FUSION) study is a long-term effort to identify genetic variants that predispose to type 2 diabetes (T2D) or that impact the variability of T2D-related quantitative traits. The FUSION GWAS sample consists of 1,161 Finnish T2D cases and 1,174 Finnish normal glucose-tolerant (NGT) controls (Scott et al. Science 2007). Cases are defined by fasting plasma glucose ≥ 7.0 mmol/l or 2-h plasma glucose ≥ 11.1 mmol/l, by report of diabetes medication use, or based on medical record review. 789 FUSION cases each reported at least one T2D sibling; 372 Finrisk 2002 T2D cases came from a Finnish population-based risk factor survey. NGT controls are defined by fasting glucose < 6.1 mmol/l and 2-h glucose < 7.8 mmol/l. FUSION controls include 119 subjects from Vantaa, Finland who were NGT at ages 65 and 70 years, 304 NGT spouses from FUSION families, and 651 Finrisk 2002 subjects. The controls were approximately frequency matched to the cases by age, sex, and birth province. Smoking and alcohol data are only available in the FUSION subset of our GWAS samples.

Scott, L.J. et al. A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. Science 316, 1341–1345, 2007.

GeneSTAR (Genetic Studies of Atherosclerosis Risk): GeneSTAR is a family-based prospective study of more than 4000 participants begun in 1983 to determine phenotypic and genetic causes of premature cardiovascular disease. Families were identified from 1983-2006 from probands with a premature coronary disease event prior to 60 years of age who were identified at the time of hospitalization in any of 10 hospitals in the Baltimore, Maryland area. Their apparently healthy 30-59 year old siblings without known coronary disease were recruited and screened between 1983 and 2006. From 2003-2006, adult offspring over 21 years of age of all participating siblings and probands, as well as the coparents of the offspring were recruited and screened. Genotyping was performed in 3,232 participants on the Illumina 1Mv1_c platform.

GLACIER (Gene x Lifestyle Interactions and Complex Traits Involved in Elevated Disease Risk): The Gene-Lifestyle interactions And Complex traits Involved in Elevated disease Risk (GLACIER) Study is nested within the Västerbotten Health Survey, which is part of the Northern Sweden Health and Disease Study, a population-based prospective cohort study from northern Sweden. Participants were genotyped with Illumina CardioMetaboChip array. This array contains ~200,000 variants, the majority being common variants. Systolic and diastolic blood pressures were measured once following a period of five minutes rest with the participant in the supine position using a mercury-gauge sphygmomanometer. Analysis of serum lipids (HDL-C, triglycerides and total cholesterol) were undertaken at the Department of Clinical Chemistry at Umeå University Hospital using routine methods. LDL-C was determined using the Friedewald formula. All participants completed a detailed, optically

readable, health and lifestyle questionnaire including questions about smoking status and alcohol intake (FFQ). Cohort description - PMID: 25396097

GRAPHIC (Genetic Regulation of Arterial Pressure of Humans in the Community): The GRAPHIC Study comprises 2024 individuals from 520 nuclear families recruited from the general population in Leicestershire, UK between 2003-2005 for the purpose of investigating the genetic determinants of blood pressure and related cardiovascular traits. A detailed medical history was obtained from study subjects by standardized questionnaires and clinical examination was performed by research nurses following standard procedures. Measurements obtained included height, weight, waist-hip ratio, clinic and ambulatory blood pressure and a 12-lead ECG.

HCHS/SOL (Hispanic Community Health Study/ Study of Latinos): The HCHS/SOL is a community-based cohort study of 16,415 self-identified Hispanic/Latino persons aged 18–74 years and selected from households in predefined census-block groups across four US field centers (in Chicago, Miami, the Bronx, and San Diego). The census-block groups were chosen to provide diversity among cohort participants with regard to socioeconomic status and national origin or background. The HCHS/SOL cohort includes participants who self-identified as having a Hispanic/Latino background; the largest groups are Central American (n = 1,730), Cuban (n = 2,348), Dominican (n = 1,460), Mexican (n = 6,471), Puerto Rican (n = 2,728), and South American (n = 1,068). The HCHS/SOL baseline clinical examination occurred between 2008 and 2011 and included comprehensive biological, behavioral, and sociodemographic assessments. Consenting HCHS/SOL subjects were genotyped at Illumina on the HCHS/SOL custom 15041502 B3 array. The custom array comprised the Illumina Omni 2.5M array (HumanOmni2.5-8v.1-1) ancestry-informative markers, known GWAS hits and drug absorption, distribution, metabolism, and excretion (ADME) markers, and additional custom content including ~150,000 SNPs selected from the CLM (Colombian in Medellin, Colombia), MXL (Mexican Ancestry in Los Angeles, California), and PUR (Puerto Rican in Puerto Rico) samples in the 1000Genomes phase 1 data to capture a greater amount of Amerindian genetic variation. QA/QC procedures yielded a total of 12,803 unique study participants for imputation and downstream association analyses.

HRS (Health & Retirement Study): The Health and Retirement Study (HRS) is a longitudinal survey of a representative sample of Americans over the age of 50. The current sample is over 26,000 persons in 17,000 households. Respondents are interviewed every two years about income and wealth, health and use of health services, work and retirement, and family connections. DNA was extracted from saliva collected during a face-to-face interview in the respondents' homes. These data represent respondents who provided DNA samples and signed consent forms in 2006, 2008, and 2010. Respondents were removed if they had missing genotype or phenotype data.

Juster, F. T., Suzman, R. (1995). An Overview of the Health and Retirement Study, Journal of Human Resources 30:Suppl: S7-S56.

Sonnega A, Faul JD, Ofstedal MB, Langa KM, Phillips JWR, Weir DR. Cohort Profile: the Health and Retirement Study (HRS). Int. J. Epidemiol. 2014; 43 (2): 576-585. PMID: 24671021

Crimmins, E.M., Guyer H., Langa K.M., Ofstedal M.B., Wallace R.B., and Weir D.R. (2008). Documentation of Physical Measures, Anthropometrics and Blood Pressure in the Health and Retirement Study. HRS Documentation Report DR-011. http://hrsonline.isr.umich.edu/sitedocs/userg/dr-011.pdf

HyperGEN-AXIOM (Hypertension Genetic Epidemiology Network): HyperGEN is a family-based study that investigates the genetic causes of hypertension and related conditions in EA and AA subjects. HyperGEN recruited hypertensive sibships, along with their normotensive adult offspring, and an age-matched random sample. HyperGEN has collected data on 2,471 Caucasian-American subjects and 2,300 African-American subjects, from five field centers in Alabama, Massachusetts, Minnesota,

North Carolina, and Utah. HyperGEN participates as a discovery study using GWAS available in a large subset of the samples. The remaining AA subjects without GWAS data were genotyped on the Affymetrix Axiom chip as part of a HyperGEN admixture mapping ancillary study. After excluding subjects already included in the original HyperGEN (or with family members included), this subset of approximately 450 AA subjects are included in the HyperGEN-AXIOM study which participates in replications.

INGI-CARL (Italian Network Genetic Isolates): The Carlantino cohort (INGI-CARL) is a population-based study including approximately 1000 samples from an isolated village of Southern Italy.

INGI-FVG (Italian Network Genetic Isolates): INGI-FVG is a population-based study including approximately 1700 samples from six isolated villages of Northern Italy.

InterAct (The EPIC-InterAct Case-Cohort Study): The large prospective InterAct type 2 diabetes case-cohort study is coordinated by the MRC Epidemiology Unit in Cambridge and nested within the European Prospective Investigation into Cancer and Nutrition (EPIC). EPIC was initiated in the late 1980s and involves collaboration between 23 research institutions across Europe in 10 countries (Denmark, France, Germany, Greece, Italy, the Netherlands, Norway, Spain, Sweden and the United Kingdom). The majority of EPIC cohorts were recruited from the general population, with some exceptions. French cohorts included women who were members of a health insurance scheme for school and university employees; Turin and Ragusa (Italy) and the Spanish centres included some blood donors. Participants from Utrecht (Netherlands) and Florence (Italy) were recruited via a breast cancer screening program. The majority of participants recruited by the EPIC Oxford (UK) centre consisted of vegetarian and “health conscious” volunteers from England, Wales, Scotland, and Northern Ireland.

IRAS (Insulin Resistance Atherosclerosis Study): The Insulin Resistance Atherosclerosis Study (IRAS) was an epidemiologic cohort study designed to examine the relationship between insulin resistance and carotid atherosclerosis across a range of glucose tolerance. Individuals of self-reported Mexican-American ethnicity were recruited in San Antonio, TX and San Luis Valley, CO. Recruitment was balanced across age and glucose tolerance status. Inclusion of IRAS data is limited to 194 normoglycemic individuals with genotype data from the Illumina OmniExpress and Omni 1S arrays and imputation to the 1000 Genome Integrated Reference Panel (phase I).

IRAS Family Study (Insulin Resistance Atherosclerosis Study): The IRASFS was a family study designed to examine the genetic and epidemiologic basis of glucose homeostasis traits and abdominal adiposity. Briefly, self-reported Mexican pedigrees were recruited in San Antonio, TX and San Luis Valley, CO. Probands with large families were recruited from the initial non-family-based IRAS, which was modestly enriched for impaired glucose tolerance and T2D. Inclusion of IRASFS data is limited to 1040 normoglycemic individuals in 88 pedigrees with genotype data from the Illumina OmniExpress and Omni 1S arrays and imputation to the 1000 Genome Integrated Reference Panel (phase I).

JUPITER (Justification for the Use of Statins in Primary Prevention: An Intervention Trial Evaluating Rosuvastatin): Genetic analysis was performed in a sub-population from JUPITER (Justification for the Use of statins in Prevention: an Intervention Trial Evaluating Rosuvastatin), an international, randomized, placebo-controlled trial of rosuvastatin (20mg/day) in the primary prevention of cardiovascular disease conducted among apparently healthy men and women with LDL-C < 130 mg/dL and hsCRP ≥ 2 mg/L (1, 2). Individuals with diabetes or triglyceride concentration >500mg/dL were excluded. The present analysis includes only individuals who provided consent for genetic analysis, had successfully collected genotype information, and who had either verified European or verified South African black ancestry.

1) Ridker PM, Danielson E, Fonseca FA, Genest J, Gotto AM Jr, Kastelein JJ, Koenig W, Libby P, Lorenzatti AJ, MacFadyen JG, Nordestgaard BG, Shepherd J, Willerson JT, Glynn RJ; JUPITER Study Group. Rosuvastatin to prevent vascular events in men and women with elevated C-reactive protein. N Engl J Med 2008 Nov 20; 359(21):2195-207

2) Chasman DI, Giulianini F, MacFadyen J, Barratt BJ, Nyberg F, Ridker PM. Genetic determinants of statin-induced low-density lipoprotein cholesterol reduction: the Justification for the Use of Statins in Prevention: an Intervention Trial Evaluating Rosuvastatin (JUPITER) trial. Circ Cardiovasc Genet. 2012 Apr 1;5(2):257-64. doi: 10.1161/CIRCGENETICS.111.961144. Epub 2012 Feb 13. Erratum in: Circ Cardiovasc Genet. 2012 Jun;5(3):e27. PubMed PMID: 22331829.

KORA (Cooperative Health Research in the Augsburg Region): The KORA study is a series of independent population-based epidemiological surveys of participants living in the region of Augsburg, Southern Germany. All survey participants are residents of German nationality identified through the registration office and were examined in 1994/95 (KORA S3) and 1999/2001 (KORA S4). In the KORA S3 and S4 studies 4,856 and 4,261 subjects have been examined implying response rates of 75% and 67%, respectively. 3,006 subjects participated in a 10-year follow-up examination of S3 in 2004/05 (KORA F3), and 3080 of S4 in 2006/2008 (KORA F4). The age range of the participants was 25 to 74 years at recruitment. Informed consent has been given by all participants. The study has been approved by the local ethics committee. Individuals for genotyping in KORA F3 and KORA F4 were randomly selected and these genotypes are taken for the analysis of the phenotypes in KORA S3 and KORA S4.

LBC1921 (Lothian Birth Cohort 1921): LBC1921 consists of 550 (234 male) relatively healthy individuals, assessed on cognitive and medical traits at a mean age of 79.1 years (SD = 0.6). They were born in 1921, most took part in the Scottish Mental Survey of 1932, and almost all lived independently in the Lothian region (Edinburgh City and surrounding area) of Scotland.[1]

LBC1936 (Lothian Birth Cohort 1936): LBC1936 consists of 1091 (548 male) relatively healthy individuals who underwent cognitive and medical testing at a mean age of 69.6 years (SD = 0.8). They were born in 1936, most took part in the Scottish Mental Survey of 1947, and almost all lived independently in the Lothian region of Scotland.[1]

(1) Deary IJ, Gow AJ, Pattie A, Starr JM. Cohort profile: the Lothian Birth Cohorts of 1921 and 1936. Int J Epidemiol 2012;41:1576-1584.

Lifelines (Netherlands Biobank): Lifelines (https://lifelines.nl/) is a multi-disciplinary prospective population-based cohort study using a unique three-generation design to examine the health and health-related behaviors of 165,000 persons living in the North East region of The Netherlands. It employs a broad range of investigative procedures in assessing the biomedical, socio-demographic, behavioral, physical and psychological factors which contribute to the health and disease of the general population, with a special focus on multimorbidity. In addition, the Lifelines project comprises a number of cross-sectional sub-studies which investigate specific age-related conditions. These include investigations into metabolic and hormonal diseases, including obesity, cardiovascular and renal diseases, pulmonary diseases and allergy, cognitive function and depression, and musculoskeletal conditions. All survey participants are between 18 and 90 years old at the time of enrollment. Recruitment has been going on since the end of 2006, and over 130,000 participants had been included by April 2013. At the baseline examination, the participants in the study were asked to fill in a questionnaire (on paper or online) before the first visit. During the first and second visit, the first or second part of the questionnaire, respectively, are checked for completeness, a number of investigations are conducted, and blood and urine samples are taken. Lifelines is a facility that is open for all researchers. Information on application and data access procedure is summarized on www.lifelines.nl.

Scholtens S, Smidt N, Swertz MA, Bakker SJ, Dotinga A, Vonk JM, et al. Cohort Profile: LifeLines, a three-generation cohort study and biobank. Int J Epidemiol. 2014 Dec 14.

LLFS (The Long Life Family Study): LLFS is a family-based cohort study, including four clinical centers: Boston University Medical Center in Boston, MA, USA, Columbia College of Physicians and Surgeons in New York City, NY, USA, the University of Pittsburgh in Pittsburgh PA, USA, and University of Southern Denmark, Denmark. The study characteristics, recruitment, eligibility and enrollment have been previously described (Pedersen et al., 2006, PMID: 17150149; Sebastiani et al., 2009, PMID: 19910380; Newman et al., 2011, PMID: 21258136). In brief, the LLFS was designed to determine genetic, behavioral, and environmental factors related to families of exceptionally healthy, elderly individuals. Phase 1 was conducted between 2006 and 2009 recruiting 4,953 individuals from 539 families. The probands were at least 79 years old in the USA centers, and 90 years old or above in Denmark. The families were selected to participate in the study based on The Family Longevity Selection Score (FLoSS) (Sebastiani et al., 2009, PMID: 19910380), a score generated according to birth-year cohort survival probabilities of the proband and siblings; probands and their families with FLoSS score of 7 or higher, at least one living sibling, and at least one living offspring (minimum family size of 3), who were able to give informed consent and willing to participate were recruited. The individuals were genotyped using ~2.3 million SNPs from the Illumina Omni chip, and then imputed on phased 1000 Genomes with Cosmopolitan data as a reference using MACH and MINIMAC. After excluding participants with 80 years and older, ~3,200 individuals have been included in the analyses for replication.

LOLIPOP (London Life Sciences Prospective Population Study): LOLIPOP is a population based prospective study of about 28K Indian Asian and European men and women, recruited from the lists of 58 General Practitioners in West London, United Kingdom between 2003 and 2008 [1]. Indian Asians had all four grandparents born on the Indian subcontinent. Europeans were of self-reported white ancestry. At enrolment all participants completed an interviewer-administered questionnaire for demographic data, medical history, and smoking and alcohol drinking habits. Anthropometric data were collected and blood pressure measured using an Omron 705CP with the mean of three measurements recorded. Blood samples were collected for the measurement of lipid profile after an overnight fasting of at least 8 hours. Aliquots of whole blood were stored at -80C for extraction of genomic DNA. The LOLIPOP study is approved by the local Research Ethics Committees and all participants provided written informed consent.

Loyola GxE (Kingston Gene-by-environment; subset of International Collaborative Study of Hypertension in Blacks (ICSHIB)): The Kingston GxE cohort was obtained from a survey conducted in Kingston, Jamaica as part of a larger project to examine gene by environment interactions in the determination of blood pressure among adults 25-74 years [PMID: 9103091]. The principal criterion for eligibility was a body mass index in either the top or bottom third of BMI for the Jamaican population. Participants were identified principally from the records of the Heart Foundation of Jamaica, a non-governmental organization based in Kingston, which provides low-cost screening services (height and weight, blood pressure, glucose, cholesterol) to the general public. Other participants were identified from among participants in family studies of blood pressure at the Tropical Metabolism Research Unit (TMRU) and from among staff members at the University of the West Indies, Mona.

Loyola SPT (Spanish Town; subset of International Collaborative Study of Hypertension in Blacks (ICSHIB)): Participants were recruited from Spanish Town, a stable, residential urban area neighboring the capital city of Kingston, Jamaica as part of the ICSHIB [PMID: 9103091]. A stratified random sampling scheme was used to recruit adult males and females aged 25–74 years from the general population. Spanish Town was chosen because its demographic make-up was broadly representative of Jamaica as a whole.

METSIM (Metabolic Syndrome In Men): The METSIM Study includes 10,197 men, aged from 45 to 73 years at recruitment, randomly selected from the population register of the Kuopio town, Eastern Finland, and examined in 2005-2010 (Stancakova A, et al. Diabetes 2009). The aim of the study is to investigate genetic and non-genetic factors associated with type 2 diabetes and cardiovascular disease and its risk factors.

Stancakova A, Javorsky M, Kuulasmaa T, Haffner SM, Kuusisto J, Laakso M: Changes in insulin sensitivity and insulin release in relation to glycemia and glucose tolerance in 6416 Finnish men. Diabetes 58:1212-1221, 2009.

NESDA (Netherlands Study of Depression and Anxiety): NESDA is a multi-center study designed to examine the long-term course and consequences of depressive and anxiety disorders (http://www.nesda.nl). NESDA included both individuals with depressive and/or anxiety disorders and controls without psychiatric conditions. Inclusion criteria were age 18-65 years and self-reported western European ancestry while exclusion criteria were not being fluent in Dutch and having a primary diagnosis of another psychiatric condition (psychotic disorder, obsessive compulsive disorder, bipolar disorder, or severe substance use disorder).

OBA (French obese cases): Study of the genetic of obesity in adults.

PROCARDIS (Precocious Coronary Artery Disease): The PROCARDIS (European collaborative study of the genetics of precocious coronary artery disease) study is a multi-centre case-control study in which CAD cases and controls were recruited from the United Kingdom, Italy, Sweden and Germany. Cases were defined as symptomatic CAD before age 66 years and 80% of cases also had a sibling in whom CAD had been diagnosed before age 66 years. CAD was defined as clinically documented evidence of myocardial infarction (MI) (80%), coronary artery bypass graft (CABG) (10%), acute coronary syndrome (ACS) (6%), coronary angioplasty (CA) (1%) or stable angina (hospitalization for angina or documented obstructive coronary disease) (3%). The cases included 2,136 cases who were half or full siblings. PROCARDIS controls had no personal or sibling history of CAD before age 66 years.

RHS (Ragama Health Study): The Ragama Health Study (RHS) is a population-based study of South Asian men and women aged 35-64yrs living in the Ragama Medical Officer of Health (MOH) area, near Colombo, Sri Lanka.* Consenting adults attended a clinic after a 12-h fast with available health records, and were interviewed by trained personnel to obtain information on medical, sociodemographic, and lifestyle variables. A 10-mL sample of venous blood was obtained from each subject. The concurrent study was performed in two tea plantation estates in the Lindula MOH area, near Nuwara Eliya (180 km from Colombo), to investigate the gene-environment interaction in a community with differing lifestyles (e.g., physical activity and diet). BP was measured using the Omron 750CP (Omron Co., Japan) in the seated position. The average of two readings was used for the analysis. The RHS is a collaborative effort between the Faculty of Medicine, University of Kelaniya and the National Center for Global Health and Medicine, Japan.

*Reference: Dassanayake, A.S. et al. Prevalence and risk factors for non-alcoholic fatty liver disease among adults in an urban Sri Lankan population. J Gastroenterol Hepatol 24, 1284-8 (2009).

SHEEP (Stockholm Heart Epidemiology Project): The SHEEP is a population based case-control study of risk factors for first episode of acute myocardial infarction. The study base comprised all Swedish citizens resident in the Stockholm county 1992-1994 who were 45-70 years of age and were free of previous clinically diagnosed myocardial infarction.

Cases were identified using three different sources: 1) coronary units and internal medicine wards for acute care in all Stockholm hospitals; 2) the National Patient Register; and 3) death certificates. For the present study, only cases who survived at least 28 days were considered (n=1213).

First time incident myocardial infarction cases (n=1213) were identified during a 2-year period (1992-1993) for men and during a 3-year period (1992-1994) for women. Controls (n=1561) were randomly recruited from the study population continuously over time within 2 days of the case occurrence and matched to cases on age (5-years interval), sex and hospital catchment area using computerized registers of the population of Stockholm. Five control candidates were sampled simultaneously to be able to replace potential non-respondent controls. Occasionally, because of late response of the initial control, both the first and alternative controls were considered resulting in the inclusion of more controls than cases. Postal questionnaires covering a wide range of exposure areas including occupational exposures, life style factors, social factors and health related factors were distributed to the participants. Clinical investigations were performed at least three months after myocardial infarction of cases and their matched controls. The investigations included blood samplings under fasting conditions with collection of whole blood for DNA extraction, serum and plasma. A biobank was established containing DNA, serum and plasma.

Exposure information based on both the questionnaire and biological data from the health examination was available for 78% of the male and 67% of the female non-fatal cases; the corresponding figures for their controls were 68% and 64%.

SHIP (Study of Health in Pomerania): The Study of Health In Pomerania (SHIP) is a prospective longitudinal population-based cohort study in Mecklenburg-West Pomerania assessing the prevalence and incidence of common diseases and their risk factors (PMID: 20167617). SHIP encompasses the two independent cohorts SHIP and SHIP-TREND. Participants aged 20 to 79 with German citizenship and principal residency in the study area were recruited from a random sample of residents living in the three local cities, 12 towns as well as 17 randomly selected smaller towns. Individuals were randomly selected stratified by age and sex in proportion to population size of the city, town or small towns, respectively. A total of 4,308 participants were recruited between 1997 and 2001 in the SHIP cohort. Between 2008 and 2012 a total of 4,420 participants were recruited in the SHIP-TREND cohort. Individuals were invited to the SHIP study centre for a computer-assisted personal interviews and extensive physical examinations. The study protocol was approved by the medical ethics committee of the University of Greifswald. Oral and written informed consents was obtained from each of the study participants

Genome-wide SNP-typing was performed using the Affymetrix Genome-Wide Human SNP Array 6.0 or the Illumina Human Omni 2.5 array (SHIP-TREND samples). Array processing was carried out in accordance with the manufacturer's standard recommendations. Genotypes were determined using GenomeStudio Genotyping Module v1.0 (GenCall) for SHIP-TREND and the Birdseed2 clustering algorithm for SHIP. Imputation of genotypes in SHIP and SHIP-TREND was performed with the software IMPUTE v2.2.2 based on 1000 Genomes release March 2012.

SWHS/SMHS (Shanghai Women's Health Study/ Shanghai Men's Health Study): The Shanghai Women's Health Study (SWHS) is an ongoing population-based cohort study of approximately 75,000 women who were aged 40-70 years at study enrollment and resided in in urban Shanghai, China; 56,832 (75.8%) provided a blood samples. Recruitment for the SWHS was initiated in 1997 and completed in 2000. The self-administered questionnaire includes information on demographic characteristics, disease and surgery histories, personal habits (such as cigarette smoking, alcohol consumption, tea drinking, and ginseng use), menstrual history, residential history, occupational history, and family history of cancer.

The blood pressure were measured by trained interviewers (retired nurses) with a conventional mercury sphygmomanometer according to a standard protocol, after the participants sat quietly for 5 min at the study recruitment. Included in the current project were 2970 women who had GWAS data and blood pressure measurements at the baseline interview or 892 women who had GWAS data and lipids data.

The Shanghai Men's Health Study (SMHS) is an ongoing population-based cohort study of 61,480 Chinese men who were aged between 40 and 74 years, were free of cancer at enrollment, and lived in urban Shanghai, China; 45,766 (74.4%) provided a blood samples. Recruitment for the SMHS was initiated in 2002 and completed in 2006. The self-administered questionnaire includes information on demographic characteristics, disease and surgery histories, personal habits (such as cigarette smoking, alcohol consumption, tea drinking, and ginseng use), residential history, occupational history, and family history of cancer. The blood pressure were measured by trained interviewers (retired nurses) with a conventional mercury sphygmomanometer according to a standard protocol, after the participants sat quietly for 5 min at the study recruitment. Included in the current project were 892 men who had GWAS data and blood pressure measurements at the baseline interview or 298 men who had GWAS data and lipids data.

Genotyping and imputation: Genomic DNA was extracted from buffy coats by using a Qiagen DNA purification kit (Valencia, CA) or Puregene DNA purification kit (Minneapolis, MN) according to the manufacturers' instructions and then used for genotyping assays. The GWAS genotyping was performed using the Affymetrix Genome-Wide Human SNP Array 6.0 (Affy6.0) platform or Illumina 660, following manufacturers' protocols. After sample quality control, we exclude SNPs with 1) MAF <0.01; 2) call rate <95%; 2) bad genotyping cluster; and 3) concordance rate <95% among duplicated QC samples. Genotypes were imputed using the program MACH (http://www.sph.umich.edu/csg/abecasis/MACH/download/), which determines the probable distribution of missing genotypes conditional on a set of known haplotypes, while simultaneously estimating the fine-scale recombination map. Phased autosome SNP data from HapMap Phase II Asians (release 22) were used as the reference. To test for associations between the imputed SNP data with BMI, linear regression (additive model) was used, in which SNPs were represented by the expected allele count, an approach that takes into account the degree of uncertainty of genotype imputation (http://www.sph.umich.edu/csg/abecasis/MACH/download/).

The lipid profiles were measured at Vanderbilt Lipid Laboratory. Total cholesterol, high-density lipoprotein (HDL) cholesterol, and triglycerides (TG) were measured using an ACE Clinical Chemistry System (Alfa Wassermann, Inc, West Caldwell, NJ). Low-density lipoprotein (LDL) cholesterol levels were calculated by using the Friedwald equation. The levels of LDL cholesterol were directly measured using an ACE Clinical Chemistry System for subjects with TG levels ≥ 400 mg/dL. Fasting status was defined as an interval between the last meal and blood draw of 8 hours or longer.

TAICHI-G: The TaiChi consortium consists of 7 studies that collaborated initially in a large scale metabochip study, and became an ongoing consortium for studies of cardiometabolic disease in the Chinese population in Taiwan. The seven studies included the following: 1) HALST (Healthy Aging Longitudinal Study in Taiwan), a population based epidemiologic study of older adults living in all major geographic regions of Taiwan established by the Taiwan National Health Research Institutes (NHRI); 2) SAPPHIRe (Stanford-Asian Pacific Program in Hypertension and Insulin Resistance), a family based study established in 1995 with an initial goal of identifying major genetic loci underlying hypertension and insulin resistance in East Asian populations, with Taiwan subjects participating in the TaiChi consortium; 3) TCAGEN (Taiwan Coronary Artery Disease GENetic), a cohort study that that enrolled patients undergoing coronary angiography or percutaneous intervention at the National Taiwan University Hospital (NTUH) in the setting of either stable angina pectoris or prior myocardial infarction; 4) TACT (TAiwan Coronary and Transcatheter intervention), a cohort study enrolled patients with angina pectoris and objective documentation of myocardial ischemia who underwent diagnostic coronary angiography and/or revascularization any time after October 2000 at the National Taiwan

University Hospital (NTUH) (similar to TCAGEN but recruitment was independent of TCAGEN); 5) Taiwan DRAGON (Taiwan Diabetes and RelAted Genetic COmplicatioN), acohort study of Type 2 diabetes at Taichung Veterans General Hospital (Taichung VGH) in Taiwan, with participants including individuals with either newly diagnosed or established diabetes (subjects with hyperglycemia who did not meet diagnostic criteria for Type 2 DM were not included); 6) TCAD (Taichung CAD study), includes patients with a variety of cardiovascular diseases who received care at the Taichung Veterans General Hospital (Taichung VGH), i.e. specifically individuals who were hospitalized for diagnostic and interventional coronary angiography examinations and treatment; 7) TUDR (Taiwan US Diabetic Retinopathy) enrolled subjects with Type 2 diabetes who received care at Taichung Veteran General Hospital (Taichung VGH), and a small number of subjects from Taipei Tri-Service General Hospital (TSGH); TUDR subjects underwent a complete ophthalmic and fundus examination to carefully document the presence and extent of retinopathy. From these 7 studies, samples for over 1,800 subjects were selected based on completeness of standard metabolic phenotyping and knowledge of cardiac disease status, to undergo GWAS genotyping with an Illumina human-omni 'chip' specific for Asian population (Illumina, San Diego, CA; cat. No. 20004337), hence TAICHI-G.

THRV (Taiwan study of Hypertensives Rare Variants): THRV proposed to identify rare and low frequency genetic variants for blood pressure and hypertension through whole exome sequencing of a subset of highly enriched Taiwan Chinese hypertensive families and as many matched controls. The Taiwan Chinese families (approximately N=1,200 subjects) were previously recruited as part of the NHLBI-sponsored SAPPHIRe Network which is part of the Family Blood Pressure Program (FBPP). The SAPPHIRe families were recruited to have multiple hypertensive sibs and some of them also included one normotensive/hypotensive sib. The matched controls (N=1,200) were selected from the large population-based HALST Study and a Hospital-based population, both in Taipei, Taiwan.

TRAILS (Tracking Adolescents' Individual Lives Survey): TRAILS is a prospective cohort study of Dutch adolescents and young adults, with bi- or triennial measurements from age 11 onwards, which started in 2001. TRAILS consists of a general population and a clinical cohort (https://www.trails.nl/en/home). In the population cohort, six assessment waves have been completed to date, at mean ages 11.1 (SD = 0.6), 13.6 (SD = 0.5), 16.3 (SD = 0.7), 19.1 (SD = 0.6), 22.3 (SD = 0.6), and 25.8 (SD = 0.6). Data for the present study were collected in the population cohort only, during the third assessment wave. The study was approved by the Dutch Central Committee on Research Involving Human Subjects.

TUDR (Taiwan-US Diabetic Retinopathy): 2009 to present, is a cohort that enrolled subjects with Type 2 diabetes receiving care at Taichung Veteran General Hospital (Taichung VGH), and a small number of subjects from Taipei Tri-Service General Hospital. All TUDR subjects underwent a complete ophthalmic and fundus examination to carefully document the presence and extent of retinopathy.

TWINGENE (TwinGene of the Swedish Twin Registry): The aim of the TwinGene project has been to systematically transform the oldest cohorts of the Swedish Twin Registry (STR) into a molecular-genetic resource. Beginning in 2004, about 200 twins were contacted each month until the data collection was completed in 2008. A total of 21 500 twins were contacted where of 12 600 participated. Invitations to the study contained information of the study and its purpose. Along with the invitations consent forms and health questionnaire were sent to the subjects. When the signed consent forms where returned, the subjects were sent blood sampling equipment and asked to contact a local health facility for blood sampling. The study population was recruited among twins participating in the Screening Across the Lifespan Twin Study (SALT) which was a telephone interview study conducted in 1998-2002. Other inclusion criteria were that both twins in the pair had to be alive and living in Sweden. Subjects were excluded from the study if they preciously declined participation in future studies or if they had been enrolled in other STR DNA sampling projects. The subjects were asked to make an appointment for a health check-up at their local health-care facility on the morning Monday to Thursday and not the day before a national holiday, this to ensure that the sample would reach the KI biobank the

following morning by overnight mail. The subjects were instructed to fast from 20.00 the previous night. By venipuncture a total of 50 ml of blood was drawn from each subject. Tubes with serum and blood for biobanking as well as for clinical chemistry tests were sent to KI by overnight mail. One 7ml EDTA tube of whole blood is stored in -80°C while a second 7ml EDTA tube of blood is used for DNA extraction using Puregene extraction kit (Gentra systems, Minneapolis, USA). After excluding subjects in which the DNA concentration in the stock-solution was below 20ng/µl as well as subset of 302 female monozygous twin pairs participating in a previous genome wide effort DNA from 9896 individual subjects was sent to SNP&SEQ Technology Platform Uppsala, Sweden for genome wide genotyping with Illumina OmniExpress bead chip (all available dizygous twins + one twin from each available MZ twin pair).

UKB (United Kingdom Biobank, www.ukbiobank.ac.uk): UK Biobank is a major national health resource with the aim of improving the prevention, diagnosis and treatment of a wide range of serious and life-threatening illnesses. UK Biobank includes data from 502,682 individuals (94% of self-reported European ancestry), with extensive health and lifestyle questionnaire data, physical measures and genetic data. A total of 152,249 participants had genetic and phenotypic (blood pressure) data. Central genotyping quality control (QC) had been performed by UK Biobank [The UK Biobank. UK Biobank Genotyping QC documentation. (2015)]. Further QC was also performed locally.

UKHLS (Understanding Society / The UK Household Longitudinal Study): The United Kingdom Household Longitudinal Study, also known as Understanding Society (https://www.understandingsociety.ac.uk) is a longitudinal panel survey of 40.000 UK households (England, Scotland, Wales and Northern Ireland) representative of the UK population. Participants are surveyed annually since 2009 and contribute information relating to their socioeconomic circumstances, attitudes, and behaviours via a computer assisted interview. The study includes phenotypical data for a representative sample of participants for a wide range of social and economic indicators as well as a biological sample collection encompassing biometric, physiological, biochemical, and haematological measurements and self-reported medical history and medication use. The United Kingdom Household Longitudinal Study has been approved by the University of Essex Ethics Committee and informed consent was obtained from every participant.

YFS (The Cardiovascular Risk in Young Finns Study): The YFS is a population-based follow up-study started in 1980. The main aim of the YFS is to determine the contribution made by childhood lifestyle, biological and psychological measures to the risk of cardiovascular diseases in adulthood. In 1980, over 3,500 children and adolescents all around Finland participated in the baseline study. The follow-up studies have been conducted mainly with 3-year intervals. The latest 30-year follow-up study was conducted in 2010-11 (ages 33-49 years) with 2,063 participants. The study was approved by the local ethics committees (University Hospitals of Helsinki, Turku, Tampere, Kuopio and Oulu) and was conducted following the guidelines of the Declaration of Helsinki. All participants gave their written informed consent.

NOTE: Baependi, NEO, Pelotas, and WHI (EA) also participated in replications since they did not contribute to Smoking-BP discovery analysis.

Stage 1 Study Acknowledgments

Infrastructure for the CHARGE Consortium is supported in part by the National Heart, Lung, and Blood Institute grant R01HL105756. Infrastructure for the Gene-Lifestyle Working Group is supported by the National Heart, Lung, and Blood Institute grant R01HL118305.

AGES (Age Gene/Environment Susceptibility Reykjavik Study): This study has been funded by NIH contract N01-AG012100, the NIA Intramural Research Program, an Intramural Research Program Award (ZIAEY000401) from the National Eye Institute, an award from the National Institute on Deafness and Other Communication Disorders (NIDCD) Division of Scientific Programs (IAA Y2-DC_1004-02), Hjartavernd (the Icelandic Heart Association), and the Althingi (the Icelandic Parliament). The study is approved by the Icelandic National Bioethics Committee, VSN: 00-063. The researchers are indebted to the participants for their willingness to participate in the study.

ARIC (Atherosclerosis Risk in Communities) Study: The ARIC study is carried out as a collaborative study supported by National Heart, Lung, and Blood Institute contracts (HHSN268201100005C, HHSN268201100006C, HHSN268201100007C, HHSN268201100008C, HHSN268201100009C, HHSN268201100010C, HHSN268201100011C, and HHSN268201100012C), R01HL087641, R01HL59367 and R01HL086694; National Human Genome Research Institute contract U01HG004402; and National Institutes of Health contract HHSN268200625226C. The authors thank the staff and participants of the ARIC study for their important contributions. Infrastructure was partly supported by Grant Number UL1RR025005, a component of the National Institutes of Health and NIH Roadmap for Medical Research.

Baependi Heart Study (Brazil): The Baependi Heart Study was supported by Fundação de Amparo a Pesquisa do Estado de São Paulo (FAPESP) (Grant 2013/17368-0), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) and Hospital Samaritano Society (Grant 25000.180.664/2011-35), through Ministry of Health to Support Program Institutional Development of the Unified Health System (SUS-PROADI).

Bio*Me* Biobank (Bio*Me* Biobank of Institute for Personalized Medicine at Mount Sinai): The Mount Sinai IPM Biobank Program is supported by The Andrea and Charles Bronfman Philanthropies.

CARDIA (Coronary Artery Risk Development in Young Adults): The CARDIA Study is conducted and supported by the National Heart, Lung, and Blood Institute in collaboration with the University of Alabama at Birmingham (HHSN268201300025C & HHSN268201300026C), Northwestern University (HHSN268201300027C), University of Minnesota (HHSN268201300028C), Kaiser Foundation Research Institute (HHSN268201300029C), and Johns Hopkins University School of Medicine (HHSN268200900041C). CARDIA is also partially supported by the Intramural Research Program of the National Institute on Aging. Genotyping was funded as part of the NHLBI Candidate-gene Association Resource (N01-HC-65226) and the NHGRI Gene Environment Association Studies (GENEVA) (U01-HG004729, U01-HG04424, and U01-HG004446). This manuscript has been reviewed and approved by CARDIA for scientific content.

CHS (Cardiovascular Health Study): This CHS research was supported by NHLBI contracts HHSN268201200036C, HHSN268200800007C, HHSN268200960009C, N01HC55222, N01HC85079, N01HC85080, N01HC85081, N01HC85082, N01HC85083, N01HC85086; and NHLBI grants U01HL080295, R01HL085251, R01HL087652, R01HL105756, R01HL103612, R01HL120393 and R01HL130114 with additional contribution from the National Institute of Neurological Disorders and Stroke (NINDS). Additional support was provided through R01AG023629 from the National Institute on Aging (NIA). A full list of principal CHS investigators and institutions can be found at CHS-NHLBI.org. The provision of genotyping data was supported in part by the National Center for Advancing Translational Sciences, CTSI grant UL1TR000124, and the National Institute of Diabetes and Digestive and Kidney Disease Diabetes Research Center (DRC) grant DK063491 to the Southern California

Diabetes Endocrinology Research Center. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

ERF (Erasmus Rucphen Family study): The ERF study as a part of EUROSPAN (European Special Populations Research Network) was supported by European Commission FP6 STRP grant number 018947 (LSHG-CT-2006-01947) and also received funding from the European Community's Seventh Framework Programme (FP7/2007-2013)/grant agreement HEALTH-F4-2007-201413 by the European Commission under the programme "Quality of Life and Management of the Living Resources" of 5th Framework Programme (no. QLG2-CT-2002-01254). The ERF study was further supported by ENGAGE consortium and CMSB. High-throughput analysis of the ERF data was supported by joint grant from Netherlands Organisation for Scientific Research and the Russian Foundation for Basic Research (NWO-RFBR 047.017.043). ERF was further supported by the ZonMw grant (project 91111025). We are grateful to all study participants and their relatives, general practitioners and neurologists for their contributions and to P. Veraart for her help in genealogy, J. Vergeer for the supervision of the laboratory work, P. Snijders for his help in data collection and E.M. van Leeuwen for genetic imputation.

FamHS (Family Heart Study): The FamHS is funded by R01HL118305 and R01HL117078 NHLBI grants, and 5R01DK07568102 and 5R01DK089256 NIDDK grant.

FHS (Framingham Heart Study): This research was conducted in part using data and resources from the Framingham Heart Study of the National Heart Lung and Blood Institute of the National Institutes of Health and Boston University School of Medicine. The analyses reflect intellectual input and resource development from the Framingham Heart Study investigators participating in the SNP Health Association Resource (SHARe) project. This work was partially supported by the National Heart, Lung and Blood Institute's Framingham Heart Study (Contract Nos. N01-HC-25195 and HHSN268201500001I) and its contract with Affymetrix, Inc for genotyping services (Contract No. N02-HL-6-4278). A portion of this research utilized the Linux Cluster for Genetic Analysis (LinGA-II) funded by the Robert Dawson Evans Endowment of the Department of Medicine at Boston University School of Medicine and Boston Medical Center. This research was partially supported by grant R01-DK089256 from the National Institute of Diabetes and Digestive and Kidney Diseases (MPIs: Ingrid B. Borecki, L. Adrienne Cupples, Kari North).

GENOA (Genetic Epidemiology Network of Arteriopathy): Support for GENOA was provided by the National Heart, Lung and Blood Institute (HL119443, HL118305, HL054464, HL054457, HL054481, HL071917 and HL087660) of the National Institutes of Health. Genotyping was performed at the Mayo Clinic (Stephen T. Turner, MD, Mariza de Andrade PhD, Julie Cunningham, PhD). We thank Eric Boerwinkle, PhD and Megan L. Grove from the Human Genetics Center and Institute of Molecular Medicine and Division of Epidemiology, University of Texas Health Science Center, Houston, Texas, USA for their help with genotyping. We would also like to thank the families that participated in the GENOA study.

GenSalt (Genetic Epidemiology Network of Salt Sensitivity): The Genetic Epidemiology Network of Salt Sensitivity is supported by research grants (U01HL072507, R01HL087263, and R01HL090682) from the National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, MD.

GOLDN (Genetics of Diet and Lipid Lowering Network): Support for the genome-wide association studies in GOLDN was provided by the National Heart, Lung, and Blood Institute grant U01HL072524-04 and R01HL091357.

HANDLS (Healthy Aging in Neighborhoods of Diversity across the Life Span): The Healthy Aging in Neighborhoods of Diversity across the Life Span (HANDLS) study was supported by the Intramural Research Program of the NIH, National Institute on Aging and the National Center on Minority Health

and Health Disparities (project # Z01-AG000513 and human subjects protocol number 09-AG-N248). Data analyses for the HANDLS study utilized the high-performance computational resources of the Biowulf Linux cluster at the National Institutes of Health, Bethesda, MD. (http://biowulf.nih.gov; http://hpc.nih.gov)).

Health ABC (Health, Aging, and Body Composition): Health ABC was funded by the National Institutes of Aging. This research was supported by NIA contracts N01AG62101, N01AG62103, and N01AG62106. The GWAS was funded by NIA grant 1R01AG032098-01A1 to Wake Forest University Health Sciences and genotyping services were provided by the Center for Inherited Disease Research (CIDR). CIDR is fully funded through a federal contract from the National Institutes of Health to The Johns Hopkins University, contract number HHSN268200782096C. This research was supported in part by the Intramural Research Program of the NIH, National Institute on Aging.

HERITAGE (Health, Risk Factors, Exercise Training and Genetics): The HERITAGE Family Study was supported by National Heart, Lung, and Blood Institute grant HL-45670.

HUFS (Howard University Family Study): The Howard University Family Study was supported by National Institutes of Health grants S06GM008016-320107 to Charles Rotimi and S06GM008016-380111 to Adebowale Adeyemo. We thank the participants of the study, for which enrollment was carried out at the Howard University General Clinical Research Center, supported by National Institutes of Health grant 2M01RR010284. The contents of this publication are solely the responsibility of the authors and do not necessarily represent the official view of the National Institutes of Health. This research was supported in part by the Intramural Research Program of the Center for Research on Genomics and Global Health (CRGGH). The CRGGH is supported by the National Human Genome Research Institute, the National Institute of Diabetes and Digestive and Kidney Diseases, the Center for Information Technology, and the Office of the Director at the National Institutes of Health (Z01HG200362). Genotyping support was provided by the Coriell Institute for Medical Research.

HyperGEN (Hypertension Genetic Epidemiology Network): The hypertension network is funded by cooperative agreements (U10) with NHLBI: HL54471, HL54472, HL54473, HL54495, HL54496, HL54497, HL54509, HL54515, and 2 R01 HL55673-12. The study involves: University of Utah: (Network Coordinating Center, Field Center, and Molecular Genetics Lab); Univ. of Alabama at Birmingham: (Field Center and Echo Coordinating and Analysis Center); Medical College of Wisconsin: (Echo Genotyping Lab); Boston University: (Field Center); University of Minnesota: (Field Center and Biochemistry Lab); University of North Carolina: (Field Center); Washington University: (Data Coordinating Center); Weil Cornell Medical College: (Echo Reading Center); National Heart, Lung, & Blood Institute. For a complete list of HyperGEN Investigators: http://www.biostat.wustl.edu/hypergen/Acknowledge.html

IGMM (Institute of Genetics and Molecular Medicine): CROATIA-Korcula: We would like to acknowledge the staff of several institutions in Croatia that supported the field work, including but not limited to The University of Split and Zagreb Medical Schools and the Croatian Institute for Public Health. We would like to acknowledge the invaluable contributions of the recruitment team in Korcula, the administrative teams in Croatia and Edinburgh and the participants. The SNP genotyping for the CROATIA-Korcula cohort was performed in Helmholtz Zentrum München, Neuherberg, Germany. CROATIA-Korcula (CR-Korcula) was funded by the Medical Research Council UK, The Croatian Ministry of Science, Education and Sports (grant 216-1080315-0302), the European Union framework program 6 EUROSPAN project (contract no. LSHG-CT-2006-018947) and the Croatian Science Foundation (grant 8875). **CROATIA-Vis:** We would like to acknowledge the staff of several institutions in Croatia that supported the field work, including but not limited to The University of Split and Zagreb Medical Schools, the Institute for Anthropological Research in Zagreb and Croatian Institute for Public Health. The SNP genotyping for the CROATIA-Vis cohort was performed in the core genotyping laboratory of the Wellcome Trust Clinical Research Facility at the Western General Hospital, Edinburgh,

Scotland. CROATIA-Vis (CR-Vis) was funded by the Medical Research Council UK, The Croatian Ministry of Science, Education and Sports (grant 216-1080315-0302), the European Union framework program 6 EUROSPAN project (contract no. LSHG-CT-2006-018947) and the Croatian Science Foundation (grant 8875). **GS:SFHS:** Generation Scotland received core support from the Chief Scientist Office of the Scottish Government Health Directorates [CZD/16/6] and the Scottish Funding Council [HR03006]. Genotyping of the GS:SFHS samples was carried out by the Genetics Core Laboratory at the Wellcome Trust Clinical Research Facility, Edinburgh, Scotland and was funded by the Medical Research Council UK and the Wellcome Trust (Wellcome Trust Strategic Award "STratifying Resilience and Depression Longitudinally" (STRADL) Reference 104036/Z/14/Z). Ethics approval for the study was given by the NHS Tayside committee on research ethics (reference 05/S1401/89). We are grateful to all the families who took part, the general practitioners and the Scottish School of Primary Care for their help in recruiting them, and the whole Generation Scotland team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists, healthcare assistants and nurses.

JHS (Jackson Heart Study): The Jackson Heart Study is supported by contracts HSN268201300046C, HHSN268201300047C, HHSN268201300048C, HHSN268201300049C, HHSN268201300050C from the National Heart, Lung, and Blood Institute on Minority Health and Health Disparities. The authors acknowledge the Jackson Heart Study team institutions (University of Mississippi Medical Center, Jackson State University and Tougaloo College) and participants for their long-term commitment that continues to improve our understanding of the genetic epidemiology of cardiovascular and other chronic diseases among African Americans.

Maywood-Loyola Study: Maywood African-American study is supported in part by the National Institutes of Health grant numbers HL074166, R01HL074166, R01HG003054, R37HL45508 and R01HL53353.

Maywood-Nigeria Study: The Loyola-Nigeria study was supported by National Institutes of Health grant number R01HL053353 and the Intramural Research Program of the Center for Research on Genomics and Global Health, National Human Genome Research Institute (Z01HG200362). The authors acknowledge the assistance of the research staff and participants in Ibadan and Igbo-Ora, Oyo State, Nigeria.

MESA (Multi-Ethnic Study of Atherosclerosis): This research was supported by the Multi-Ethnic Study of Atherosclerosis (MESA) contracts N01-HC-95159, N01-HC-95160, N01-HC-95161, N01-HC-95162, N01-HC-95163, N01-HC-95164, N01-HC-95165, N01-HC-95166, N01-HC-95167, N01-HC-95168, N01-HC-95169, by grant HL071205 and by UL1-DR-001079 from NCRR . Funding for MESA SHARe genotyping was provided by NHLBI Contract N02-HL-6-4278. This publication was partially developed under a STAR research assistance agreement, No. RD831697 (MESA Air), awarded by the U.S Environmental Protection Agency. It has not been formally reviewed by the EPA. The views expressed in this document are solely those of the authors and the EPA does not endorse any products or commercial services mentioned in this publication. The provision of genotyping data was supported in part by the National Center for Advancing Translational Sciences, CTSI grant UL1TR000124, and the National Institute of Diabetes and Digestive and Kidney Disease Diabetes Research Center (DRC) grant DK063491 to the Southern California Diabetes Endocrinology Research Center. The authors thank the participants of the MESA study, the Coordinating Center, MESA investigators, and study staff for their valuable contributions. A full list of participating MESA investigators and institutions can be found at http://www.mesa-nhlbi.org.

NEO (The Netherlands Epidemiology of Obesity study): The authors of the NEO study thank all individuals who participated in the Netherlands Epidemiology in Obesity study, all participating general practitioners for inviting eligible participants and all research nurses for collection of the data. We thank the NEO study group, Petra Noordijk, Pat van Beelen and Ingeborg de Jonge for the coordination, lab

and data management of the NEO study. The genotyping in the NEO study was supported by the Centre National de Génotypage (Paris, France), headed by Jean-Francois Deleuze. The NEO study is supported by the participating Departments, the Division and the Board of Directors of the Leiden University Medical Center, and by the Leiden University, Research Profile Area Vascular and Regenerative Medicine. Dennis Mook-Kanamori is supported by Dutch Science Organization (ZonMW-VENI Grant 916.14.023).

Pelotas Birth Cohort Study (The 1982 Pelotas Birth Cohort Study, Brazil): The 1982 Pelotas Birth Cohort Study is conducted by the Postgraduate Program in Epidemiology at Universidade Federal de Pelotas with the collaboration of the Brazilian Public Health Association (ABRASCO). From 2004 to 2013, the Wellcome Trust supported the study. The International Development Research Center, World Health Organization, Overseas Development Administration, European Union, National Support Program for Centers of Excellence (PRONEX), the Brazilian National Research Council (CNPq), and the Brazilian Ministry of Health supported previous phases of the study.

Genotyping of 1982 Pelotas Birth Cohort Study participants was supported by the Department of Science and Technology (DECIT, Ministry of Health) and National Fund for Scientific and Technological Development (FNDCT, Ministry of Science and Technology), Funding of Studies and Projects (FINEP, Ministry of Science and Technology, Brazil), Coordination of Improvement of Higher Education Personnel (CAPES, Ministry of Education, Brazil).

RS (Rotterdam Study): The Rotterdam Study is funded by Erasmus Medical Center and Erasmus University, Rotterdam, Netherlands Organization for the Health Research and Development (ZonMw), the Research Institute for Diseases in the Elderly (RIDE), the Ministry of Education, Culture and Science, the Ministry for Health, Welfare and Sports, the European Commission (DG XII), and the Municipality of Rotterdam. The authors are grateful to the study participants, the staff from the Rotterdam Study and the participating general practitioners and pharmacists.

The generation and management of GWAS genotype data for the Rotterdam Study was executed by the Human Genotyping Facility of the Genetic Laboratory of the Department of Internal Medicine, Erasmus MC, Rotterdam, The Netherlands. The GWAS datasets are supported by the Netherlands Organisation of Scientific Research NWO Investments (nr. 175.010.2005.011, 911-03-012), the Genetic Laboratory of the Department of Internal Medicine, Erasmus MC, the Research Institute for Diseases in the Elderly (014-93-015; RIDE2), the Netherlands Genomics Initiative (NGI)/Netherlands Organisation for Scientific Research (NWO) Netherlands Consortium for Healthy Aging (NCHA), project nr. 050-060-810. We thank Pascal Arp, Mila Jhamai, Marijn Verkerk, Lizbeth Herrera, Marjolein Peters and Carolina Medina-Gomez for their help in creating the GWAS database, and Karol Estrada, Yurii Aulchenko and Carolina Medina-Gomez for the creation and analysis of imputed data.

SCHS-CHD (Singapore Chinese Health Study - Coronary Heart Disease): The Singapore Chinese Health Study is supported by the National Institutes of Health, USA (RO1 CA144034 and UM1 CA182876), the nested case-control study of myocardial infarction by the Singapore National Medical Research Council (NMRC 1270/2010) and genotyping by the HUJ-CREATE Programme of the National Research Foundation, Singapore (Project Number 370062002).

SCES (Singapore Chinese Eye Study), SiMES (Singapore Malay Eye Study), (SINDI) Singapore Indian Eye Study: The Singapore Malay Eye Study (SiMES), the Singapore Indian Eye Study (SINDI), and the Singapore Chinese Eye Study (SCES) are supported by the National Medical Research Council (NMRC), Singapore (grants 0796/2003, 1176/2008, 1149/2008, STaR/0003/2008, 1249/2010, CG/SERI/2010, CIRG/1371/2013, and CIRG/1417/2015), and Biomedical Research Council (BMRC), Singapore (08/1/35/19/550 and 09/1/35/19/616). Ching-Yu Cheng is supported by an award from NMRC (CSA/033/2012). The Singapore Tissue Network and the Genome Institute of Singapore, Agency for Science, Technology and Research, Singapore provided services for tissue archival and

genotyping, respectively. **SP2 (Singapore Prospective Study Program):** SP2 is supported by the individual research grant and clinician scientist award schemes from the National Medical Research Council and the Biomedical Research Councils of Singapore.

WGHS (Women's Genome Health Study): The WGHS is supported by the National Heart, Lung, and Blood Institute (HL043851 and HL080467) and the National Cancer Institute (CA047988 and UM1CA182913), with collaborative scientific support and funding for genotyping provided by Amgen.

WHI (Women's Health Initiative): The WHI program is funded by the National Heart, Lung, and Blood Institute, National Institutes of Health, U.S. Department of Health and Human Services through contracts HHSN268201100046C, HHSN268201100001C, HHSN268201100002C, HHSN268201100003C, HHSN268201100004C, and HHSN271201100004C. The authors thank the WHI investigators and staff for their dedication, and the study participants for making the program possible. A full listing of WHI investigators can be found at: http://www.whi.org/researchers/Documents%20%20Write%20a%20Paper/WHI%20Investigator%20Short%20List.pdf

Stage 2 Study Acknowledgments

Infrastructure for the CHARGE Consortium is supported in part by the National Heart, Lung, and Blood Institute grant R01HL105756. Infrastructure for the Gene-Lifestyle Working Group is supported by the National Heart, Lung, and Blood Institute grant R01HL118305.

AA-DHS (African American Diabetes Heart Study): The investigators acknowledge the cooperation of our Diabetes Heart Study (DHS) and AA-DHS participants. This work was supported by NIH R01 DK071891, R01 HL092301 and the General Clinical Research Center of Wake Forest School of Medicine M01-RR-07122.

Airwave (The Airwave Health Monitoring Study): We thank all participants in the Airwave Health Monitoring Study. The study is funded by the Home Office (Grant number 780-TETRA) with additional support from the National Institute for Health Research (NIHR), Imperial College Healthcare NHS Trust (ICHNT) and Imperial College Biomedical Research Centre (BRC). The study has ethical approval from the National Health Service Multi-site Research Ethics Committee (MREC/13/NW/0588). This work used computing resources provided by the MRC- funded UK MEDical Bioinformatics partnership programme (UK MED-BIO) (MR/L01632X/1). P.E. would like to acknowledge support from the Medical Research Council and Public Health England for the MRC-PHE Centre for Environment and Health (MR/L01341X/1) and from the NIHR NIHR Health Protection Research Unit in Health Impact of Environmental Hazards (HPRU-2012-10141).

ASCOT (Anglo-Scandinavian Cardiac Outcomes Trial): The ASCOT study was supported by Pfizer, New York, NY, USA for the ASCOT study and the collection of the ASCOT DNA repository; by Servier Research Group, Paris, France; and by Leo Laboratories, Copenhagen, Denmark. We thank all ASCOT trial participants, physicians, nurses, and practices in the participating countries for their important contribution to the study. In particular we thank Clare Muckian and David Toomey for their help in DNA extraction, storage, and handling. Genotyping was funded by the CNG, MRC and the National Institutes of Health Research (NIHR). We would also like to acknowledge the Barts and The London Genome Centre staff for genotyping. This work forms part of the research programme of the NIHR Cardiovascular Biomedical Research Unit at Barts and The London, QMUL. H.R.W, M.J.C and P.B.M. wishes to acknowledge the NIHR Cardiovascular Biomedical Research Unit at Barts and The London, Queen Mary University of London, UK for support.

BBJ (Biobank Japan Project): BioBank Japan project is supported by the Japan Agency for Medical Research and Development and by the Ministry of Education, Culture, Sports, Sciences and Technology of the Japanese government.

BES (Beijing Eye Study): BES was supported by the National Key Laboratory Fund, Beijing, China.

BRIGHT (British Genetics of Hypertension): This work was supported by the Medical Research Council of Great Britain (grant number G9521010D) and the British Heart Foundation (grant number PG/02/128). The BRIGHT study is extremely grateful to all the patients who participated in the study and the BRIGHT nursing team. This work forms part of the research program of the National Institutes of Health Research (NIHR Cardiovascular Biomedical Research) Cardiovascular Biomedical Unit at Barts and The London, QMUL.

CAGE-Amagasaki (Cardio-metabolic Genome Epidemiology Network, Amagasaki Study): The CAGE Network studies were supported by grants for the Core Research for Evolutional Science and Technology (CREST) from the Japan Science Technology Agency; the Program for Promotion of Fundamental Studies in Health Sciences, National Institute of Biomedical Innovation Organization (NIBIO); and the Grant of National Center for Global Health and Medicine (NCGM).

CFS (Cleveland Family Study): The CFS was supported by the National Institutes of Health, the National Heart, Lung, Blood Institute grant HL113338, R01HL098433, HL46380.

CoLaus (Cohorte Lausannoise): The CoLaus study was and is supported by research grants from GlaxoSmithKline, the Faculty of Biology and Medicine of Lausanne, and the Swiss National Science Foundation (grants 33CSCO-122661, 33CS30-139468 and 33CS30-148401).

DESIR (Data from an Epidemiological Study on the Insulin Resistance): The DESIR Study Group is composed of Inserm-U1018 (Paris: B. Balkau, P. Ducimetière, E. Eschwège), Inserm-U367 (Paris: F. Alhenc-Gelas), CHU d'Angers (A. Girault), Bichat Hospital (Paris: F. Fumeron, M. Marre, R. Roussel), CHU de Rennes (F. Bonnet), CNRS UMR-8199 (Lille: A. Bonnefond, P. Froguel), Medical Examination Services (Alençon, Angers, Blois, Caen, Chartres, Chateauroux, Cholet, LeMans, Orléans and Tours), Research Institute for General Medicine (J. Cogneau), the general practitioners of the region and the Cross- Regional Institute for Health (C. Born, E. Caces, M. Cailleau, N. Copin, J.G. Moreau, F. Rakotozafy, J. Tichet, S. Vol).

The DESIR study was supported by Inserm contracts with CNAMTS, Lilly, Novartis Pharma and Sanofi-aventis, and by Inserm (Réseaux en Santé Publique, Interactions entre les déterminants de la santé, Cohortes Santé TGIR 2008), the Association Diabète Risque Vasculaire, the Fédération Française de Cardiologie, La Fondation de France, ALFEDIAM, ONIVINS, Société Francophone du Diabète, Ardix Medical, Bayer Diagnostics, Becton Dickinson, Cardionics, Merck Santé, Novo Nordisk, Pierre Fabre, Roche and Topcon.

DFTJ (Dongfeng-Tongji Cohort Study): This work was supported by grants from the National Basic Research Program grant (2011CB503800), the Programme of Introducing Talents of Discipline, the grants from the National Natural Science Foundation (grant NSFC-81473051, 81522040 and 81230069), and the Program for the New Century Excellent Talents in University (NCET-11-0169).

DHS (Diabetes Heart Study): The authors thank the investigators, staff, and participants of the DHS for their valuable contributions. This study was supported by the National Institutes of Health through HL67348 and HL092301.

DR's EXTRA (Dose Responses to Exercise Training): The study was supported by grants from Ministry of Education and Culture of Finland (722 and 627; 2004-2010); Academy of Finland (102318, 104943, 123885, 211119); European Commission FP6 Integrated Project (EXGENESIS), LSHM-CT-2004-005272; City of Kuopio; Juho Vainio Foundation; Finnish Diabetes Association; Finnish Foundation for Cardiovascular Research; Kuopio University Hospital; Päivikki and Sakari Sohlberg Foundation; Social Insurance Institution of Finland 4/26/2010.

EGCUT (Estonian Genome Center - University of Tartu (Estonian Biobank)): This study was supported by EU H2020 grants 692145, 676550, 654248, Estonian Research Council Grant IUT20-60, NIASC, EIT – Health and NIH-BMI Grant No: 2R01DK075787-06A1 and EU through the European Regional Development Fund (Project No. 2014-2020.4.01.15-0012 GENTRANSMED.

EPIC (European Prospective Investigation into Cancer and Nutrition): The EPIC Norfolk Study is funded by Cancer Research, United Kingdom, British Heart Foundation, the Medical Research Council, the Ministry of Agriculture, Fisheries and Food, and the Europe against Cancer Programme of the Commission of the European Communities. We thank all EPIC participants and staff for their contribution to the study. We thank staff from the Technical, Field Epidemiology and Data Functional Group Teams of the Medical Research Council Epidemiology Unit in Cambridge, UK, for carrying out sample preparation, DNA provision and quality control, genotyping and data handling work. We specifically thank Sarah Dawson for coordinating the sample provision for biomarker measurements, Abigail Britten for coordinating DNA sample provision and genotyping of candidate markers, Nicola

Kerrison, Chris Gillson and Abigail Britten for data provision and genotyping quality control, Matt Sims for writing the technical laboratory specification for the intermediate pathway biomarker measurements and for overseeing the laboratory work.

FENLAND (The Fenland Study): The Fenland Study is funded by the Wellcome Trust and the Medical Research Council (MC_U106179471). We are grateful to all the volunteers for their time and help, and to the General Practitioners and practice staff for assistance with recruitment. We thank the Fenland Study Investigators, Fenland Study Co-ordination team and the Epidemiology Field, Data and Laboratory teams. We further acknowledge support from the Medical research council (MC_UU_12015/1).

FUSION (Finland-United States Investigation of NIDDM Genetics): The FUSION study was supported by DK093757, DK072193, DK062370, and ZIA-HG000024.

Genotyping was conducted at the Genetic Resources Core Facility (GRCF) at the Johns Hopkins Institute of Genetic Medicine.

GeneSTAR (Genetic Studies of Atherosclerosis Risk): [for the smoking/lipids and smoking/BP analyses] GeneSTAR was supported by National Institutes of Health grants from the National Heart, Lung, and Blood Institute (HL49762, HL59684, HL58625, HL071025, U01 HL72518, and HL087698), National Institute of Nursing Research (NR0224103), and by a grant from the National Center for Research Resources to the Johns Hopkins General Clinical Research Center (M01-RR000052).

[for the alcohol/lipids and alcohol/BP analyses] GeneSTAR was supported by National Institutes of Health grants from the National Heart, Lung, and Blood Institute (HL49762, HL59684, HL58625, HL071025, U01 HL72518, HL087698, HL092165, HL099747, and K23HL105897), National Institute of Nursing Research (NR0224103), National Institute of Neurological Disorders and Stroke (NS062059), and by grants from the National Center for Research Resources to the Johns Hopkins General Clinical Research Center (M01-RR000052) and the Johns Hopkins Institute for Clinical & Translational Research (UL1 RR 025005).

GLACIER (Gene x Lifestyle Interactions and Complex Traits Involved in Elevated Disease Risk): We thank the participants, health professionals and data managers involved in the Västerbottens Intervention Project. We are also grateful to the staff of the Northern Sweden Biobank for preparing materials and to K Enqvist and T Johansson (Västerbottens County Council, Umeå, Sweden) for DNA preparation. The current study was supported by Novo Nordisk (PWF), the Swedish Research Council (PWF), the Swedish Heart Lung Foundation (PWF), the European Research Council (PWF), and the Skåne Health Authority (PWF).

GRAPHIC (Genetic Regulation of Arterial Pressure of Humans in the Community): The GRAPHIC Study was funded by the British Heart Foundation (BHF/RG/2000004). This work falls under the portfolio of research supported by the NIHR Leicester Cardiovascular Biomedical Research Unit. CPN and NJS are funded by the BHF and NJS is a NIHR Senior Investigator.

HCHS/SOL (Hispanic Community Health Study/ Study of Latinos): The baseline examination of HCHS/SOL was supported by contracts from the National Heart, Lung, and Blood Institute (NHLBI) to the University of North Carolina (N01-HC65233), University of Miami (N01-HC65234), Albert Einstein College of Medicine (N01-HC65235), Northwestern University (N01-HC65236), and San Diego State University (N01-HC65237). The National Institute on Minority Health and Health Disparities, National Institute on Deafness and Other Communication Disorders, National Institute of Dental and Craniofacial Research (NIDCR), National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), National Institute of Neurological Disorders and Stroke, and NIH Office of Dietary Supplements additionally contributed funding to HCHS/SOL. The Genetic Analysis Center at the University of

Washington was supported by NHLBI and NIDCR contracts (HHSN268201300005C AM03 and MOD03). Additional analysis support was provided by 1R01DK101855-01 and 13GRNT16490017. Genotyping was also supported by National Center for Advancing Translational Sciences UL1TR000124 and NIDDK DK063491 to the Southern California Diabetes Endocrinology Research Center. This research was also supported in part by the Intramural Research Program of the NIDDK, contract no. HHSB268201200054C, and Illumina.

HRS (Health & Retirement Study): HRS is supported by the National Institute on Aging (NIA U01AG009740 and R03 AG046389). Genotyping was funded separately by NIA (RC2 AG036495, RC4 AG039029). Our genotyping was conducted by the NIH Center for Inherited Disease Research (CIDR) at Johns Hopkins University. Genotyping quality control and final preparation of the data were performed by the Genetics Coordinating Center at the University of Washington.

HyperGEN-AXIOM (Hypertension Genetic Epidemiology Network): The study was support by the National Institutes of Health, the National Heart, Lung, Blood Institute grant HL086718.

INGI-CARL (Italian Network Genetic Isolates): This study was partially supported by Regione FVG (L.26.2008) and Italian Ministry of Health (GR-2011-02349604).

INGI-FVG (Italian Network Genetic Isolates): This study was partially supported by Regione FVG (L.26.2008) and Italian Ministry of Health (GR-2011-02349604).

InterAct (The EPIC-InterAct Case-Cohort Study): We thank all EPIC participants and staff for their contribution to the study. The InterAct study received funding from the European Union (Integrated Project LSHM-CT-2006-037197 in the Framework Programme 6 of the European Community).

IRAS (Insulin Resistance Atherosclerosis Study): The IRAS is supported by the National Heart Lung Institute (HL047887, HL047889, HL047890, and HL47902). Genotyping for this study was supported by the GUARDIAN Consortium with grant support from the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK; DK085175) and in part by UL1TR000124 (CTSI) and DK063491 (DRC). The authors thank study investigators, staff, and participants for their valuable contributions.

IRAS Family Study (Insulin Resistance Atherosclerosis Study): The IRASFS is supported by the National Heart Lung and Blood Institute (HL060944, HL061019, and HL060919). Genotyping for this study was supported by the GUARDIAN Consortium with grant support from the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK; DK085175) and in part by UL1TR000124 (CTSI) and DK063491 (DRC). The authors thank study investigators, staff, and participants for their valuable contributions.

JUPITER (Justification for the Use of Statins in Primary Prevention: An Intervention Trial Evaluating Rosuvastatin): Support for genotype data collection and collaborative genetic analysis in JUPITER was provided by Astra-Zeneca.

KORA (Cooperative Health Research in the Augsburg Region): The KORA study was initiated and financed by the Helmholtz Zentrum München – German Research Center for Environmental Health, which is funded by the German Federal Ministry of Education and Research (BMBF) and by the State of Bavaria. Furthermore, KORA research was supported within the Munich Center of Health Sciences (MC-Health), Ludwig-Maximilians-Universität, as part of LMUinnovativ.

LBC1921 (Lothian Birth Cohort 1921): We thank the LBC1921 cohort participants and team members who contributed to these studies. Phenotype collection was supported by the UK's Biotechnology and Biological Sciences Research Council (BBSRC), The Royal Society, and The Chief Scientist Office of

the Scottish Government. Genotyping was funded by the BBSRC (BB/F019394/1). The work was undertaken by The University of Edinburgh Centre for Cognitive Ageing and Cognitive Epidemiology, part of the cross council Lifelong Health and Wellbeing Initiative (MR/K026992/1). Funding from the BBSRC and Medical Research Council (MRC) is gratefully acknowledged.

LBC1936 (Lothian Birth Cohort 1936): We thank the LBC1936 cohort participants and team members who contributed to these studies. Phenotype collection was supported by Age UK (The Disconnected Mind project). Genotyping was funded by the BBSRC (BB/F019394/1). The work was undertaken by The University of Edinburgh Centre for Cognitive Ageing and Cognitive Epidemiology, part of the cross council Lifelong Health and Wellbeing Initiative (MR/K026992/1). Funding from the BBSRC and Medical Research Council (MRC) is gratefully acknowledged.

LifeLines (Netherlands Biobank): The Lifelines Cohort Study, and generation and management of GWAS genotype data for the Lifelines Cohort Study is supported by the Netherlands Organization of Scientific Research NWO (grant 175.010.2007.006), the Economic Structure Enhancing Fund (FES) of the Dutch government, the Ministry of Economic Affairs, the Ministry of Education, Culture and Science, the Ministry for Health, Welfare and Sports, the Northern Netherlands Collaboration of Provinces (SNN), the Province of Groningen, University Medical Center Groningen, the University of Groningen, Dutch Kidney Foundation and Dutch Diabetes Research Foundation.

The authors wish to acknowledge the services of the Lifelines Cohort Study, the contributing research centers delivering data to Lifelines, and all the study participants.

LLFS (The Long Life Family Study): The study is supported by the National Institute on Aging (NIA) grant U01AG023746.

LOLIPOP (London Life Sciences Prospective Population Study): The LOLIPOP study is supported by the National Institute for Health Research (NIHR) Comprehensive Biomedical Research Centre Imperial College Healthcare NHS Trust, the British Heart Foundation (SP/04/002), the Medical Research Council (G0601966, G0700931), the Wellcome Trust (084723/Z/08/Z), the NIHR (RP-PG-0407-10371), European Union FP7 (EpiMigrant, 279143), and Action on Hearing Loss (G51). We thank the participants and research staff who made the study possible.

Loyola GxE (Kingston Gene-by-environment; subset of International Collaborative Study of Hypertension in Blacks (ICSHIB)): The Loyola GxE project was supported by NIH Grant R01HL53353.

Loyola SPT (Spanish Town; subset of International Collaborative Study of Hypertension in Blacks (ICSHIB)): The Loyola SPT project was supported by NIH Grant R01HL53353.

METSIM (Metabolic Syndrome In Men): The METSIM study was supported by the Academy of Finland (contract 124243), the Finnish Heart Foundation, the Finnish Diabetes Foundation, Tekes (contract 1510/31/06), and the Commission of the European Community (HEALTH-F2-2007 201681), and the US National Institutes of Health grants DK093757, DK072193, DK062370, and ZIA-HG000024. Genotyping was conducted at the Genetic Resources Core Facility (GRCF) at the Johns Hopkins Institute of Genetic Medicine.

NESDA (Netherlands Study of Depression and Anxiety): The infrastructure for the NESDA study is funded through the Geestkracht programme of the Dutch Scientific Organization (ZON-MW, grant number 10-000-1002) and matching funds from participating universities and mental health care organizations. Genotyping in NESDA was funded by the Genetic Association Information Network (GAIN) of the Foundation for the US National Institutes of Health. Statistical analyses were carried out on the Genetic Cluster Computer (http://www.geneticcluster.org), which is financially supported by the

Netherlands Scientific Organization (NWO 480-05-003) along with a supplement from the Dutch Brain Foundation.

OBA (French obese cases): The obese French adults were recruited by the laboratory "Integrated Genomics and Metabolic Diseases Modeling" (UMR 8199 CNRS / Université de Lille 2 / Institut Pasteur de Lille) of Pr. Philippe Froguel.

PROCARDIS (Precocious Coronary Artery Disease): PROCARDIS was supported by the European Community Sixth Framework Program (LSHM-CT- 2007-037273), AstraZeneca, the British Heart Foundation, the Swedish Research Council, the Knut and Alice Wallenberg Foundation, the Swedish Heart-Lung Foundation, the Torsten and Ragnar Söderberg Foundation, the Strategic Cardiovascular Program of Karolinska Institutet and Stockholm County Council, the Foundation for Strategic Research and the Stockholm County Council (560283). M.F and H.W acknowledge the support of the Wellcome Trust core award (090532/Z/09/Z) and the BHF Centre of Research Excellence. A.G and H.W acknowledge European Union Seventh Framework Programme FP7/2007-2013 under grant agreement no. HEALTH-F2-2013-601456 (CVGenes@Target) & and A.G, the Wellcome Trust Institutional strategic support fund.

RHS (Ragama Health Study): The RHS was supported by the Grant of National Center for Global Health and Medicine (NCGM).

SHEEP (Stockholm Heart Epidemiology Project): This study was supported by grants from the Swedish Research Council for Health, Working Life and Welfare (http://www.forte.se/en/), the Stockholm County Council (http://www.sll.se/om-landstinget/Information-in-English1/), the Swedish Research Council (http://www.vr.se/inenglish.4.12fff4451215cbd83e4800015152.html), the Swedish Heart and Lung Foundation (https://www.hjart-lungfonden.se/HLF/Om-Hjart-lungfonden/About-HLF/), and the Cardiovascular Programme at Karolinska Institutet (http://ki.se/en/mmk/cardiovascular-research-networks).

SHIP (Study of Health in Pomerania): SHIP is part of the Community Medicine Research net of the University of Greifswald, Germany, which is funded by the Federal Ministry of Education and Research (grants no. 01ZZ9603, 01ZZ0103, and 01ZZ0403), the Ministry of Cultural Affairs as well as the Social Ministry of the Federal State of Mecklenburg-West Pomerania, and the network ‘Greifswald Approach to Individualized Medicine (GANI_MED)’ funded by the Federal Ministry of Education and Research (grant 03IS2061A). Genome-wide data were supported by the Federal Ministry of Education and Research (grant no. 03ZIK012) and a joint grant from Siemens Healthcare, Erlangen, Germany and the Federal State of Mecklenburg- West Pomerania. The University of Greifswald is a member of the ‘Center of Knowledge Interchange’ program of the Siemens AG.

SWHS/SMHS (Shanghai Women's Health Study/ Shanghai Men's Health Study): We thank all the individuals who took part in these studies and all the researchers who have enabled this work to be carried out. The Shanghai Women’s Health Study and the Shanghai Men’s Health Study are supported by research grants UM1CA182910 and UM1CA173640 from the U.S. National Cancer Institute, respectively.

TAICHI_G: This study was supported by the National Health Research Institutes, Taiwan (PH-100-SP-01, BS-094-PP-01, PH-100-PP-03), the National Science Council, Taiwan (Grant Nos NSC 98-2314-B-075A-002-MY3, NSC 96-2314-B-002-151, NSC 96-2314-B-002-152, NSC 98-2314-B-002-122-MY2, NSC 100-2314-B-002-115, NSC 101-2325-002-078, 101-2314-B-075A-006-MY3), the National Taiwan University Hospital, Taiwan (NTUH 98-N1266, NTUH 100-N1775, NTUH 101-N2010, NTUH 101-N, VN101-04, NTUH 101-S1784).

THRV (Taiwan study of Hypertensives Rare Variants): The THRV study is supported by National Heart, Lung, and Blood Institute grant R01HL111249.

TRAILS (Tracking Adolescents' Individual Lives Survey): TRAILS (TRacking Adolescents' Individual Lives Survey) is a collaborative project involving various departments of the University Medical Center and University of Groningen, the Erasmus University Medical Center Rotterdam, the University of Utrecht, the Radboud Medical Center Nijmegen, and the Parnassia Bavo group, all in the Netherlands. TRAILS has been financially supported by grants from the Netherlands Organization for Scientific Research NWO (Medical Research Council program grant GB-MW 940-38-011; ZonMW Brainpower grant 100-001-004; ZonMw Risk Behavior and Dependence grant 60-60600-97-118; ZonMw Culture and Health grant 261-98-710; Social Sciences Council medium-sized investment grants GB-MaGW 480-01-006 and GB-MaGW 480-07-001; Social Sciences Council project grants GB-MaGW 452-04-314 and GB-MaGW 452-06-004; NWO large-sized investment grant 175.010.2003.005; NWO Longitudinal Survey and Panel Funding 481-08-013); the Dutch Ministry of Justice (WODC), the European Science Foundation (EuroSTRESS project FP-006), Biobanking and Biomolecular Resources Research Infrastructure BBMRI-NL (CP 32), the participating universities, and Accare Center for Child and Adolescent Psychiatry. Statistical analyses were carried out on the Genetic Cluster Computer (http://www.geneticcluster.org), which is financially supported by the Netherlands Scientific Organization (NWO 480-05-003) along with a supplement from the Dutch Brain Foundation.

We are grateful to all adolescents who participated in this research and to everyone who worked on this project and made it possible.

TUDR (Taiwan-US Diabetic Retinopathy): This study was supported by the National Eye Institute of the National Institutes of Health (EY014684 to J.I.R. and Y.-D.I.C.) and ARRA Supplement (EY014684-03S1, -04S1), the Eye Birth Defects Foundation Inc., the National Science Council, Taiwan (NSC 98-2314-B-075A-002-MY3 to W.H.S.) and the Taichung Veterans General Hospital, Taichung, Taiwan (TCVGH-1003001C to W.H.S .). DNA handling and genotyping were supported in part by the National Center for Advancing Translational Sciences, CTSI grant UL1TR000124 and the National Institute of Diabetes and Digestive and Kidney Disease Diabetes Research Center (DRC) grant DK063491 to the Southern California Diabetes Endocrinology Research Center.

TWINGENE (TwinGene of the Swedish Twin Registry): The Swedish Twin Registry is financially supported by Karolinska Institutet. TwinGene project received funding from the Swedish Research Council (M-2005-1112), GenomEUtwin (EU/QLRT-2001-01254; QLG2-CT-2002-01254), NIH DK U01-066134, The Swedish Foundation for Strategic Research (SSF) and the Heart and Lung foundation no. 20070481

UKB (United Kingdom Biobank, www.ukbiobank.ac.uk): This research has been conducted using the UK Biobank Resource. The UK Biobank data was analysed from the data set corresponding to UK Biobank access application no. 236, application title "Genome-wide association study of blood pressure", with Paul Elliott as the PI/applicant. Central analysts from Queen Mary University of London (QMUL) and Imperial College London are funded by the NIHR Cardiovascular Biomedical Research Unit at Barts and The London School of Medicine, QMUL, and the NIHR Imperial College Health Care NHS Trust and Imperial College London Biomedical Research Centre respectively. This work was supported by the UK Biobank CardioMetabolic Consortium (UKB-CMC) and the BP working group. The UKB-CMC was supported by the British Heart Foundation (grant SP/13/2/30111).

UKHLS (Understanding Society / The UK Household Longitudinal Study): These data are from Understanding Society: The UK Household Longitudinal Study, which is led by the Institute for Social and Economic Research at the University of Essex and funded by the Economic and Social Research Council. The data were collected by NatCen and the genome wide scan data were analysed by the

Wellcome Trust Sanger Institute. The Understanding Society DAC have an application system for genetics data and all use of the data should be approved by them. The application form is at:

https://www.understandingsociety.ac.uk/about/health/data.

YFS (The Cardiovascular Risk in Young Finns Study): The Young Finns Study has been financially supported by the Academy of Finland: grants 286284, 134309 (Eye), 126925, 121584, 124282, 129378 (Salve), 117787 (Gendi), and 41071 (Skidi); the Social Insurance Institution of Finland; Kuopio, Tampere and Turku University Hospital Medical Funds (grant X51001); Juho Vainio Foundation; Paavo Nurmi Foundation; Finnish Foundation for Cardiovascular Research ; Finnish Cultural Foundation; Tampere Tuberculosis Foundation; Emil Aaltonen Foundation; Yrjö Jahnsson Foundation; Signe and Ane Gyllenberg Foundation; and Diabetes Research Foundation of Finnish Diabetes Association.

The expert technical assistance in the statistical analyses by Leo-Pekka Lyytikäinen and Irina Lisinen is gratefully acknowledged.

NOTE: Baependi, NEO, Pelotas, and WHI (EA) also participated in replications since they did not contribute to Smoking-BP discovery analysis.

Supplemental Figures

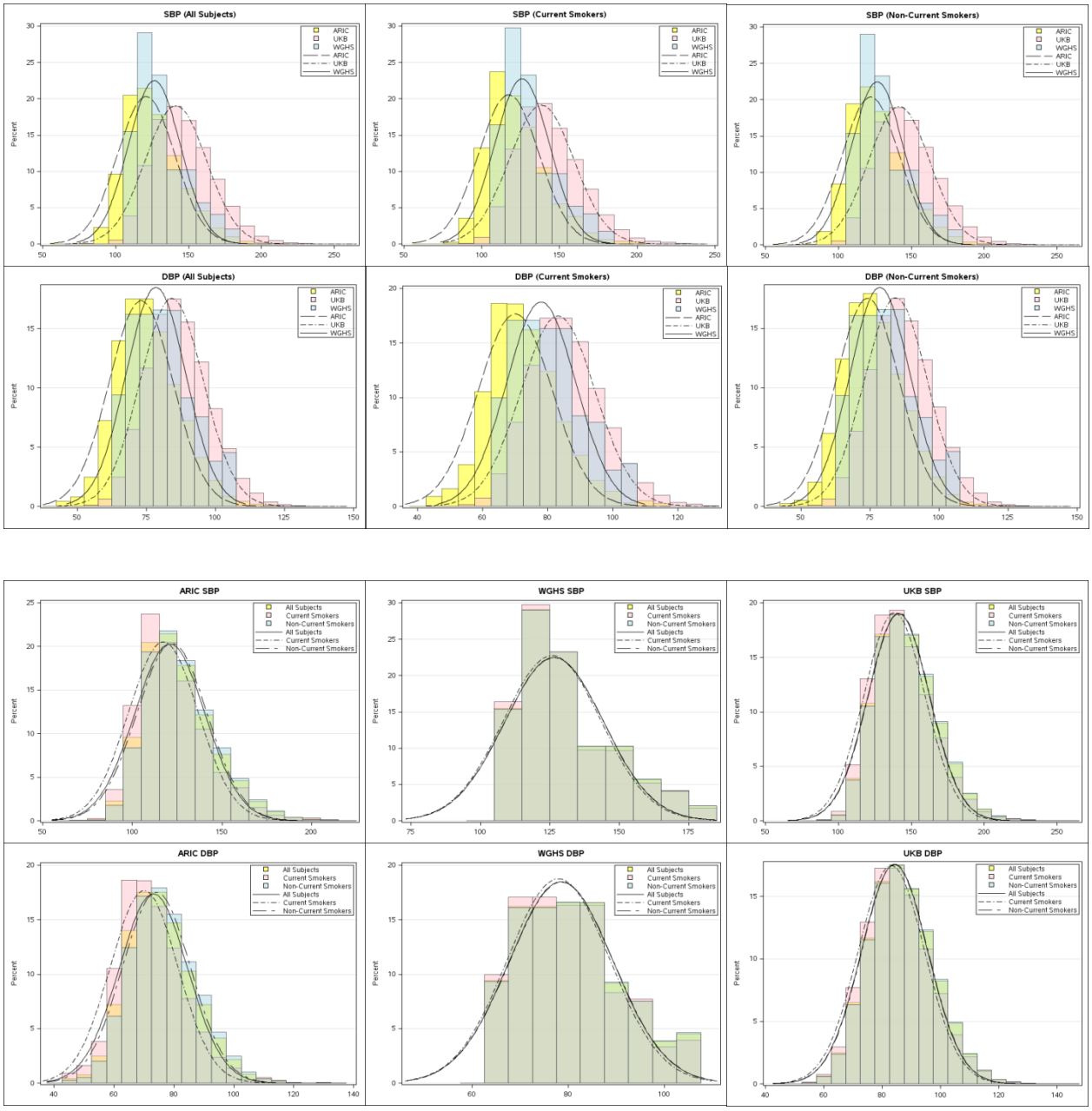


Figure S1: BP distributions across 3 large cohorts (ARIC, WGHS, and UK Biobank).

The top 6 panels (a panel for each smoking status) show some variations among the cohorts. This is because there are variations in covariates, which are adjusted within each cohort. The bottom 6 panels (a panel for each cohort) show almost identical distributions across smoking status within each cohort.

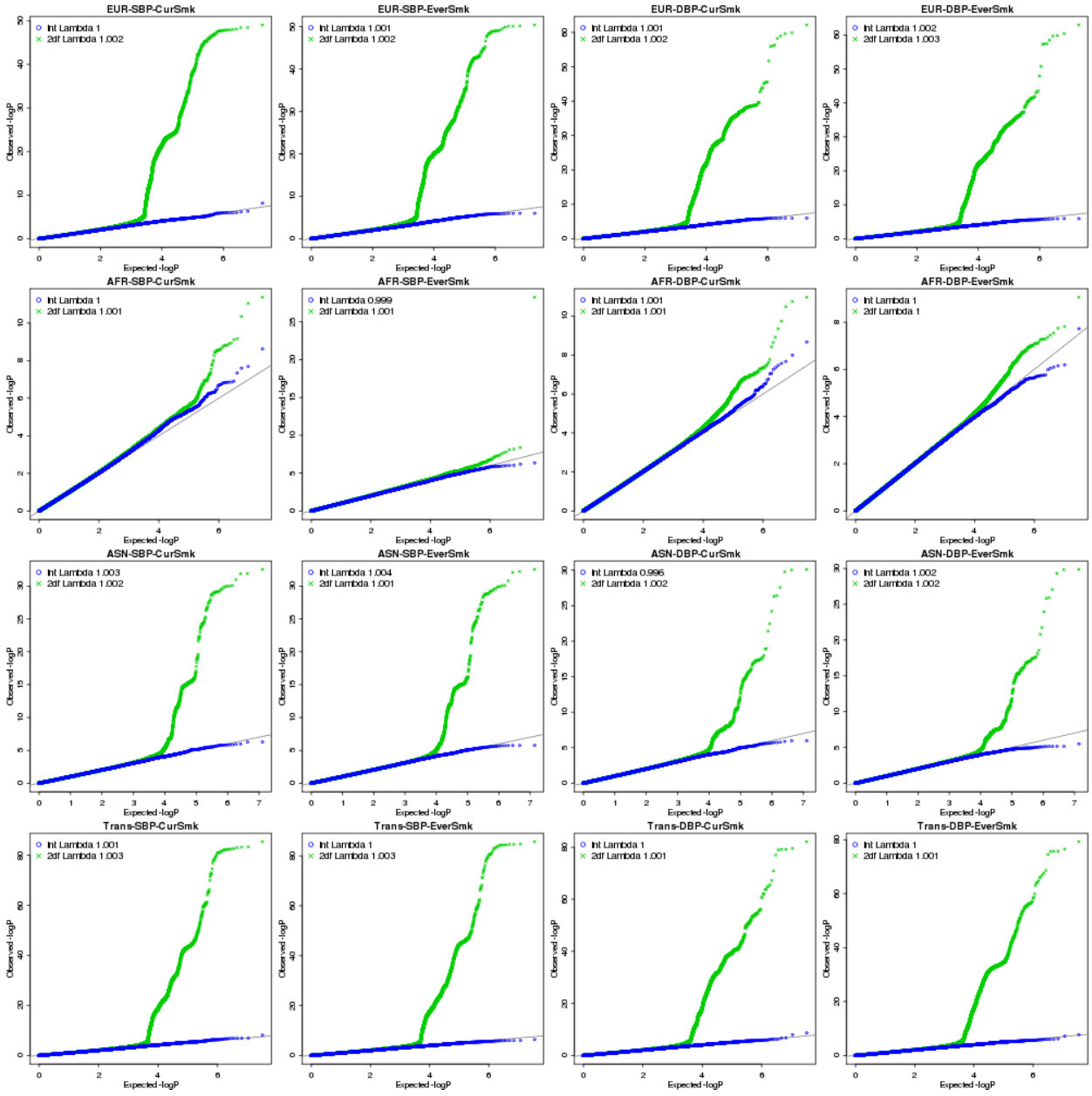


Figure S2: QQ plots of the Stage 1 discovery meta-analyses.

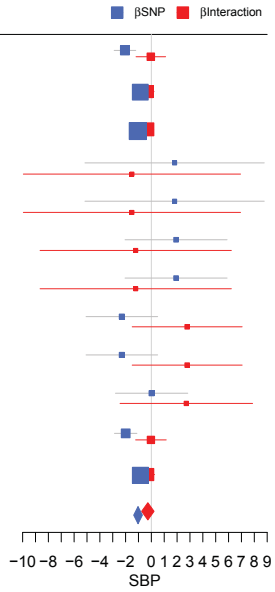
The combination of BP traits and smoking exposures were used: SBP-CurSmk (1st column), SBP-EverSmk (2nd column), DBP-CurSmk (3rd column), and DBP-EverSmk (4th column). Each plot displays p-values (blue circles for the 1 DF test of interaction effect; green crosses for the 2 DF joint test) and their genomic inflation factor.

Figure S3: Forest plots that examine consistencies between Stage 1 and 2 and across ancestries

Forest plots are ordered by tables (2-5) then by loci within each table. If both traits reach genome-wide significance at a locus, then the most significant result is shown.

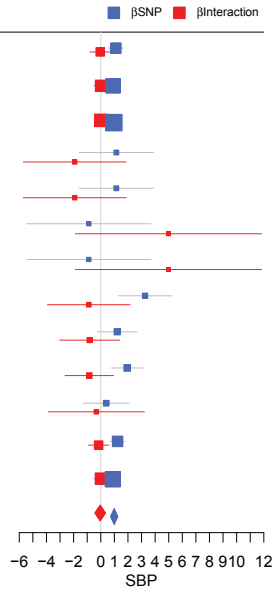
Effect of rs72640287 (T2−L3) and its interaction with EverSmk on SBP

Study	N	EAF	2df P
EUR S1	73913	0.039	1.585e−10
EUR S2	261161	0.04	8.185e−13
EUR S1+2	335074	0.039	2.787e−21
AFR S1	7928	0.01	0.8725
AFR S1+2	7928	0.01	0.8725
ASN S2	6548	0.019	0.6217
ASN S1+2	6548	0.019	0.6217
HIS S2	12380	0.01	0.2583
HIS S1+2	12380	0.01	0.2583
BRZ S2	3541	0.016	0.4621
Trans S1	81841	0.036	5.913e−10
Trans S2	283630	0.038	2.898e−12
Trans S1+2	**365471**	**0.037**	**2.164e−20**



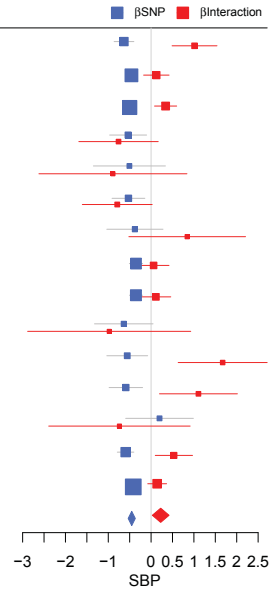
Effect of rs202071545 (T2−L1) and its interaction with EverSmk on SBP

Study	N	EAF	2df P
EUR S1	75450	0.933	1.079e−06
EUR S2	210600	0.946	1.426e−11
EUR S1+2	286050	0.943	6.049e−18
AFR S1	15485	0.976	0.604
AFR S1+2	15485	0.976	0.604
ASN S2	2468	0.967	0.3315
ASN S1+2	2468	0.967	0.3315
HIS S1	8742	0.955	0.0008513
HIS S2	13348	0.964	0.2583
HIS S1+2	22090	0.961	0.0018
BRZ S2	3541	0.954	0.8957
Trans S1	102168	0.942	3.774e−08
Trans S2	229957	0.947	7.44e−12
Trans S1+2	**329634**	**0.946**	**9.394e−20**



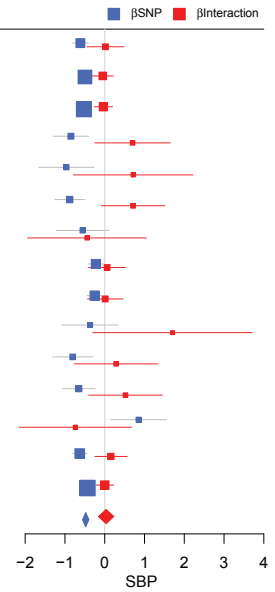
Effect of rs351364 (T2−L4) and its interaction with CurSmk on SBP

Study	N	EAF	2df P
EUR S1	79732	0.265	4.039e−07
EUR S2	289723	0.26	7.252e−14
EUR S1+2	369455	0.261	6.386e−19
AFR S1	25821	0.321	0.000748
AFR S2	6269	0.322	0.1654
AFR S1+2	32090	0.321	7.599e−05
ASN S1	11546	0.283	0.4004
ASN S2	148143	0.373	1.278e−05
ASN S1+2	159689	0.366	1.025e−05
HIS S1	8742	0.345	0.04215
HIS S2	13337	0.367	0.003939
HIS S1+2	22079	0.358	0.005371
BRZ S2	4414	0.275	0.6778
Trans S1	125841	0.284	1.672e−08
Trans S2	461886	0.3	5.378e−19
Trans S1+2	**587727**	**0.297**	**1.195e−26**



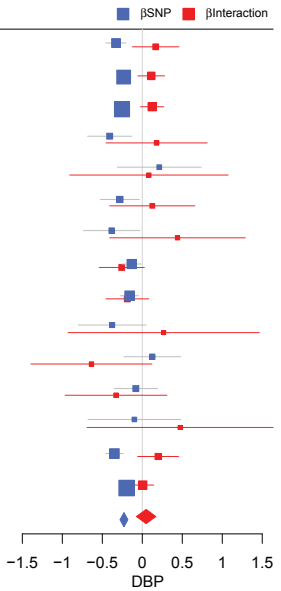
Effect of rs3753581 (T2−L2) and its interaction with CurSmk on SBP

Study	N	EAF	2df P
EUR S1	79732	0.373	3.435e−09
EUR S2	291464	0.362	2.756e−22
EUR S1+2	371196	0.364	3.497e−31
AFR S1	25821	0.62	0.001281
AFR S2	7785	0.643	0.02497
AFR S1+2	33606	0.625	2.098e−05
ASN S1	10798	0.275	0.09832
ASN S2	148099	0.174	0.06717
ASN S1+2	158897	0.181	0.01873
HIS S1	8742	0.349	0.228
HIS S2	13337	0.313	0.004836
HIS S1+2	22079	0.327	0.007626
BRZ S2	4414	0.4	0.0582
Trans S1	125093	0.414	4.337e−12
Trans S2	465099	0.306	5.522e−23
Trans S1+2	**590192**	**0.329**	**1.311e−34**



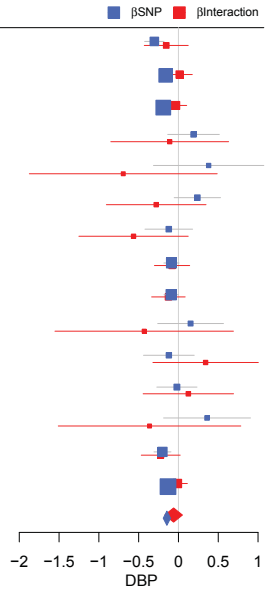
Effect of rs3897821 (T2−L5*) and its interaction with CurSmk on DBP

Study	N	EAF	2df P
EUR S1	79732	0.664	3.804e−06
EUR S2	284501	0.666	1.181e−11
EUR S1+2	364233	0.666	4.029e−17
AFR S1	25818	0.757	0.01474
AFR S2	6269	0.764	0.6781
AFR S1+2	32087	0.758	0.07042
ASN S1	11546	0.682	0.1076
ASN S2	148642	0.783	0.001257
ASN S1+2	160188	0.776	0.0005438
HIS S1	8742	0.756	0.2241
HIS S2	13337	0.763	0.2573
HIS S1+2	22079	0.76	0.326
BRZ S2	4414	0.685	0.7292
Trans S1	125838	0.691	2.485e−09
Trans S2	457163	0.708	1.506e−12
Trans S1+2	**583001**	**0.705**	**1.67e−20**



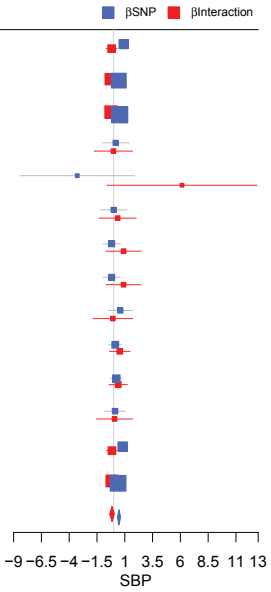
Effect of rs7599598 (T2−L6*) and its interaction with CurSmk on DBP

Study	N	EAF	2df P
EUR S1	79732	0.573	5.93e−08
EUR S2	286555	0.562	4.099e−07
EUR S1+2	366287	0.564	4.253e−13
AFR S1	24132	0.213	0.5303
AFR S2	5708	0.142	0.3566
AFR S1+2	29840	0.2	0.2657
ASN S1	13438	0.579	0.07915
ASN S2	148929	0.601	0.05323
ASN S1+2	162367	0.6	0.01312
HIS S1	8742	0.587	0.6694
HIS S2	13337	0.63	0.5651
HIS S1+2	22079	0.613	0.9122
BRZ S2	4414	0.591	0.4342
Trans S1	126044	0.506	1.607e−06
Trans S2	458943	0.572	1.825e−07
Trans S1+2	**584987**	**0.558**	**2.013e−12**



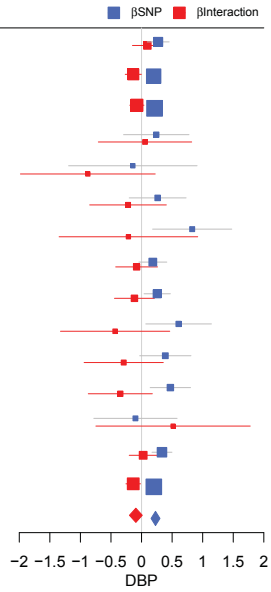
Effect of rs62493780 (T2−L9*) and its interaction with EverSmk on SBP

Study	N	EAF	2df P
EUR S1	77241	0.236	3.471e−08
EUR S2	269826	0.239	2.374e−07
EUR S1+2	347067	0.238	2.948e−13
AFR S1	25204	0.062	0.9429
AFR S2	1993	0.054	0.2008
AFR S1+2	27197	0.061	0.8519
ASN S2	17405	0.175	0.5448
ASN S1+2	17405	0.175	0.5448
HIS S1	8742	0.219	0.4423
HIS S2	13529	0.253	0.1627
HIS S1+2	22271	0.24	0.09371
BRZ S2	3541	0.22	0.923
Trans S1	111187	0.195	5.532e−08
Trans S2	306294	0.234	3.739e−07
Trans S1+2	**417481**	**0.224**	**3.324e−13**



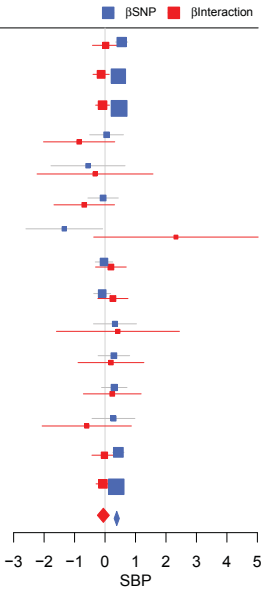
Effect of rs13063291 (T2−L7) and its interaction with EverSmk on DBP

Study	N	EAF	2df P
EUR S1	79732	0.263	4.848e−06
EUR S2	275207	0.272	4.981e−05
EUR S1+2	354939	0.27	1.133e−08
AFR S1	25818	0.085	0.42
AFR S2	5568	0.077	0.1805
AFR S1+2	31386	0.084	0.5395
ASN S1	11546	0.091	0.02114
ASN S2	145631	0.07	0.2095
ASN S1+2	157177	0.072	0.03621
HIS S1	8742	0.204	0.08251
HIS S2	13529	0.203	0.1817
HIS S1+2	22271	0.204	0.0174
BRZ S2	4414	0.215	0.713
Trans S1	125838	0.207	4.002e−08
Trans S2	444349	0.201	3.751e−06
Trans S1+2	**570187**	**0.202**	**1.668e−11**



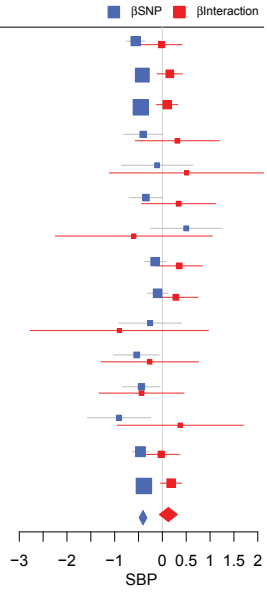
Effect of rs13271489 (T2−L10*) and its interaction with CurSmk on SBP

Study	N	EAF	2df P
EUR S1	79732	0.476	6.372e−08
EUR S2	284854	0.478	9.349e−16
EUR S1+2	364586	0.478	4.562e−23
AFR S1	25821	0.802	0.3378
AFR S2	6269	0.85	0.5739
AFR S1+2	32090	0.811	0.2752
ASN S1	6857	0.85	0.0855
ASN S2	141234	0.851	0.7563
ASN S1+2	148091	0.851	0.5568
HIS S1	8742	0.675	0.5035
HIS S2	13337	0.7	0.3579
HIS S1+2	22079	0.69	0.1844
BRZ S2	4414	0.574	0.6578
Trans S1	121152	0.581	3.712e−06
Trans S2	450108	0.608	6.56e−14
Trans S1+2	**571260**	**0.602**	**9.461e−20**



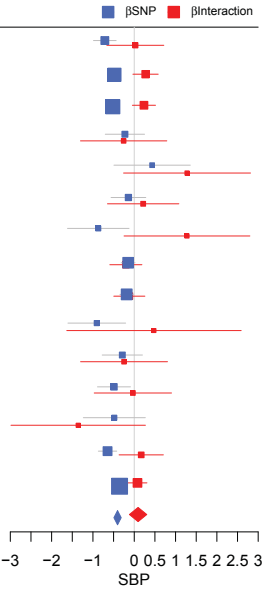
Effect of rs7823056 (T2−L8*) and its interaction with CurSmk on SBP

Study	N	EAF	2df P
EUR S1	79732	0.51	1.537e−08
EUR S2	278186	0.5	1.545e−14
EUR S1+2	357918	0.502	3.007e−22
AFR S1	25821	0.443	0.1614
AFR S2	5792	0.464	0.8262
AFR S1+2	31613	0.447	0.1667
ASN S1	9654	0.216	0.4305
ASN S2	148099	0.157	0.2512
ASN S1+2	157753	0.16	0.4372
HIS S1	8742	0.383	0.3315
HIS S2	13337	0.34	0.02791
HIS S1+2	22079	0.357	0.0122
BRZ S2	4414	0.455	0.0182
Trans S1	123949	0.464	5.664e−08
Trans S2	449828	0.381	9.154e−16
Trans S1+2	**573777**	**0.399**	**3.119e−23**



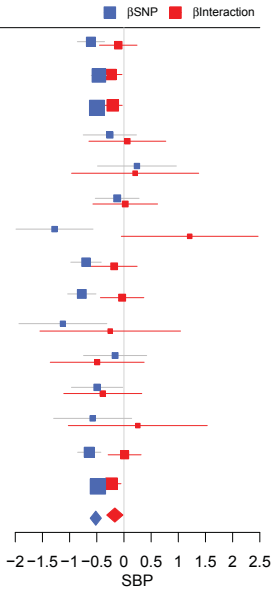
Effect of rs7483477 (T2−L11) and its interaction with CurSmk on SBP

Study	N	EAF	2df P
EUR S1	65647	0.756	7.863e−07
EUR S2	287900	0.754	1.357e−13
EUR S1+2	353547	0.755	2.831e−19
AFR S1	25821	0.771	0.4175
AFR S2	6269	0.797	0.08694
AFR S1+2	32090	0.776	0.7824
ASN S1	10798	0.786	0.06405
ASN S2	147692	0.742	0.03371
ASN S1+2	158490	0.745	0.02049
HIS S1	8742	0.688	0.03764
HIS S2	13337	0.67	0.2835
HIS S1+2	22079	0.677	0.02824
BRZ S2	4414	0.737	0.02045
Trans S1	111008	0.757	2.252e−08
Trans S2	459612	0.748	1.766e−13
Trans S1+2	**570620**	**0.75**	**2.119e−20**



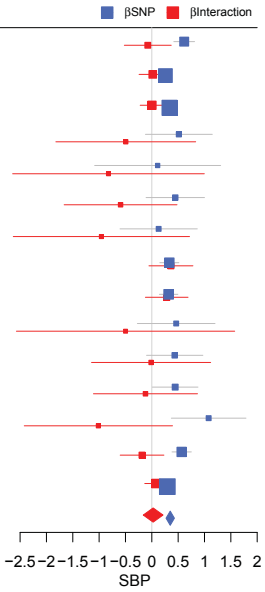
Effect of rs7313874 (T2−L12) and its interaction with EverSmk on SBP

Study	N	EAF	2df P
EUR S1	79732	0.398	1.591e−11
EUR S2	276436	0.386	1.73e−29
EUR S1+2	356168	0.388	4.117e−41
AFR S1	25821	0.478	0.4651
AFR S2	6180	0.47	0.6562
AFR S1+2	32001	0.477	0.772
ASN S1	11546	0.239	0.002175
ASN S2	148931	0.142	8.383e−12
ASN S1+2	160477	0.15	1.46e−13
HIS S1	8742	0.332	0.00104
HIS S2	13529	0.298	0.1502
HIS S1+2	22271	0.311	0.001474
BRZ S2	4414	0.367	0.2443
Trans S1	125841	0.395	1.85e−14
Trans S2	449490	0.303	1.07e−39
Trans S1+2	**575331**	**0.324**	**2.486e−54**



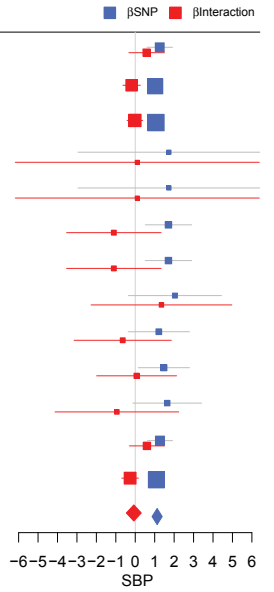
Effect of rs11642015 (T2−L15*) and its interaction with CurSmk on SBP

Study	N	EAF	2df P
EUR S1	79732	0.409	1.742e−09
EUR S2	291464	0.411	7.223e−07
EUR S1+2	371196	0.411	5.967e−14
AFR S1	25204	0.112	0.3082
AFR S2	7173	0.09	0.6743
AFR S1+2	32377	0.107	0.2564
ASN S1	10798	0.214	0.5333
ASN S2	148099	0.214	2.794e−06
ASN S1+2	158897	0.214	1.109e−05
HIS S1	8742	0.269	0.473
HIS S2	13337	0.25	0.213
HIS S1+2	22079	0.257	0.1086
BRZ S2	4414	0.367	0.01158
Trans S1	124476	0.322	2.783e−09
Trans S2	464487	0.338	6.742e−13
Trans S1+2	**588963**	**0.335**	**9.905e−21**



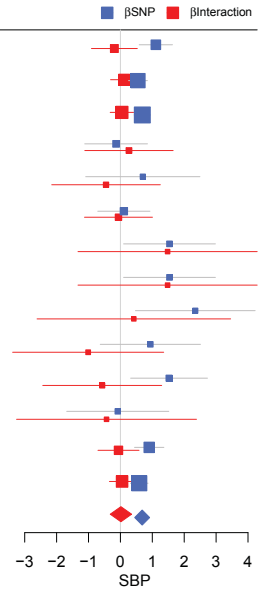
Effect of rs111337717 (T2−L13) and its interaction with EverSmk on SBP

Study	N	EAF	2df P
EUR S1	74153	0.943	1.221e−09
EUR S2	272302	0.943	1.454e−15
EUR S1+2	346455	0.943	1.373e−23
AFR S1	10790	0.985	0.5348
AFR S1+2	10790	0.985	0.5348
ASN S2	10007	0.88	0.01669
ASN S1+2	10007	0.88	0.01669
HIS S1	8742	0.964	0.01346
HIS S2	12380	0.966	0.2777
HIS S1+2	21122	0.966	0.01575
BRZ S2	4414	0.957	0.1747
Trans S1	96176	0.948	9.234e−11
Trans S2	299103	0.942	2.863e−18
Trans S1+2	**392788**	**0.944**	**1.269e−27**



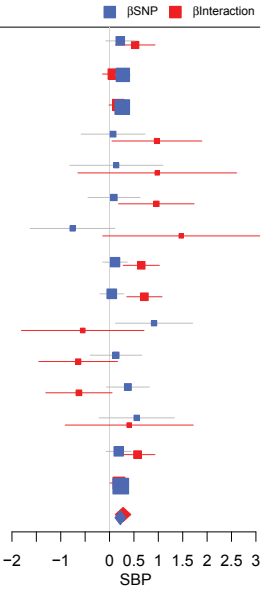
Effect of rs12741980 (T3−L1) and its interaction with EverSmk on SBP

Study	N	EAF	2df P
EUR S1	78762	0.939	1.384e−06
EUR S2	283471	0.944	1.651e−07
EUR S1+2	362233	0.943	4.782e−13
AFR S1	25821	0.928	0.9346
AFR S2	5150	0.936	0.7048
AFR S1+2	30971	0.93	0.965
ASN S2	11075	0.95	0.01201
ASN S1+2	11075	0.95	0.01201
HIS S1	8742	0.961	0.004193
HIS S2	12380	0.966	0.5011
HIS S1+2	21122	0.964	0.02012
BRZ S2	4414	0.947	0.9107
Trans S1	115816	0.939	1.663e−06
Trans S2	316490	0.945	2.436e−08
Trans S1+2	**429815**	**0.943**	**3.042e−14**



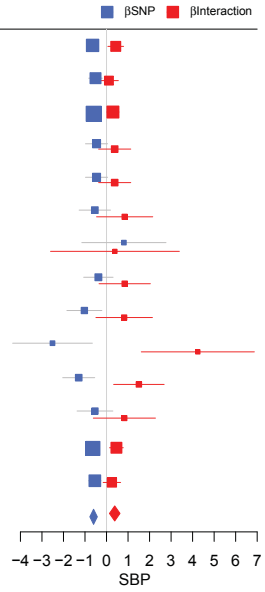
Effect of rs7974266 (T2−L14) and its interaction with EverSmk on SBP

Study	N	EAF	2df P
EUR S1	79732	0.695	2.107e−06
EUR S2	271525	0.699	2.088e−07
EUR S1+2	351257	0.698	2.512e−12
AFR S1	25821	0.201	0.01247
AFR S2	6180	0.176	0.4156
AFR S1+2	32001	0.196	0.003923
ASN S1	11546	0.146	0.1418
ASN S2	148644	0.157	3.296e−07
ASN S1+2	160190	0.156	2.644e−07
HIS S1	8742	0.519	0.06641
HIS S2	13529	0.514	0.2486
HIS S1+2	22271	0.516	0.1498
BRZ S2	4414	0.628	0.08385
Trans S1	125841	0.531	3.578e−08
Trans S2	444292	0.504	6.499e−12
Trans S1+2	**570133**	**0.51**	**5.914e−19**



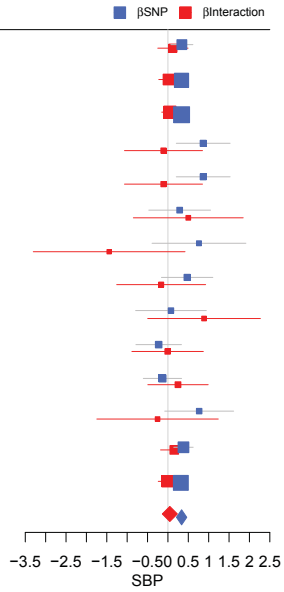
Effect of rs201851995 (T3−L2*) and its interaction with EverSmk on SBP

Study	N	EAF	2df P
EUR S1	76420	0.63	1.052e−05
EUR S2	55319	0.656	0.0001329
EUR S1+2	131739	0.641	8.905e−10
AFR S1	24685	0.609	0.2007
AFR S1+2	24685	0.609	0.2007
ASN S1	11546	0.774	0.3052
ASN S2	2552	0.815	0.4214
ASN S1+2	14098	0.782	0.3585
HIS S1	8742	0.658	0.04543
HIS S2	1149	0.686	0.004578
HIS S1+2	9891	0.662	0.002663
BRZ S2	3541	0.646	0.3994
Trans S1	121393	0.641	1.435e−07
Trans S2	62561	0.662	0.0001094
Trans S1+2	**183954**	**0.648**	**4.646e−12**



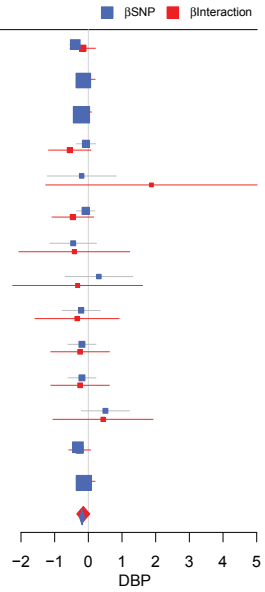
Effect of rs78763922 (T3−L3*) and its interaction with EverSmk on SBP

Study	N	EAF	2df P
EUR S1	76420	0.328	0.0002915
EUR S2	210883	0.318	1.415e−07
EUR S1+2	287303	0.32	2.507e−11
AFR S1	24685	0.172	0.004321
AFR S1+2	24685	0.172	0.004321
ASN S1	11546	0.223	0.3143
ASN S2	5894	0.154	0.2451
ASN S1+2	17440	0.2	0.2982
HIS S1	8742	0.294	0.2241
HIS S2	13529	0.308	0.5919
HIS S1+2	22271	0.303	0.7915
BRZ S2	3541	0.3	0.1486
Trans S1	121393	0.284	1.197e−06
Trans S2	233847	0.313	3.637e−07
Trans S1+2	**355240**	**0.303**	**4.026e−13**



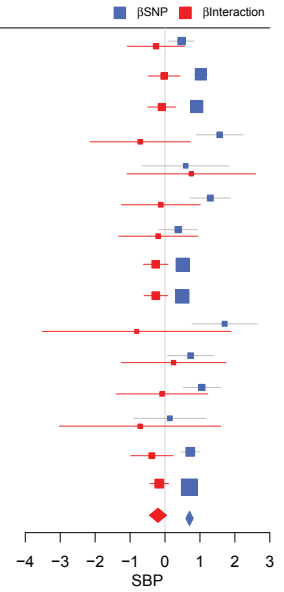
Effect of rs35904419 (T3−L6*) and its interaction with CurSmk on DBP

Study	N	EAF	2df P
EUR S1	75921	0.808	1.044e−06
EUR S2	216460	0.822	0.00208
EUR S1+2	292381	0.819	4.496e−08
AFR S1	24682	0.734	0.1031
AFR S2	1516	0.723	0.5043
AFR S1+2	26198	0.734	0.1407
ASN S1	8705	0.896	0.2421
ASN S2	4748	0.926	0.8259
ASN S1+2	13453	0.907	0.5013
HIS S2	13337	0.84	0.3781
HIS S1+2	13337	0.84	0.3781
BRZ S2	3541	0.792	0.1408
Trans S1	114077	0.8	2.291e−07
Trans S2	239602	0.824	0.003798
Trans S1+2	**348910**	**0.816**	**1.338e−08**



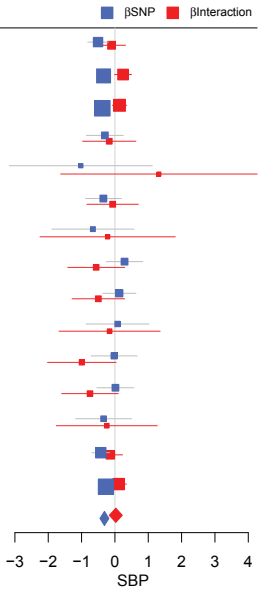
Effect of rs2023843 (T3−L4) and its interaction with CurSmk on SBP

Study	N	EAF	2df P
EUR S1	77753	0.926	0.04404
EUR S2	295747	0.923	1.121e−29
EUR S1+2	373500	0.924	6e−29
AFR S1	25087	0.889	2.53e−05
AFR S2	6612	0.896	0.38
AFR S1+2	31699	0.89	1.708e−05
ASN S1	13438	0.678	0.3969
ASN S2	147531	0.619	2.864e−10
ASN S1+2	160969	0.624	1.244e−10
HIS S1	8742	0.876	0.001428
HIS S2	13337	0.868	0.0387
HIS S1+2	22079	0.871	0.0002041
BRZ S2	4414	0.885	0.8311
Trans S1	125020	0.888	9.385e−07
Trans S2	467641	0.825	8.976e−35
Trans S1+2	**592661**	**0.838**	**3.688e−41**



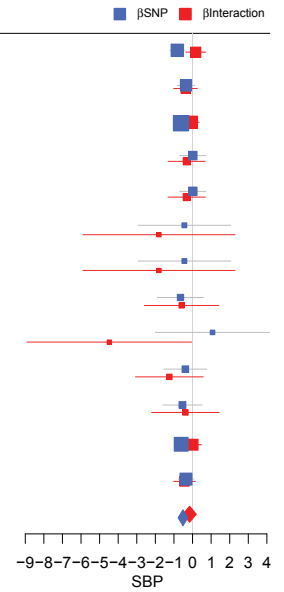
Effect of rs4841531 (T3−L7) and its interaction with EverSmk on SBP

Study	N	EAF	2df P
EUR S1	79732	0.213	6.069e−06
EUR S2	261459	0.2	0.0004417
EUR S1+2	341191	0.203	8.499e−08
AFR S1	25821	0.272	0.1787
AFR S2	1564	0.273	0.6178
AFR S1+2	27385	0.272	0.156
ASN S1	8705	0.107	0.3497
ASN S2	145159	0.032	0.4131
ASN S1+2	153864	0.036	0.4462
HIS S1	8742	0.213	0.9781
HIS S2	13529	0.185	0.04221
HIS S1+2	22271	0.196	0.0942
BRZ S2	4414	0.25	0.4752
Trans S1	123000	0.218	1.761e−06
Trans S2	426125	0.143	0.0007277
Trans S1+2	**549125**	**0.16**	**1.317e−08**



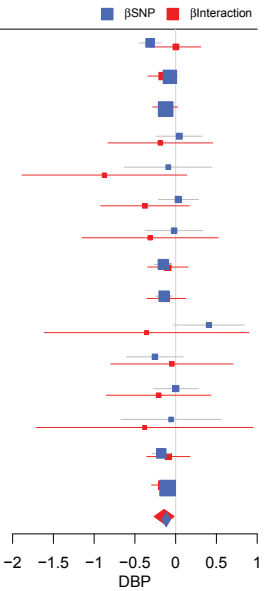
Effect of rs201133964 (T3−L5*) and its interaction with EverSmk on SBP

Study	N	EAF	2df P
EUR S1	76420	0.171	5.986e−06
EUR S2	50408	0.187	0.001544
EUR S1+2	126828	0.177	9.586e−09
AFR S1	24685	0.184	0.7089
AFR S1+2	24685	0.184	0.7089
ASN S1	5022	0.078	0.3783
ASN S1+2	5022	0.078	0.3783
HIS S1	8742	0.125	0.1903
HIS S2	1149	0.096	0.1079
HIS S1+2	9891	0.121	0.06797
BRZ S2	3541	0.178	0.2851
Trans S1	114869	0.166	8.257e−06
Trans S2	55098	0.184	0.0002744
Trans S1+2	**169967**	**0.172**	**1.239e−09**



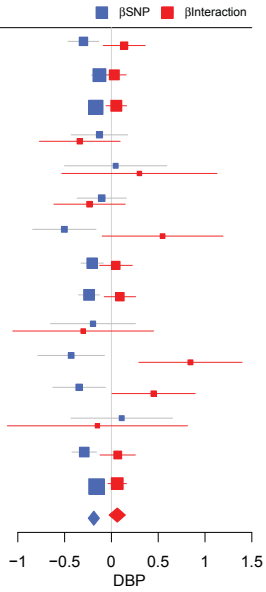
Effect of rs58429174 (T3−L8*) and its interaction with CurSmk on DBP

Study	N	EAF	2df P
EUR S1	79732	0.268	9.139e−06
EUR S2	288026	0.259	0.003226
EUR S1+2	367758	0.261	1.651e−06
AFR S1	25818	0.25	0.8553
AFR S2	7173	0.232	0.1657
AFR S1+2	32991	0.246	0.3765
ASN S1	11546	0.279	0.6998
ASN S2	148642	0.27	0.001922
ASN S1+2	160188	0.27	0.001757
HIS S1	8742	0.245	0.1938
HIS S2	13337	0.243	0.2514
HIS S1+2	22079	0.244	0.7876
BRZ S2	4414	0.263	0.7566
Trans S1	125838	0.264	0.001236
Trans S2	461592	0.262	2.029e−06
Trans S1+2	**587430**	**0.262**	**2.595e−09**



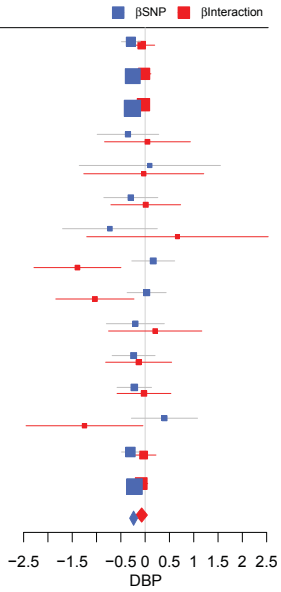
Effect of rs180940 (T3−L9) and its interaction with EverSmk on DBP

Study	N	EAF	2df P
EUR S1	79732	0.329	0.0005074
EUR S2	276437	0.318	0.001844
EUR S1+2	356169	0.321	2.905e−06
AFR S1	25818	0.438	0.01184
AFR S2	6180	0.465	0.634
AFR S1+2	31998	0.444	0.05146
ASN S1	11546	0.478	0.01602
ASN S2	148929	0.545	0.0002149
ASN S1+2	160475	0.54	1.464e−05
HIS S1	8742	0.372	0.194
HIS S2	13529	0.362	0.009175
HIS S1+2	22271	0.366	0.04453
BRZ S2	4414	0.397	0.9192
Trans S1	125838	0.368	8.601e−07
Trans S2	449489	0.398	1.082e−06
Trans S1+2	**575327**	**0.391**	**4.998e−12**



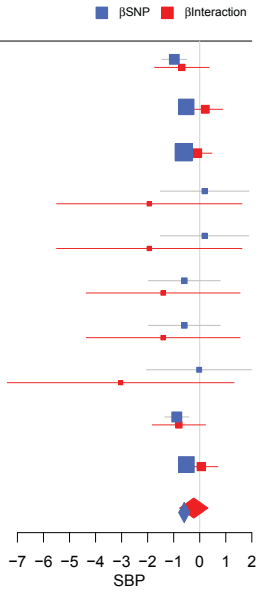
Effect of rs73073686 (T3−L12) and its interaction with EverSmk on DBP

Study	N	EAF	2df P
EUR S1	79732	0.242	2.79e−05
EUR S2	275207	0.249	1.245e−12
EUR S1+2	354939	0.248	1.94e−17
AFR S1	25818	0.072	0.3721
AFR S2	4246	0.056	0.9915
AFR S1+2	30064	0.07	0.414
ASN S1	5022	0.084	0.3566
ASN S2	11972	0.202	0.008852
ASN S1+2	16994	0.167	0.02754
HIS S1	8742	0.176	0.8071
HIS S2	13529	0.191	0.2345
HIS S1+2	22271	0.185	0.26
BRZ S2	4414	0.208	0.1253
Trans S1	119314	0.194	8.858e−06
Trans S2	309368	0.242	4.102e−13
Trans S1+2	**428682**	**0.228**	**1.677e−18**



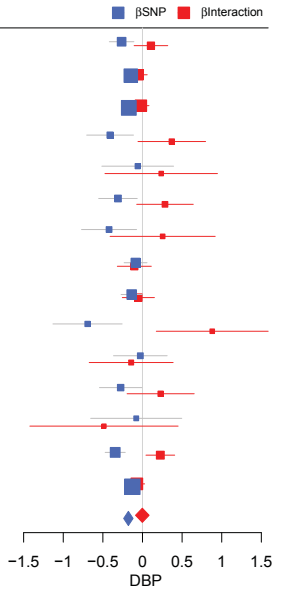
Effect of rs201316070 (T3−L10) and its interaction with CurSmk on SBP

Study	N	EAF	2df P
EUR S1	72250	0.066	2.12e−06
EUR S2	226241	0.065	0.0001282
EUR S1+2	298491	0.065	6.187e−09
AFR S1	15577	0.034	0.5357
AFR S1+2	15577	0.034	0.5357
HIS S2	12380	0.034	0.238
HIS S1+2	12380	0.034	0.238
BRZ S2	3541	0.031	0.2995
Trans S1	94331	0.059	2.121e−06
Trans S2	242162	0.063	5.282e−05
Trans S1+2	**329989**	**0.062**	**1.541e−09**



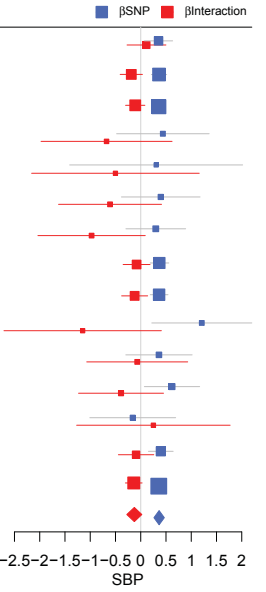
Effect of rs10858948 (T3−L13) and its interaction with EverSmk on DBP

Study	N	EAF	2df P
EUR S1	79732	0.499	0.0009504
EUR S2	276437	0.495	1.782e−08
EUR S1+2	356169	0.496	2.439e−11
AFR S1	25818	0.488	0.02933
AFR S2	6180	0.487	0.8061
AFR S1+2	31998	0.488	0.04467
ASN S1	11546	0.675	0.05486
ASN S2	146463	0.804	0.03224
ASN S1+2	158009	0.794	0.007433
HIS S1	8742	0.528	0.00711
HIS S2	13529	0.532	0.7186
HIS S1+2	22271	0.53	0.1264
BRZ S2	4414	0.51	0.3386
Trans S1	125838	0.515	2.073e−07
Trans S2	447023	0.597	4.774e−10
Trans S1+2	**572861**	**0.579**	**4.736e−15**



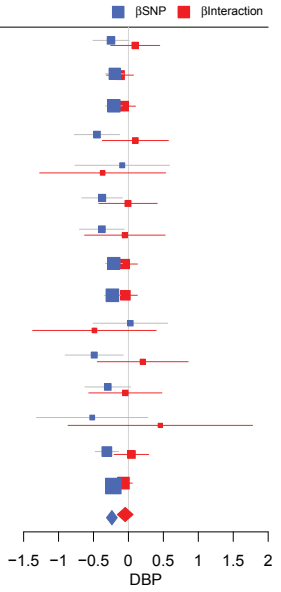
Effect of rs72656645 (T3−L11*) and its interaction with EverSmk on SBP

Study	N	EAF	2df P
EUR S1	77241	0.723	0.0004221
EUR S2	275206	0.724	1.671e−06
EUR S1+2	352447	0.724	1.734e−09
AFR S1	25821	0.914	0.5821
AFR S2	5568	0.925	0.8304
AFR S1+2	31389	0.916	0.474
ASN S1	11546	0.586	0.2062
ASN S2	148931	0.59	7.194e−06
ASN S1+2	160477	0.589	1.377e−05
HIS S1	8742	0.807	0.06004
HIS S2	13529	0.8	0.437
HIS S1+2	22271	0.802	0.07007
BRZ S2	4414	0.785	0.9259
Trans S1	123350	0.756	0.0005549
Trans S2	447648	0.685	1.43e−11
Trans S1+2	**570998**	**0.7**	**4.486e−15**



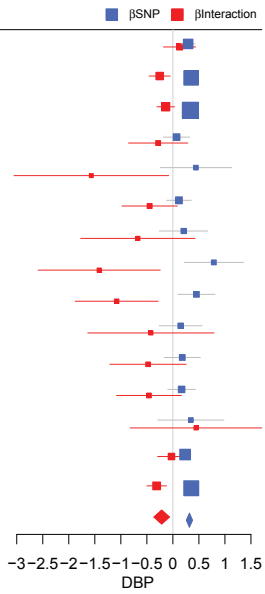
Effect of rs11067762 (T3−L14) and its interaction with EverSmk on DBP

Study	N	EAF	2df P
EUR S1	79732	0.107	0.09884
EUR S2	276154	0.101	8.922e−07
EUR S1+2	355886	0.102	1.486e−07
AFR S1	25818	0.259	0.004505
AFR S2	6180	0.263	0.5145
AFR S1+2	31998	0.26	0.002997
ASN S1	13438	0.405	0.01632
ASN S2	148929	0.313	4.425e−06
ASN S1+2	162367	0.321	1.297e−07
HIS S1	8742	0.197	0.4502
HIS S2	13529	0.203	0.04241
HIS S1+2	22271	0.201	0.06283
BRZ S2	4414	0.142	0.4445
Trans S1	127730	0.176	3.247e−05
Trans S2	449206	0.177	1.991e−13
Trans S1+2	**576936**	**0.177**	**5.302e−18**



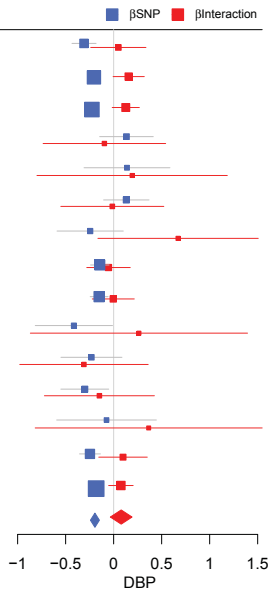
Effect of rs10628234 (T3−L15) and its interaction with CurSmk on DBP

Study	N	EAF	2df P
EUR S1	76420	0.296	4.244e−06
EUR S2	230048	0.303	3.385e−19
EUR S1+2	306468	0.301	4.15e−24
AFR S1	24682	0.395	0.6345
AFR S2	3122	0.383	0.1106
AFR S1+2	27804	0.394	0.2502
ASN S1	9654	0.168	0.4587
ASN S2	7815	0.186	0.003772
ASN S1+2	17469	0.176	0.006385
HIS S1	8742	0.276	0.7007
HIS S2	13337	0.262	0.389
HIS S1+2	22079	0.268	0.2729
BRZ S2	3541	0.334	0.2094
Trans S1	119498	0.305	9.789e−05
Trans S2	257863	0.298	4.511e−21
Trans S1+2	**377361**	**0.3**	**1.574e−24**



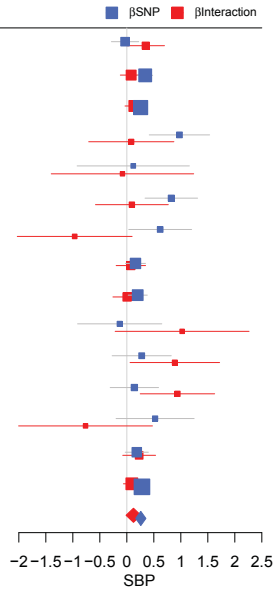
Effect of rs4968782 (T3−L18*) and its interaction with CurSmk on DBP

Study	N	EAF	2df P
EUR S1	79732	0.606	4.181e−06
EUR S2	290236	0.617	2.751e−09
EUR S1+2	369968	0.615	2.989e−14
AFR S1	24132	0.551	0.6566
AFR S2	7785	0.549	0.6796
AFR S1+2	31917	0.551	0.4993
ASN S1	11546	0.66	0.2181
ASN S2	148428	0.626	0.00267
ASN S1+2	159974	0.628	0.002149
HIS S1	8742	0.618	0.1305
HIS S2	13337	0.651	0.08263
HIS S1+2	22079	0.638	0.01629
BRZ S2	4414	0.574	0.8324
Trans S1	124152	0.601	2.636e−05
Trans S2	464200	0.619	1.665e−11
Trans S1+2	**588352**	**0.616**	**3.301e−16**



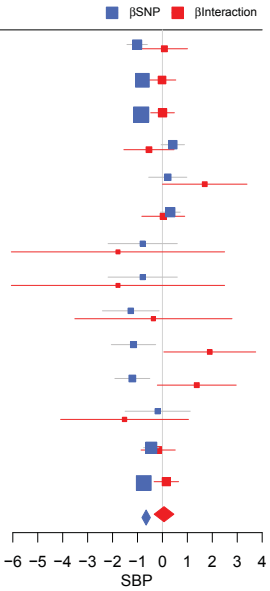
Effect of rs4888411 (T3−L16*) and its interaction with EverSmk on SBP

Study	N	EAF	2df P
EUR S1	79732	0.587	0.04776
EUR S2	273507	0.591	9.519e−13
EUR S1+2	353239	0.59	1.044e−12
AFR S1	22720	0.346	9.293e−06
AFR S2	4664	0.334	0.9749
AFR S1+2	27384	0.344	2.144e−05
ASN S1	11546	0.534	0.09058
ASN S2	148931	0.528	0.01962
ASN S1+2	160477	0.528	0.01086
HIS S1	8742	0.589	0.197
HIS S2	13529	0.604	0.001597
HIS S1+2	22271	0.598	0.000476
BRZ S2	4414	0.545	0.3276
Trans S1	122740	0.538	0.0003688
Trans S2	445045	0.567	4.329e−15
Trans S1+2	**567785**	**0.561**	**1.193e−18**



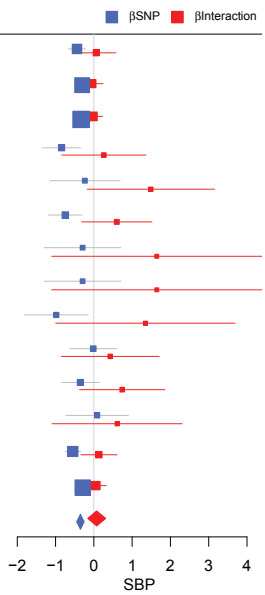
Effect of rs6541006 (T4−L1) and its interaction with CurSmk on SBP

Study	N	EAF	2df P
EUR S1	71824	0.092	2.446e−06
EUR S2	285304	0.066	2.374e−13
EUR S1+2	357128	0.071	3.165e−19
AFR S1	25821	0.286	0.2335
AFR S2	7224	0.311	0.08787
AFR S1+2	33045	0.292	0.224
ASN S2	11075	0.06	0.2892
ASN S1+2	11075	0.06	0.2892
HIS S1	8742	0.116	0.05476
HIS S2	13337	0.078	0.02323
HIS S1+2	22079	0.093	0.00315
BRZ S2	3541	0.095	0.3208
Trans S1	108878	0.14	0.002934
Trans S2	320481	0.072	5.318e−13
Trans S1+2	**426868**	**0.089**	**3.73e−15**



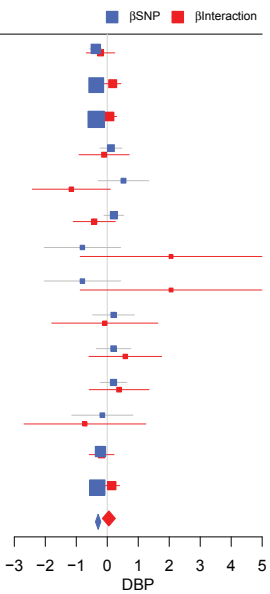
Effect of rs9899183 (T3−L17*) and its interaction with CurSmk on SBP

Study	N	EAF	2df P
EUR S1	79732	0.734	0.0003315
EUR S2	290234	0.733	8.557e−08
EUR S1+2	369966	0.733	2.663e−11
AFR S1	24135	0.745	0.003471
AFR S2	5708	0.746	0.2144
AFR S1+2	29843	0.745	0.003869
ASN S2	8607	0.884	0.4956
ASN S1+2	8607	0.884	0.4956
HIS S1	8742	0.808	0.06936
HIS S2	13337	0.823	0.7767
HIS S1+2	22079	0.817	0.2752
BRZ S2	4414	0.778	0.647
Trans S1	115100	0.746	3.161e−07
Trans S2	322300	0.741	6.199e−07
Trans S1+2	**434909**	**0.742**	**1.242e−12**



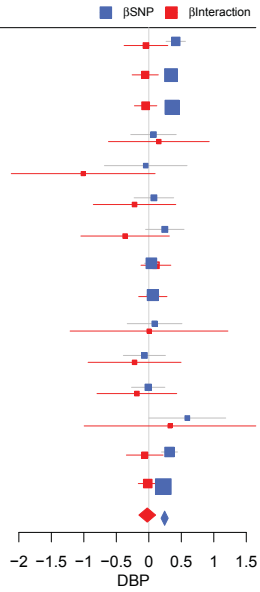
Effect of rs73923009 (T4−L2*) and its interaction with CurSmk on DBP

Study	N	EAF	2df P
EUR S1	78764	0.099	7.519e−05
EUR S2	279248	0.1	8.788e−11
EUR S1+2	358012	0.099	1.213e−14
AFR S1	25084	0.129	0.8117
AFR S2	4276	0.115	0.1306
AFR S1+2	29360	0.127	0.3181
ASN S2	5816	0.037	0.2889
ASN S1+2	5816	0.037	0.2889
HIS S1	8742	0.091	0.834
HIS S2	13337	0.08	0.2711
HIS S1+2	22079	0.084	0.2853
BRZ S2	3541	0.094	0.5706
Trans S1	116916	0.102	0.004921
Trans S2	306218	0.098	1.637e−09
Trans S1+2	**418808**	**0.1**	**1.389e−11**



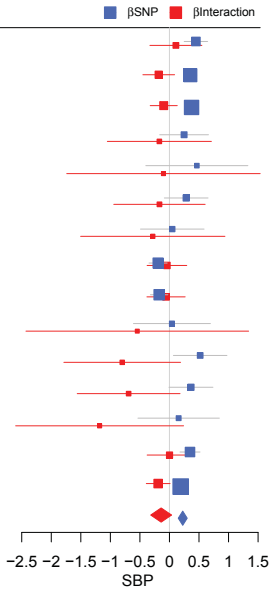
Effect of rs7753826 (T4−L3) and its interaction with CurSmk on DBP

Study	N	EAF	2df P
EUR S1	79732	0.195	8.819e−08
EUR S2	285767	0.187	4.614e−18
EUR S1+2	365499	0.189	1.715e−25
AFR S1	25084	0.144	0.7674
AFR S2	7173	0.137	0.1594
AFR S1+2	32257	0.142	0.7695
ASN S1	13438	0.475	0.2571
ASN S2	148929	0.623	0.316
ASN S1+2	162367	0.611	0.2509
HIS S1	8742	0.262	0.9085
HIS S2	13337	0.287	0.636
HIS S1+2	22079	0.277	0.797
BRZ S2	4414	0.258	0.04707
Trans S1	126996	0.219	5.18e−07
Trans S2	459620	0.331	2.063e−13
Trans S1+2	**586616**	**0.307**	**7.714e−20**



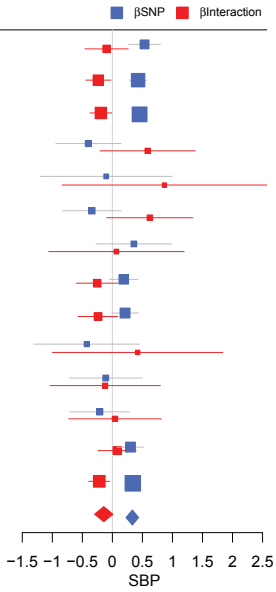
Effect of rs4841235 (T4−L6*) and its interaction with CurSmk on SBP

Study	N	EAF	2df P
EUR S1	79732	0.414	5.292e−06
EUR S2	276487	0.429	4.696e−10
EUR S1+2	356219	0.426	4.783e−15
AFR S1	25821	0.444	0.5143
AFR S2	4276	0.473	0.5687
AFR S1+2	30097	0.448	0.3151
ASN S1	11546	0.455	0.9043
ASN S2	148931	0.448	0.02127
ASN S1+2	160477	0.449	0.02739
HIS S1	8742	0.444	0.8518
HIS S2	13522	0.474	0.0641
HIS S1+2	22264	0.462	0.1116
BRZ S2	4414	0.414	0.2453
Trans S1	125841	0.426	6.378e−05
Trans S2	447630	0.437	7.561e−05
Trans S1+2	**573471**	**0.435**	**3.423e−08**



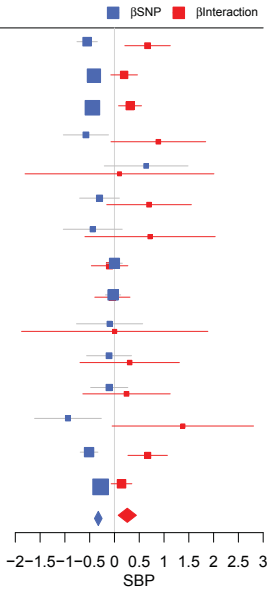
Effect of rs2243873 (T4−L4) and its interaction with EverSmk on SBP

Study	N	EAF	2df P
EUR S1	73293	0.568	5.632e−06
EUR S2	247196	0.552	3.981e−09
EUR S1+2	320489	0.556	3.331e−14
AFR S1	24135	0.622	0.3113
AFR S2	5073	0.626	0.5109
AFR S1+2	29208	0.623	0.2246
ASN S1	11546	0.683	0.3752
ASN S2	145564	0.816	0.258
ASN S1+2	157110	0.806	0.1724
HIS S1	8742	0.717	0.6421
HIS S2	13529	0.747	0.7756
HIS S1+2	22271	0.736	0.6191
Trans S1	117716	0.601	0.0002382
Trans S2	411362	0.653	6.083e−08
Trans S1+2	**529078**	**0.642**	**3.834e−11**



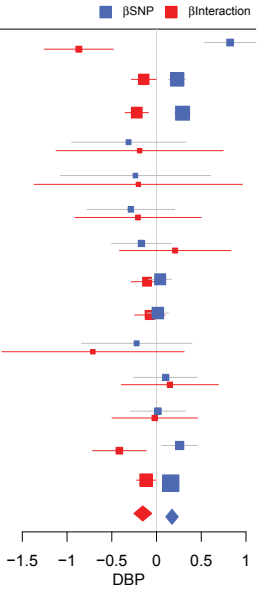
Effect of rs6995692 (T4−L7) and its interaction with CurSmk on SBP

Study	N	EAF	2df P
EUR S1	79732	0.568	2.141e−06
EUR S2	284854	0.562	1.027e−13
EUR S1+2	364586	0.563	4.11e−19
AFR S1	25821	0.725	0.04417
AFR S2	6678	0.731	0.2302
AFR S1+2	32499	0.726	0.1979
ASN S1	11546	0.488	0.327
ASN S2	147979	0.482	0.8498
ASN S1+2	159525	0.482	0.8568
HIS S1	8742	0.571	0.9547
HIS S2	13522	0.56	0.8129
HIS S1+2	22264	0.564	0.8089
BRZ S2	4414	0.553	0.019
Trans S1	125841	0.593	1.138e−07
Trans S2	457447	0.538	7.916e−10
Trans S1+2	**583288**	**0.55**	**9.927e−16**



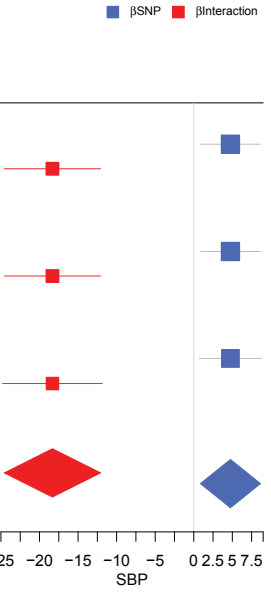
Effect of rs2071550 (T4−L5) and its interaction with EverSmk on DBP

Study	N	EAF	2df P
EUR S1	29453	0.315	4.519e−07
EUR S2	235742	0.306	3.64e−06
EUR S1+2	265195	0.307	1.167e−09
AFR S1	9849	0.184	0.2451
AFR S2	3207	0.168	0.7731
AFR S1+2	13056	0.18	0.1779
ASN S1	11546	0.4	0.619
ASN S2	135551	0.416	0.4921
ASN S1+2	147097	0.415	0.6222
HIS S1	5428	0.288	0.07137
HIS S2	13529	0.352	0.4526
HIS S1+2	18957	0.333	0.9948
Trans S1	56276	0.307	0.01652
Trans S2	388029	0.345	0.0001077
Trans S1+2	**444305**	**0.34**	**6.55e−06**



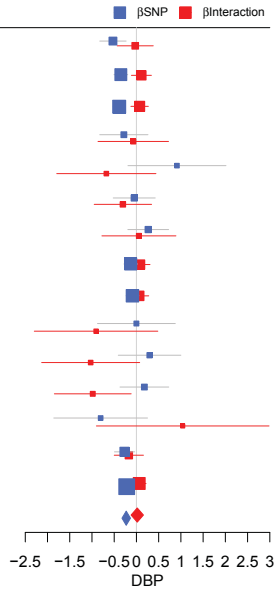
Effect of rs150155092 (T4−L8*) and its interaction with CurSmk on SBP

Study	N	EAF	2df P
EUR S1	9465	0.013	1.935e−08
EUR S1+2	9465	0.013	1.935e−08
Trans S1	9465	0.013	1.935e−08
Trans S1+2	**9465**	**0.013**	**1.935e−08**



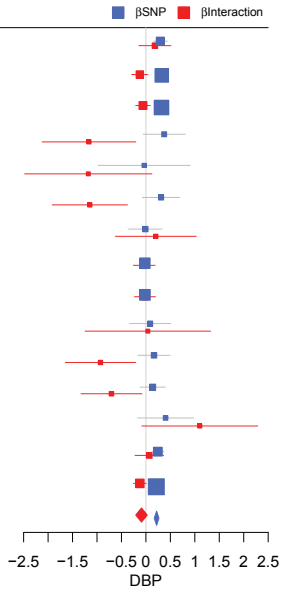
Effect of rs72941051 (T4−L9) and its interaction with EverSmk on DBP

Study	N	EAF	2df P
EUR S1	79732	0.077	5.738e−06
EUR S2	274924	0.073	9.177e−07
EUR S1+2	354656	0.074	1.753e−11
AFR S1	25818	0.076	0.3129
AFR S2	4246	0.077	0.2214
AFR S1+2	30064	0.076	0.4003
ASN S1	11546	0.159	0.3746
ASN S2	148929	0.178	0.2254
ASN S1+2	160475	0.176	0.4588
HIS S1	8742	0.067	0.2658
HIS S2	13348	0.062	0.1734
HIS S1+2	22090	0.064	0.05252
BRZ S2	4414	0.076	0.3142
Trans S1	125838	0.084	0.0001751
Trans S2	445861	0.108	9.151e−06
Trans S1+2	**571699**	**0.102**	**6.079e−09**



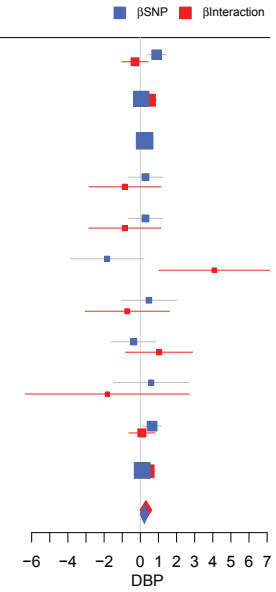
Effect of rs12050494 (T4−L12) and its interaction with CurSmk on DBP

Study	N	EAF	2df P
EUR S1	77241	0.301	4.632e−06
EUR S2	290955	0.32	4.109e−22
EUR S1+2	368196	0.316	3.005e−27
AFR S1	25818	0.095	0.05419
AFR S2	5851	0.076	0.1954
AFR S1+2	31669	0.092	0.01214
ASN S1	11546	0.487	0.886
ASN S2	148929	0.466	0.7837
ASN S1+2	160475	0.468	0.8394
HIS S1	8742	0.269	0.905
HIS S2	13337	0.28	0.04006
HIS S1+2	22079	0.276	0.0846
BRZ S2	4414	0.28	0.008919
Trans S1	123347	0.273	4.834e−05
Trans S2	463486	0.363	1.749e−14
Trans S1+2	**586833**	**0.344**	**1.169e−18**



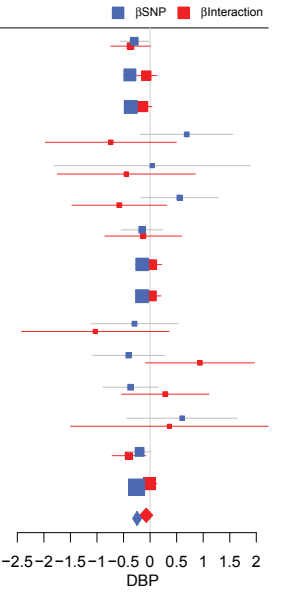
Effect of rs17713040 (T4−L10*) and its interaction with EverSmk on DBP

Study	N	EAF	2df P
EUR S1	74738	0.977	0.0007822
EUR S2	255880	0.977	2.047e−06
EUR S1+2	330618	0.977	3.443e−08
ASN S2	8607	0.954	0.6894
ASN S1+2	8607	0.954	0.6894
HIS S1	7287	0.986	0.03549
HIS S2	12380	0.985	0.8009
HIS S1+2	19667	0.985	0.5421
BRZ S2	3541	0.983	0.7157
Trans S1	84516	0.977	0.001299
Trans S2	280408	0.977	5.51e−06
Trans S1+2	**362433**	**0.977**	**5.037e−08**



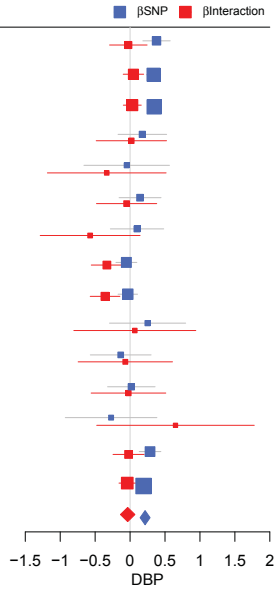
Effect of rs74439044 (T4−L13*) and its interaction with EverSmk on DBP

Study	N	EAF	2df P
EUR S1	79732	0.906	1.303e−06
EUR S2	273508	0.903	2.054e−15
EUR S1+2	353240	0.903	2.427e−21
AFR S1	24915	0.967	0.3173
AFR S2	3100	0.966	0.7914
AFR S1+2	28015	0.966	0.289
ASN S1	11546	0.737	0.5211
ASN S2	148929	0.684	0.01773
ASN S1+2	160475	0.688	0.01024
HIS S1	8742	0.928	0.05778
HIS S2	13348	0.929	0.2002
HIS S1+2	22090	0.928	0.3766
BRZ S2	4414	0.924	0.2157
Trans S1	124935	0.904	1.133e−06
Trans S2	443299	0.83	6.964e−13
Trans S1+2	**568234**	**0.847**	**7.32e−18**



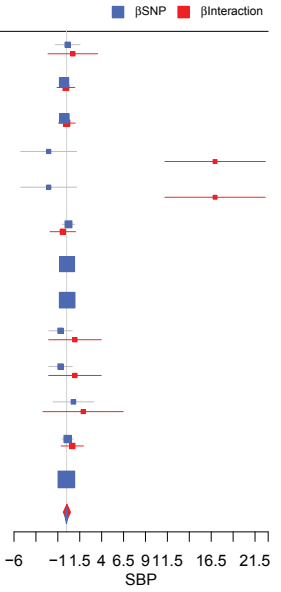
Effect of rs4375492 (T4−L11) and its interaction with EverSmk on DBP

Study	N	EAF	2df P
EUR S1	79732	0.796	2.981e−06
EUR S2	271526	0.793	1.02e−20
EUR S1+2	351258	0.794	1.034e−26
AFR S1	25818	0.748	0.3772
AFR S2	6180	0.754	0.617
AFR S1+2	31998	0.749	0.572
ASN S1	11546	0.693	0.284
ASN S2	148929	0.745	3.893e−05
ASN S1+2	160475	0.741	1.454e−05
HIS S1	8742	0.79	0.4486
HIS S2	13529	0.806	0.6356
HIS S1+2	22271	0.8	0.9944
BRZ S2	4414	0.778	0.5143
Trans S1	125838	0.776	1.184e−05
Trans S2	444578	0.777	6.143e−08
Trans S1+2	**570416**	**0.777**	**7.668e−13**



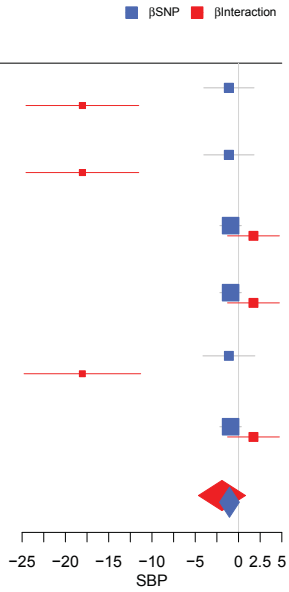
Effect of rs12135881 (T5−L1*) and its interaction with CurSmk on SBP

Study	N	EAF	2df P
EUR S1	25616	0.982	0.8148
EUR S2	254639	0.982	0.3142
EUR S1+2	280255	0.982	0.4221
AFR S1	7928	0.988	3.092e−09
AFR S1+2	7928	0.988	3.092e−09
ASN S1	12690	0.835	0.7976
ASN S2	147244	0.864	0.8753
ASN S1+2	159934	0.862	0.7853
HIS S2	12380	0.968	0.5884
HIS S1+2	12380	0.968	0.5884
BRZ S2	3541	0.971	0.3392
Trans S1	46234	0.942	0.3999
Trans S2	417804	0.94	0.8967
Trans S1+2	**464038**	**0.94**	**0.9818**



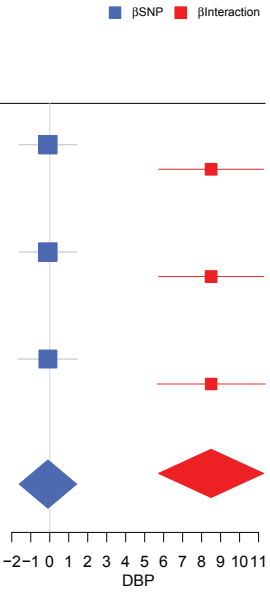
Effect of rs11809589 (T5−L2*) and its interaction with CurSmk on SBP

Study	N	EAF	2df P
AFR S1	7928	0.012	7.713e−10
AFR S1+2	7928	0.012	7.713e−10
ASN S2	7204	0.083	0.2805
ASN S1+2	7204	0.083	0.2805
Trans S1	7928	0.012	4.59e−09
Trans S2	7204	0.083	0.2805
Trans S1+2	**15132**	**0.046**	**0.01807**



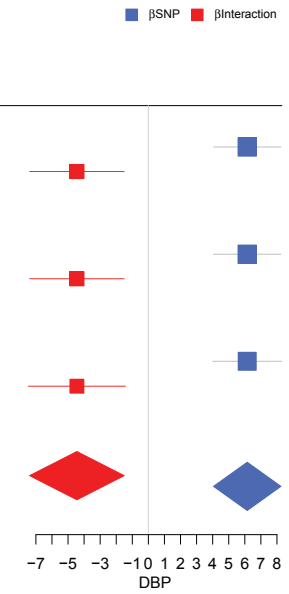
Effect of rs115234772 (T5−L5*) and its interaction with CurSmk on DBP

Study	N	EAF	2df P
AFR S1	10787	0.987	1.074e−11
AFR S1+2	10787	0.987	1.074e−11
Trans S1	10787	0.987	1.074e−11
Trans S1+2	**10787**	**0.987**	**1.074e−11**



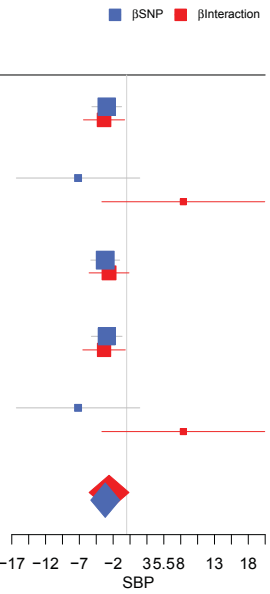
Effect of rs182662555 (T5−L3*) and its interaction with EverSmk on DBP

Study	N	EAF	2df P
AFR S1	13888	0.988	1.794e−08
AFR S1+2	13888	0.988	1.794e−08
Trans S1	13888	0.988	1.794e−08
Trans S1+2	**13888**	**0.988**	**1.794e−08**



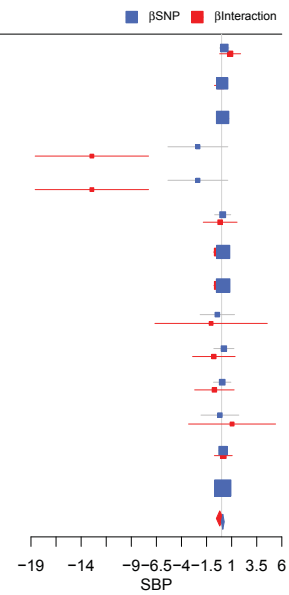
Effect of rs145162854 (T5−L6*) and its interaction with EverSmk on SBP

Study	N	EAF	2df P
AFR S1	18859	0.015	1.012e−08
AFR S2	1993	0.015	0.2904
AFR S1+2	20852	0.015	6.627e−09
Trans S1	18859	0.015	1.553e−08
Trans S2	1993	0.015	0.2904
Trans S1+2	**20852**	**0.015**	**6.734e−09**



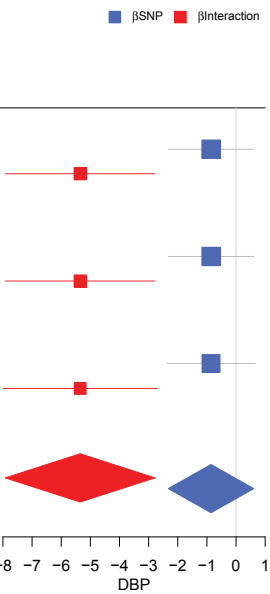
Effect of rs75247762 (T5−L4*) and its interaction with CurSmk on SBP

Study	N	EAF	2df P
EUR S1	70896	0.052	0.04753
EUR S2	269250	0.051	0.8956
EUR S1+2	340146	0.051	0.3546
AFR S1	7928	0.014	1.174e−09
AFR S1+2	7928	0.014	1.174e−09
ASN S1	10798	0.144	0.9661
ASN S2	144453	0.169	0.2201
ASN S1+2	155251	0.167	0.2173
HIS S1	7287	0.043	0.7701
HIS S2	13337	0.054	0.767
HIS S1+2	20624	0.05	0.7613
BRZ S2	3541	0.04	0.8947
Trans S1	96909	0.059	0.6082
Trans S2	430581	0.09	0.2344
Trans S1+2	**527490**	**0.085**	**0.2334**

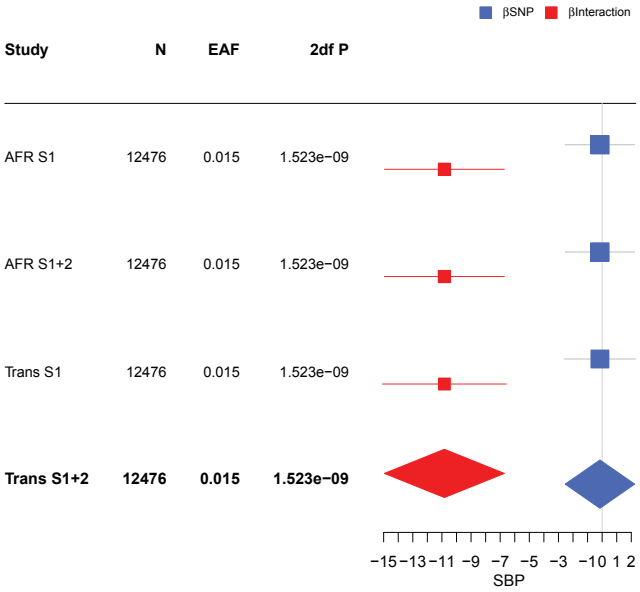


Effect of rs116008367 (T5−L7*) and its interaction with CurSmk on DBP

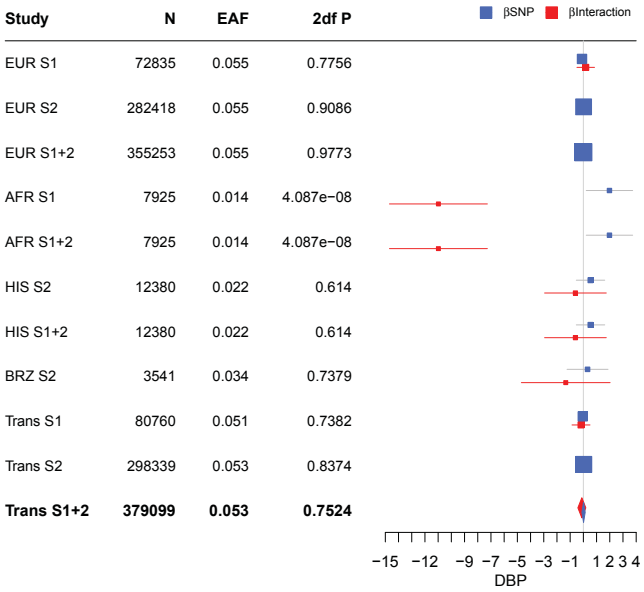
Study	N	EAF	2df P
AFR S1	10787	0.014	3.09e−08
AFR S1+2	10787	0.014	3.09e−08
Trans S1	10787	0.014	3.09e−08
Trans S1+2	**10787**	**0.014**	**3.09e−08**



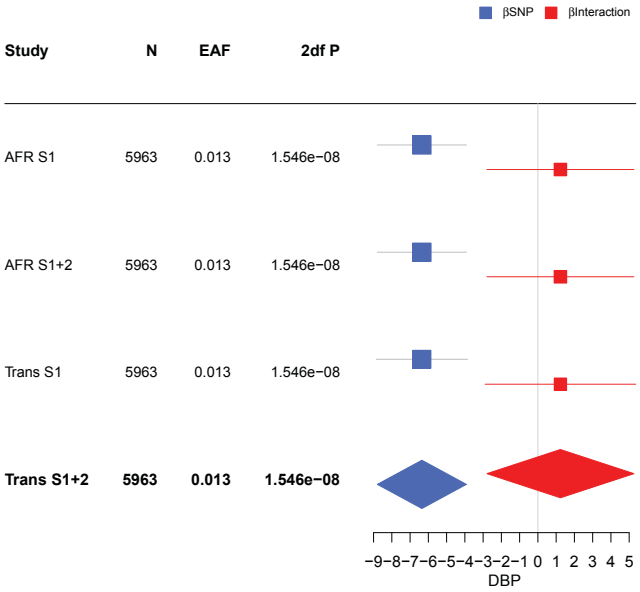
Effect of rs10166552 (T5−L8*) and its interaction with CurSmk on SBP



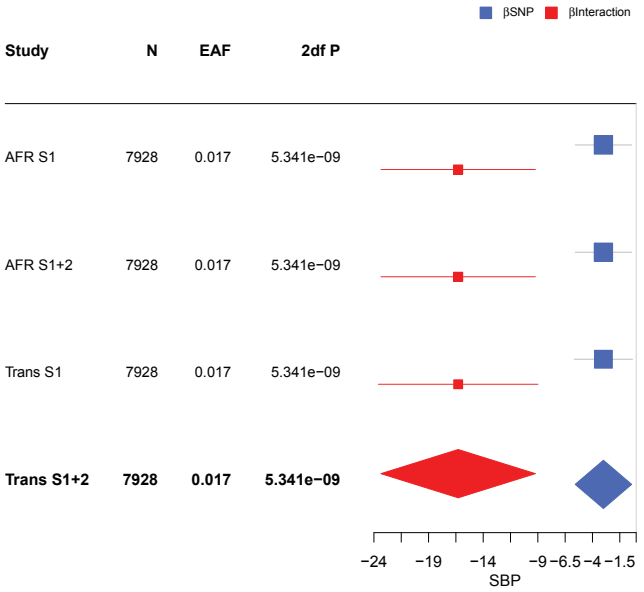
Effect of rs62319742 (T5−L11*) and its interaction with CurSmk on DBP



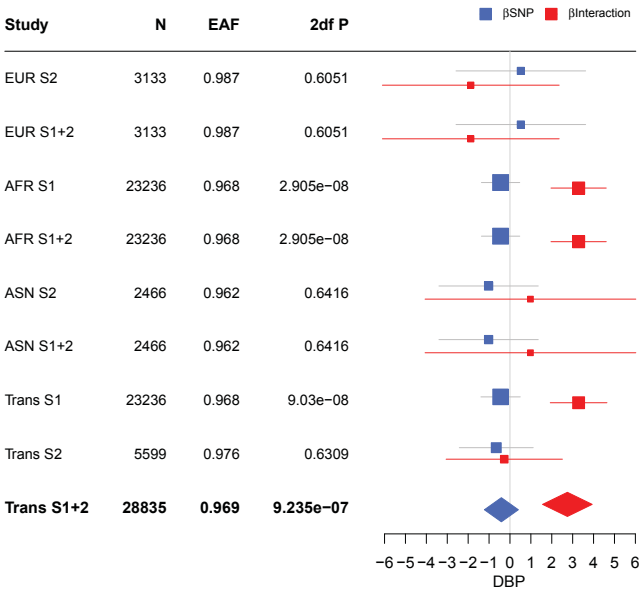
Effect of rs139963642 (T5−L9*) and its interaction with EverSmk on DBP



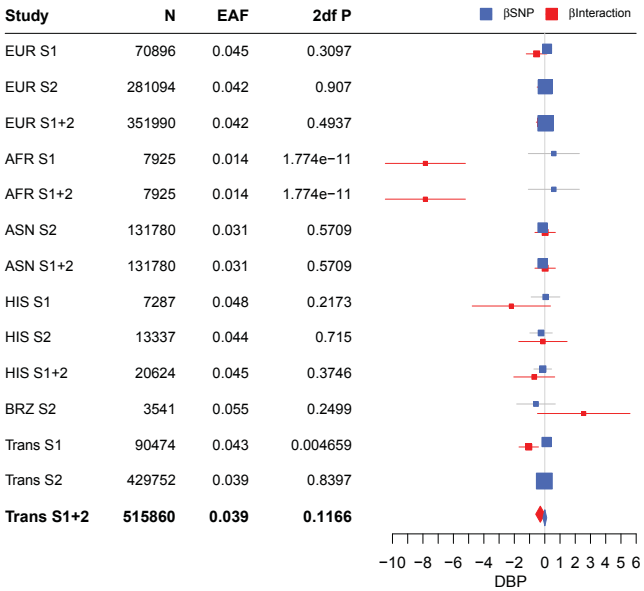
Effect of rs140543491 (T5−L12*) and its interaction with CurSmk on SBP



Effect of rs11931572 (T5−L10*) and its interaction with EverSmk on DBP

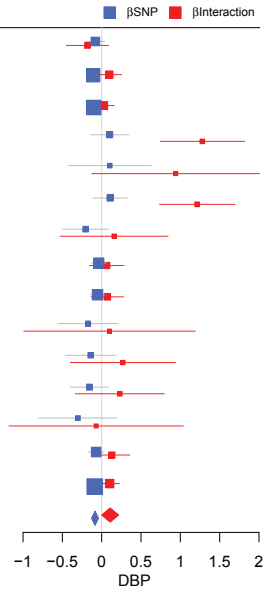


Effect of rs148387718 (T5−L13*) and its interaction with CurSmk on DBP



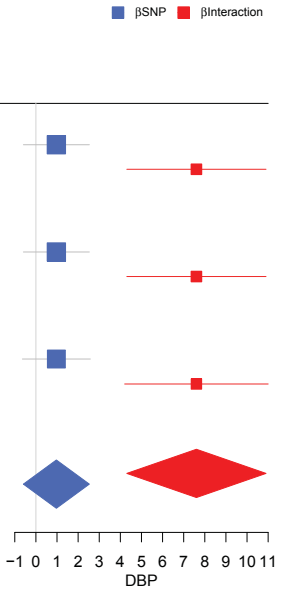
Effect of rs9348895 (T5−L14*) and its interaction with CurSmk on DBP

Study	N	EAF	2df P
EUR S1	79732	0.428	0.061
EUR S2	281869	0.416	0.003682
EUR S1+2	361601	0.419	0.0008746
AFR S1	25818	0.588	2.485e−07
AFR S2	4276	0.578	0.1502
AFR S1+2	30094	0.586	1.706e−08
ASN S1	13438	0.441	0.3907
ASN S2	148929	0.522	0.7059
ASN S1+2	162367	0.515	0.4906
HIS S1	8742	0.398	0.6643
HIS S2	13337	0.363	0.6192
HIS S1+2	22079	0.377	0.4352
BRZ S2	4414	0.442	0.3885
Trans S1	127730	0.46	0.3343
Trans S2	452825	0.451	0.003032
Trans S1+2	**580555**	**0.453**	**0.001081**



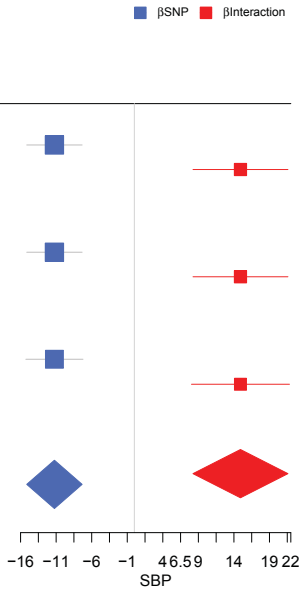
Effect of rs112140754 (T5−L17*) and its interaction with CurSmk on DBP

Study	N	EAF	2df P
AFR S1	10787	0.988	2.436e−08
AFR S1+2	10787	0.988	2.436e−08
Trans S1	10787	0.988	2.436e−08
Trans S1+2	**10787**	**0.988**	**2.436e−08**



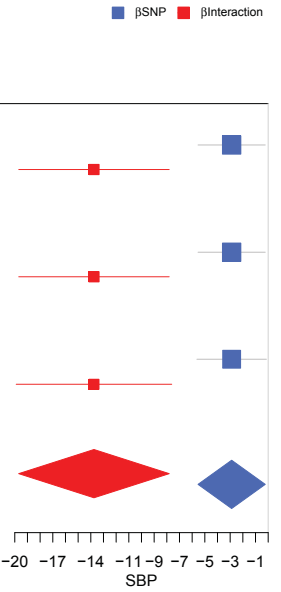
Effect of rs58806982 (T5−L15*) and its interaction with EverSmk on SBP

Study	N	EAF	2df P
AFR S1	7928	0.01	4.569e−08
AFR S1+2	7928	0.01	4.569e−08
Trans S1	7928	0.01	4.569e−08
Trans S1+2	**7928**	**0.01**	**4.569e−08**



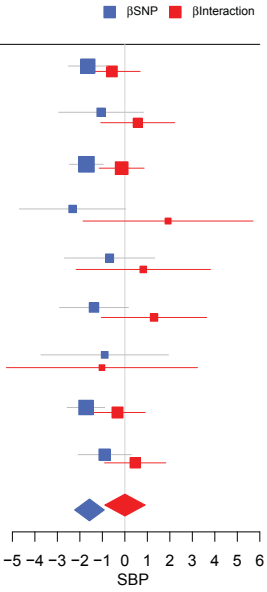
Effect of rs116196735 (T5−L18*) and its interaction with CurSmk on SBP

Study	N	EAF	2df P
AFR S1	10790	0.018	6.982e−10
AFR S1+2	10790	0.018	6.982e−10
Trans S1	10790	0.018	6.982e−10
Trans S1+2	**10790**	**0.018**	**6.982e−10**



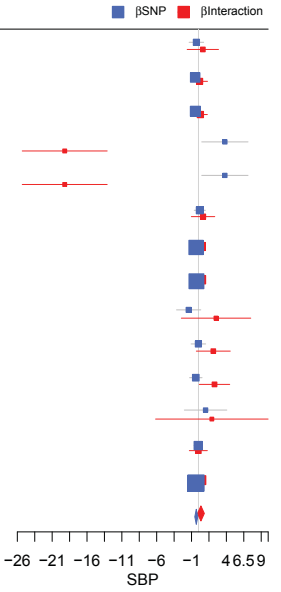
Effect of rs76987554 (T5−L16*) and its interaction with EverSmk on SBP

Study	N	EAF	2df P
AFR S1	25821	0.088	4.239e−08
AFR S2	4664	0.089	0.5344
AFR S1+2	30485	0.088	2.957e−08
HIS S1	8742	0.027	0.1566
HIS S2	12380	0.021	0.7947
HIS S1+2	21122	0.024	0.2136
BRZ S2	3541	0.02	0.4444
Trans S1	34563	0.073	2.168e−08
Trans S2	20585	0.036	0.3258
Trans S1+2	**55148**	**0.059**	**1.632e−08**



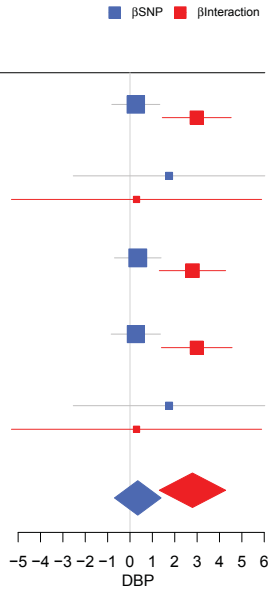
Effect of rs74701635 (T5−L19*) and its interaction with CurSmk on SBP

Study	N	EAF	2df P
EUR S1	52072	0.013	0.8216
EUR S2	238753	0.015	0.1271
EUR S1+2	290825	0.014	0.1244
AFR S1	7928	0.016	2.128e−09
AFR S1+2	7928	0.016	2.128e−09
ASN S1	10798	0.15	0.4908
ASN S2	144231	0.14	0.03645
ASN S1+2	155029	0.14	0.06466
HIS S1	7287	0.04	0.2557
HIS S2	13337	0.046	0.1783
HIS S1+2	20624	0.044	0.1186
BRZ S2	3541	0.016	0.602
Trans S1	78085	0.035	0.9853
Trans S2	399862	0.061	0.008559
Trans S1+2	**477947**	**0.057**	**0.01408**



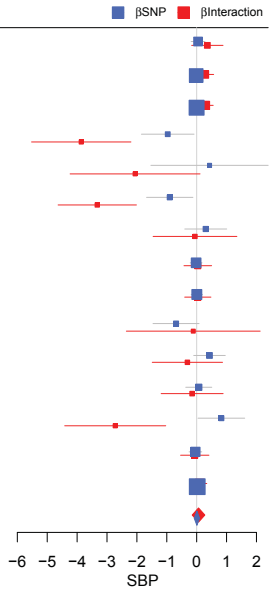
Effect of rs146250839 (T5−L20*) and its interaction with EverSmk on DBP

Study	N	EAF	2df P
AFR S1	22619	0.976	1.654e−07
AFR S2	1993	0.981	0.3962
AFR S1+2	24612	0.977	4.361e−08
Trans S1	22619	0.976	2.145e−07
Trans S2	1993	0.981	0.3962
Trans S1+2	**24612**	**0.977**	**4.377e−08**



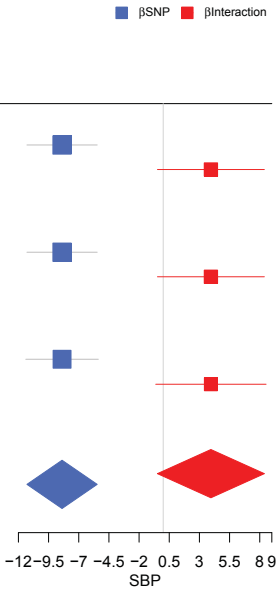
Effect of rs11599481 (T5−L23*) and its interaction with CurSmk on SBP

Study	N	EAF	2df P
EUR S1	77241	0.245	0.2594
EUR S2	289723	0.249	0.1694
EUR S1+2	366964	0.248	0.05075
AFR S1	24146	0.06	1.662e−10
AFR S2	3661	0.049	0.1709
AFR S1+2	27807	0.058	4.545e−11
ASN S1	12690	0.195	0.6548
ASN S2	148143	0.16	0.9794
ASN S1+2	160833	0.163	0.9767
HIS S1	8742	0.216	0.1672
HIS S2	13337	0.238	0.2779
HIS S1+2	22079	0.229	0.937
BRZ S2	4414	0.218	0.005224
Trans S1	122819	0.201	0.8302
Trans S2	459278	0.218	0.5403
Trans S1+2	**582097**	**0.215**	**0.7726**



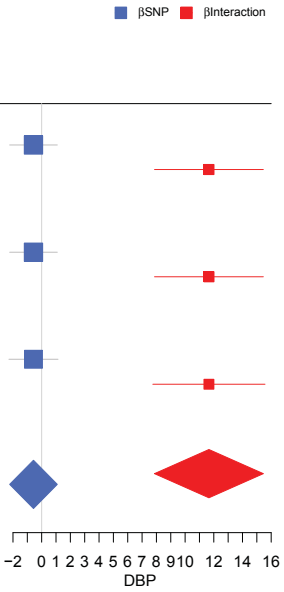
Effect of rs192642798 (T5−L21*) and its interaction with EverSmk on SBP

Study	N	EAF	2df P
AFR S1	13891	0.012	4.227e−09
AFR S1+2	13891	0.012	4.227e−09
Trans S1	13891	0.012	4.227e−09
Trans S1+2	**13891**	**0.012**	**4.227e−09**



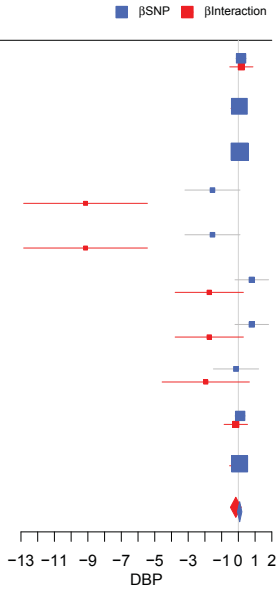
Effect of rs148772934 (T5−L24*) and its interaction with CurSmk on DBP

Study	N	EAF	2df P
AFR S1	7925	0.986	1.197e−09
AFR S1+2	7925	0.986	1.197e−09
Trans S1	7925	0.986	1.197e−09
Trans S1+2	**7925**	**0.986**	**1.197e−09**



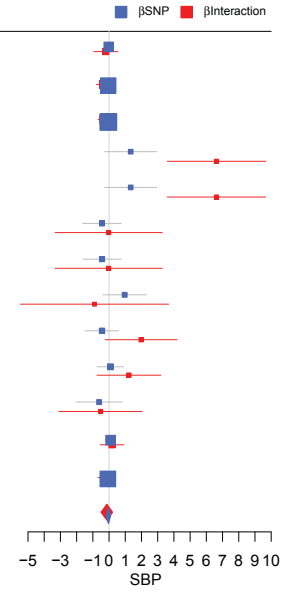
Effect of rs76726877 (T5−L22*) and its interaction with CurSmk on DBP

Study	N	EAF	2df P
EUR S1	67934	0.055	0.37
EUR S2	279915	0.049	0.7315
EUR S1+2	347849	0.05	0.4585
AFR S1	7925	0.016	4.465e−10
AFR S1+2	7925	0.016	4.465e−10
HIS S2	12380	0.027	0.1699
HIS S1+2	12380	0.027	0.1699
BRZ S2	3541	0.047	0.1816
Trans S1	79832	0.05	0.7895
Trans S2	295836	0.048	0.5449
Trans S1+2	**371695**	**0.049**	**0.4461**



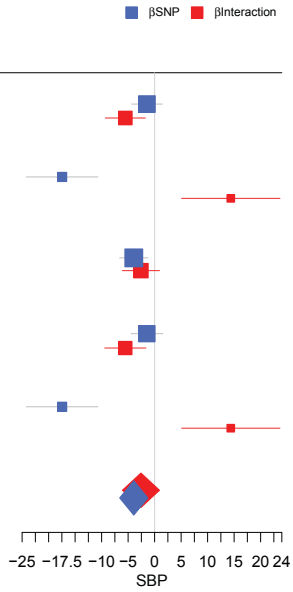
Effect of rs11601370 (T5−L25) and its interaction with CurSmk on SBP

Study	N	EAF	2df P
EUR S1	75802	0.892	0.8303
EUR S2	286617	0.887	0.131
EUR S1+2	362419	0.888	0.1192
AFR S1	16817	0.976	3.005e−09
AFR S1+2	16817	0.976	3.005e−09
ASN S2	11075	0.923	0.7517
ASN S1+2	11075	0.923	0.7517
HIS S1	8742	0.947	0.378
HIS S2	13337	0.949	0.2174
HIS S1+2	22079	0.948	0.3607
BRZ S2	3541	0.92	0.4102
Trans S1	103852	0.911	0.6226
Trans S2	314570	0.891	0.1411
Trans S1+2	**415931**	**0.896**	**0.4958**



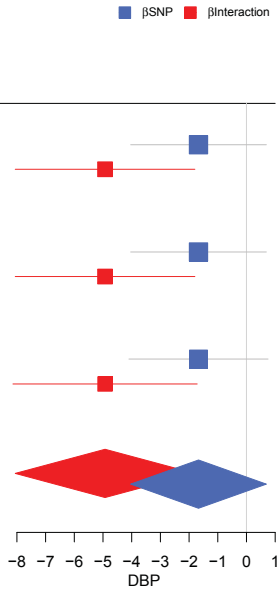
Effect of rs74601585 (T5−L26*) and its interaction with EverSmk on SBP

Study	N	EAF	2df P
AFR S1	15131	0.016	6.866e−07
AFR S2	1993	0.015	1.744e−06
AFR S1+2	17124	0.016	8.285e−09
Trans S1	15131	0.016	9.597e−07
Trans S2	1993	0.015	1.744e−06
Trans S1+2	**17124**	**0.016**	**8.06e−09**



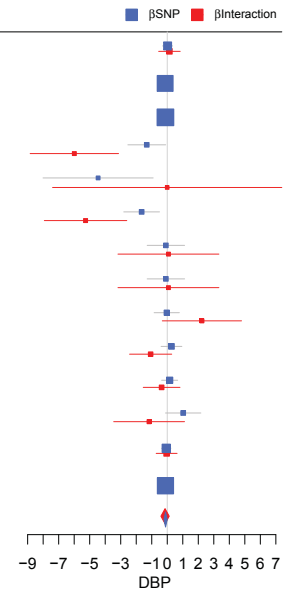
Effect of rs187852559 (T5−L29*) and its interaction with EverSmk on DBP

Study	N	EAF	2df P
AFR S1	5963	0.013	8.738e−10
AFR S1+2	5963	0.013	8.738e−10
Trans S1	5963	0.013	8.738e−10
Trans S1+2	**5963**	**0.013**	**8.738e−10**



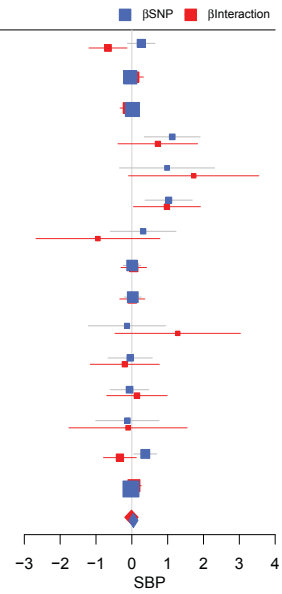
Effect of rs78103586 (T5−L27*) and its interaction with CurSmk on DBP

Study	N	EAF	2df P
EUR S1	45362	0.079	0.8802
EUR S2	271746	0.086	0.01705
EUR S1+2	317108	0.085	0.04387
AFR S1	10787	0.031	6.034e−08
AFR S2	1993	0.023	0.02001
AFR S1+2	12780	0.03	2.262e−09
ASN S2	5816	0.051	0.9897
ASN S1+2	5816	0.051	0.9897
HIS S1	8742	0.056	0.2101
HIS S2	13337	0.054	0.302
HIS S1+2	22079	0.055	0.7796
BRZ S2	3541	0.074	0.2088
Trans S1	64891	0.068	0.8665
Trans S2	296433	0.083	0.01758
Trans S1+2	**361324**	**0.08**	**0.01817**



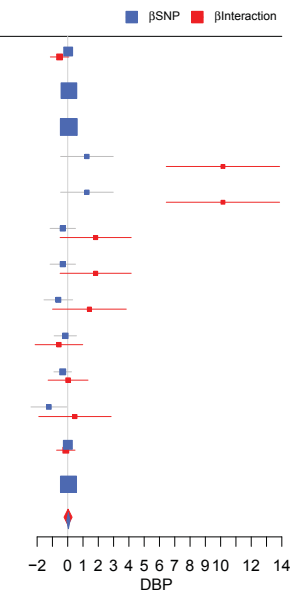
Effect of rs1257310 (T5−L30*) and its interaction with EverSmk on SBP

Study	N	EAF	2df P
EUR S1	77241	0.796	0.05319
EUR S2	269471	0.813	0.8955
EUR S1+2	346712	0.809	0.6871
AFR S1	21034	0.792	1.487e−06
AFR S2	4102	0.728	0.01716
AFR S1+2	25136	0.781	1.669e−08
ASN S1	6524	0.706	0.5642
ASN S2	147478	0.707	0.9027
ASN S1+2	154002	0.707	0.9341
HIS S1	8742	0.786	0.2643
HIS S2	13529	0.771	0.8136
HIS S1+2	22271	0.777	0.9465
BRZ S2	4414	0.818	0.9174
Trans S1	113541	0.79	0.07448
Trans S2	438994	0.775	0.8544
Trans S1+2	**552535**	**0.778**	**0.7131**



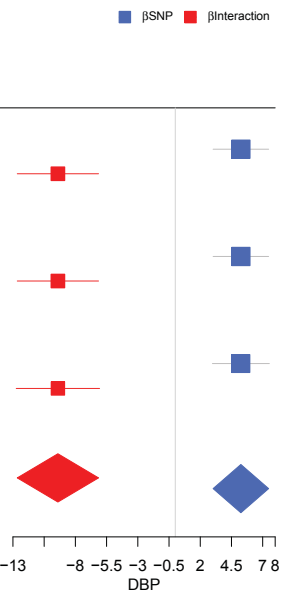
Effect of rs61935525 (T5−L28*) and its interaction with CurSmk on DBP

Study	N	EAF	2df P
EUR S1	72888	0.933	0.2212
EUR S2	283524	0.925	0.2612
EUR S1+2	356412	0.926	0.4419
AFR S1	7925	0.986	3.277e−11
AFR S1+2	7925	0.986	3.277e−11
ASN S2	11073	0.947	0.291
ASN S1+2	11073	0.947	0.291
HIS S1	7287	0.946	0.3457
HIS S2	12380	0.95	0.5394
HIS S1+2	19667	0.949	0.4743
BRZ S2	3541	0.932	0.09026
Trans S1	88100	0.939	0.9166
Trans S2	310518	0.927	0.4797
Trans S1+2	**398618**	**0.929**	**0.613**



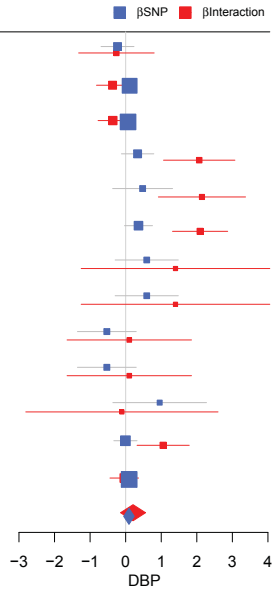
Effect of rs148753653 (T5−L31) and its interaction with EverSmk on DBP

Study	N	EAF	2df P
AFR S1	7925	0.989	6.298e−08
AFR S1+2	7925	0.989	6.298e−08
Trans S1	7925	0.989	6.298e−08
Trans S1+2	**7925**	**0.989**	**6.298e−08**



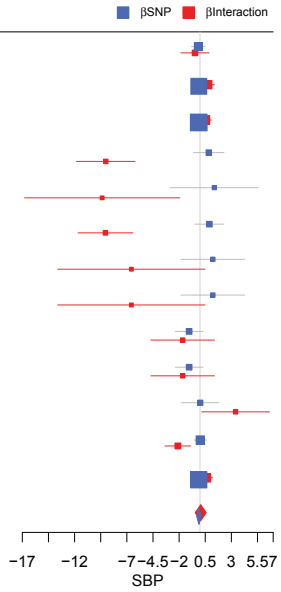
Effect of rs138973557 (T5−L32*) and its interaction with CurSmk on DBP

Study	N	EAF	2df P
EUR S1	62368	0.974	0.4085
EUR S2	274814	0.967	0.2271
EUR S1+2	337182	0.968	0.2335
AFR S1	23398	0.903	8.68e−07
AFR S2	5708	0.902	0.0004705
AFR S1+2	29106	0.903	1.805e−10
ASN S2	9014	0.958	0.1445
ASN S1+2	9014	0.958	0.1445
HIS S2	12380	0.961	0.3958
HIS S1+2	12380	0.961	0.3958
BRZ S2	3541	0.945	0.2846
Trans S1	89739	0.954	0.008253
Trans S2	305457	0.965	0.4938
Trans S1+2	**391223**	**0.963**	**0.09528**



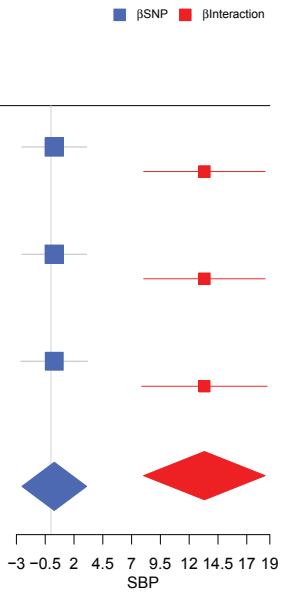
Effect of rs115893283 (T5−L35*) and its interaction with CurSmk on SBP

Study	N	EAF	2df P
EUR S1	41976	0.054	0.5377
EUR S2	269126	0.056	0.06059
EUR S1+2	311102	0.056	0.1841
AFR S1	15131	0.042	5.841e−10
AFR S2	1606	0.053	0.03063
AFR S1+2	16737	0.043	4.241e−12
ASN S2	2468	0.056	0.1286
ASN S1+2	2468	0.056	0.1286
HIS S2	12380	0.032	0.04962
HIS S1+2	12380	0.032	0.04962
BRZ S2	3541	0.049	0.04681
Trans S1	61080	0.05	0.0009311
Trans S2	289121	0.055	0.1149
Trans S1+2	**346228**	**0.054**	**0.629**



Effect of rs9965695 (T5−L33*) and its interaction with CurSmk on SBP

Study	N	EAF	2df P
AFR S1	10790	0.983	1.627e−08
AFR S1+2	10790	0.983	1.627e−08
Trans S1	10790	0.983	1.627e−08
Trans S1+2	**10790**	**0.983**	**1.627e−08**



Effect of rs10405764 (T5−L34*) and its interaction with CurSmk on SBP

Study	N	EAF	2df P
AFR S1	7928	0.017	4.302e−08
AFR S1+2	7928	0.017	4.302e−08
Trans S1	7928	0.017	4.302e−08
Trans S1+2	**7928**	**0.017**	**4.302e−08**

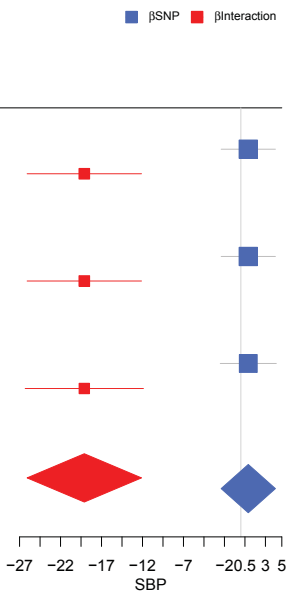
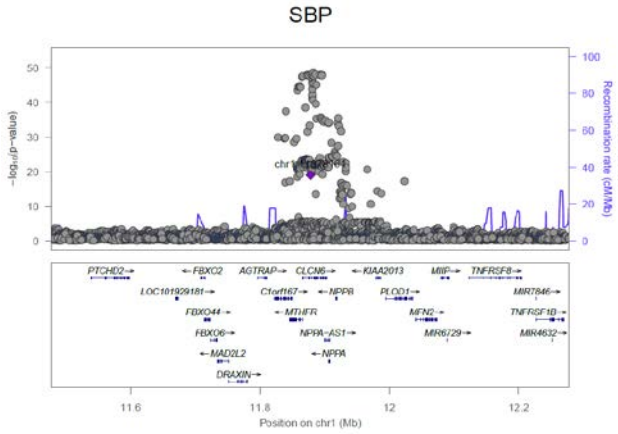


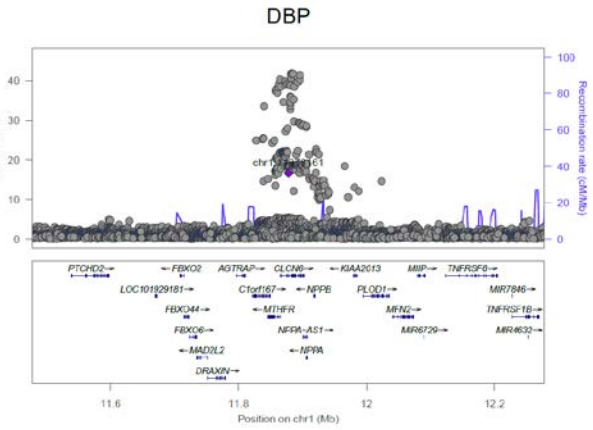
Figure S4: LocusZoom plots for the 81 newly identified loci (Tables 2-5).

LocusZoom plots are ordered by tables (2-5) then by loci within each table. Each locus has at least one BP trait reaching genome-wide significance. If both traits reach genome-wide significance at a locus, then two plots are shown. T2-L1.A and T2-L1.B refer to association with SBP and DBP, respectively, at locus 1 in Table 2.

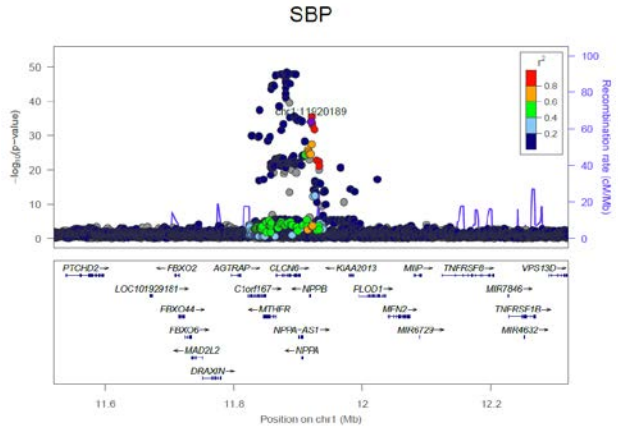
T2 - L1.A



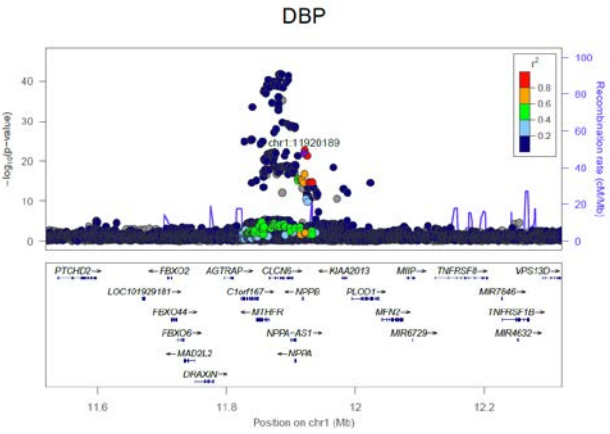
T2 – L1.B



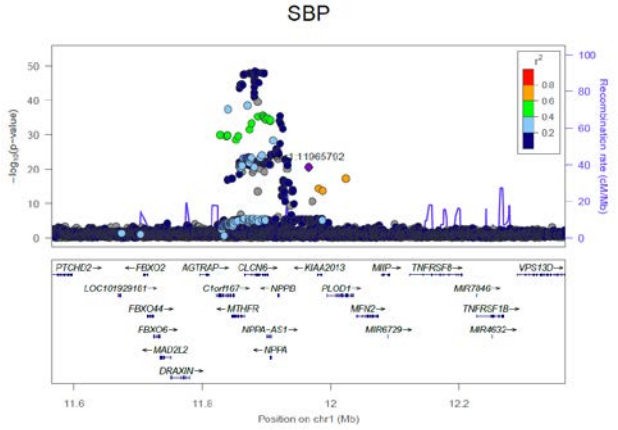
T2 - L2.A



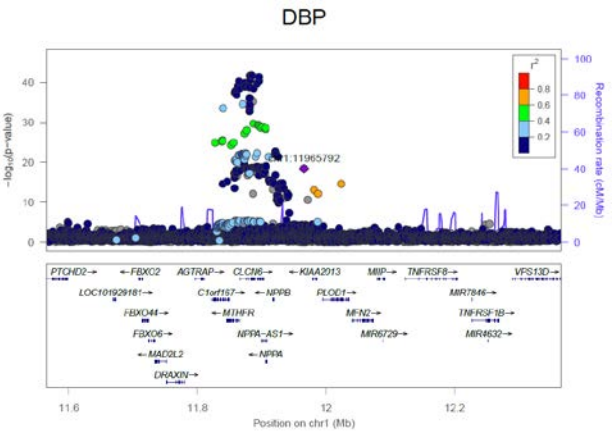
T2 – L2.B



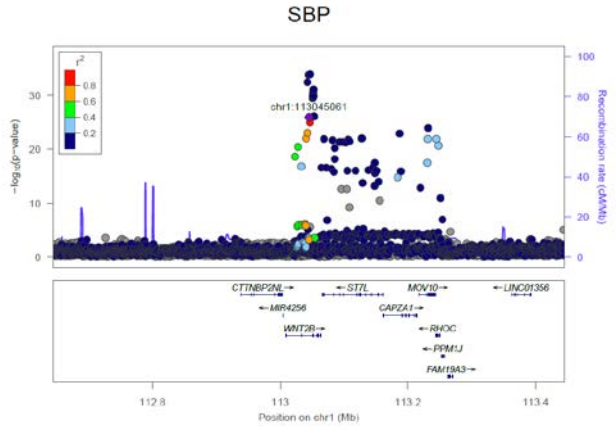
T2 – L3.A



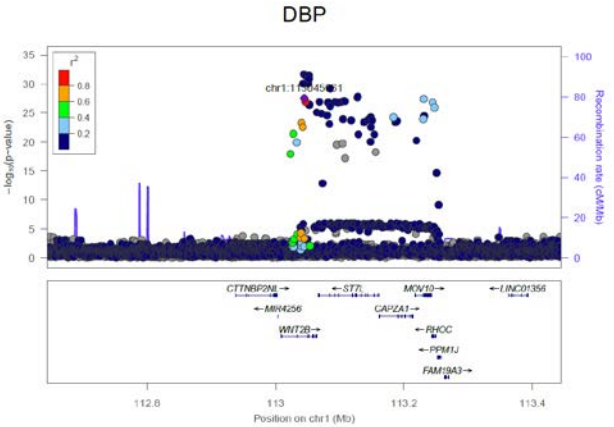
T2 – L3.B



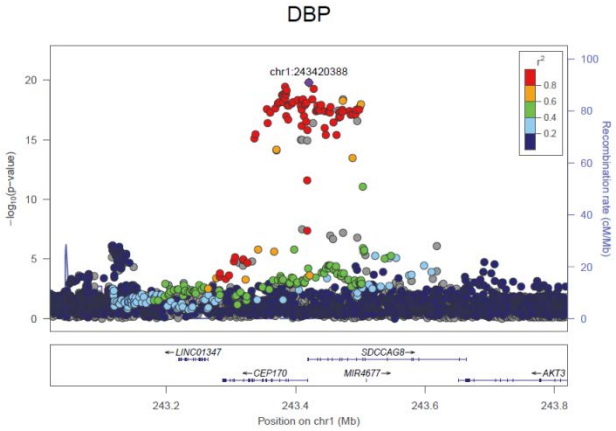
T2 - L4.A



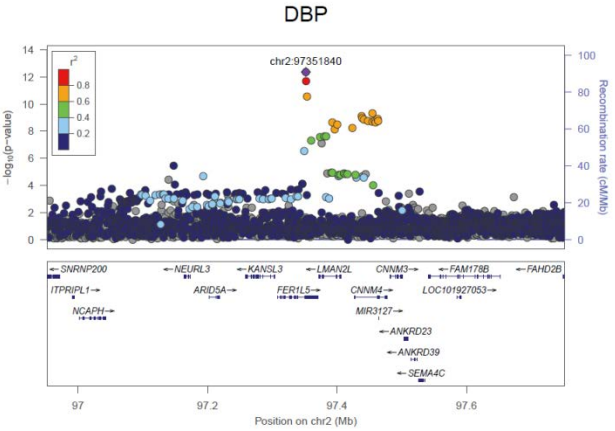
T2 - L4.B



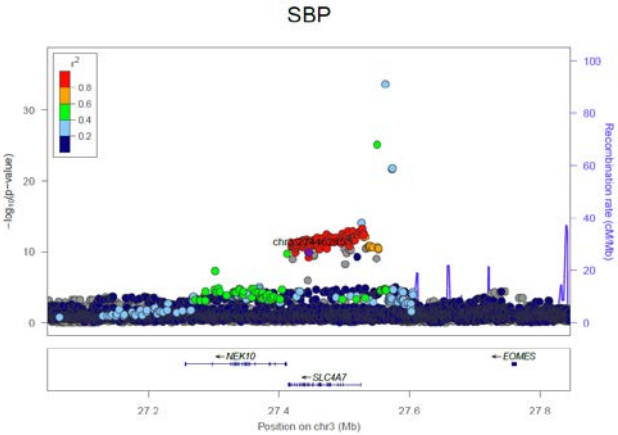
T2 - L5



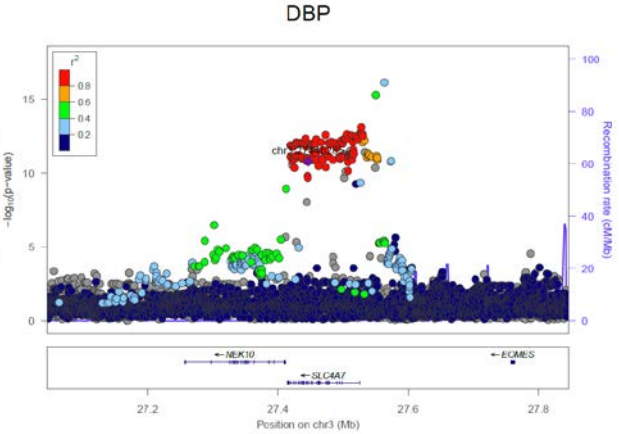
T2 - L6



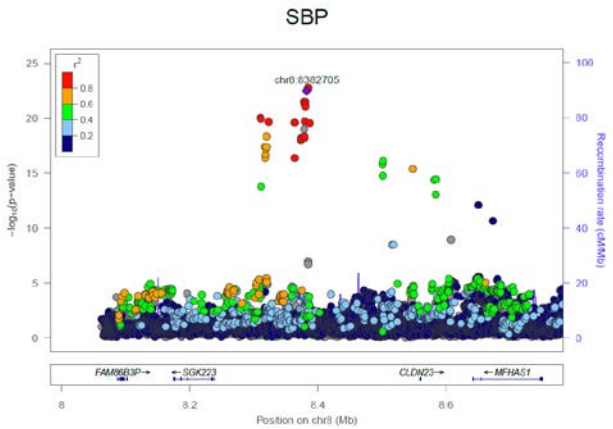
T2 – L7.A



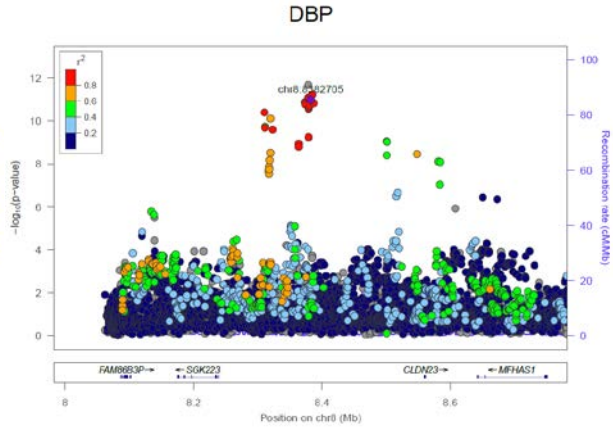
T2 – L7.B



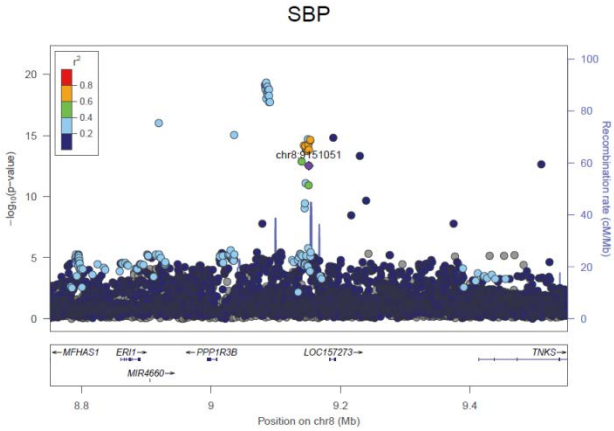
T2 – L8.A



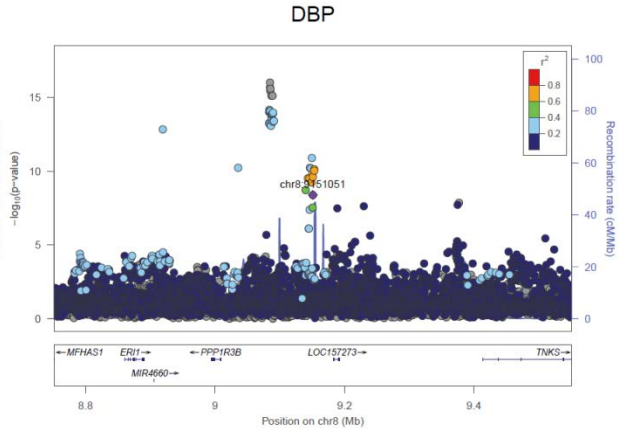
T2 – L8.B



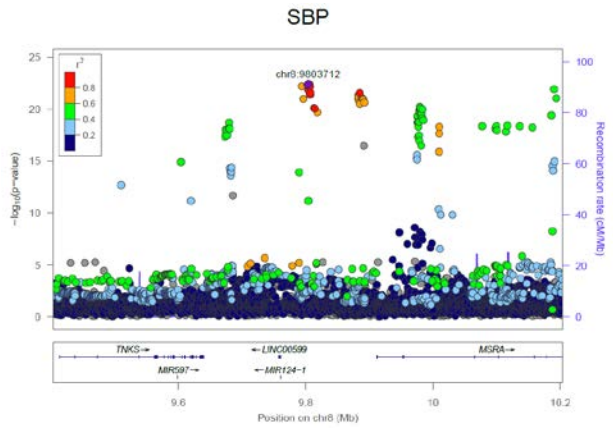
T2 – L9.A



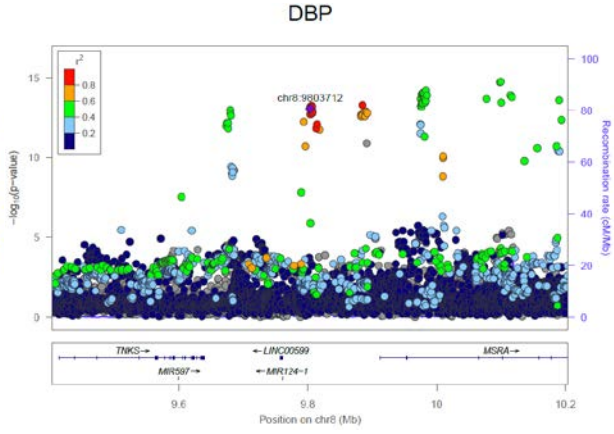
T2 – L9.B



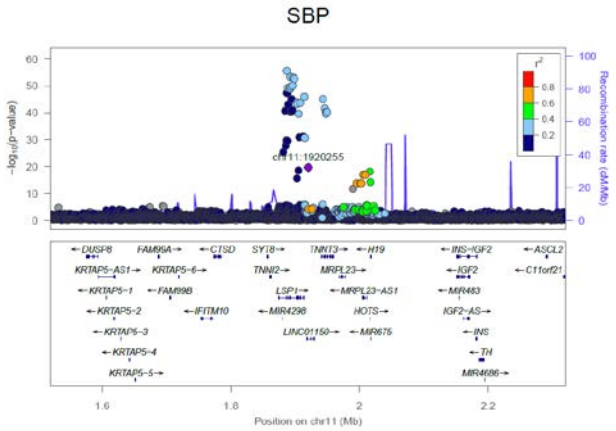
T2 – L10.A



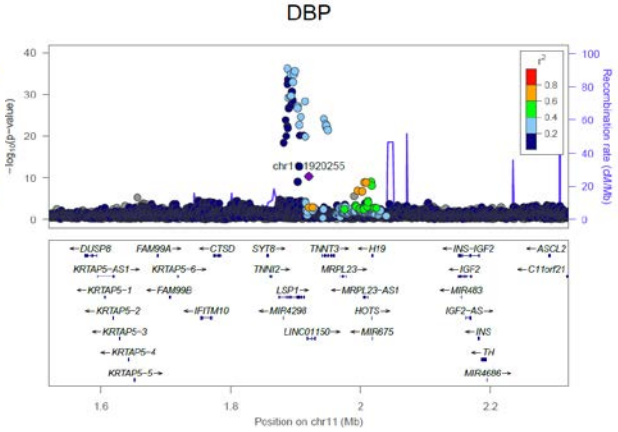
T2 – L10.B



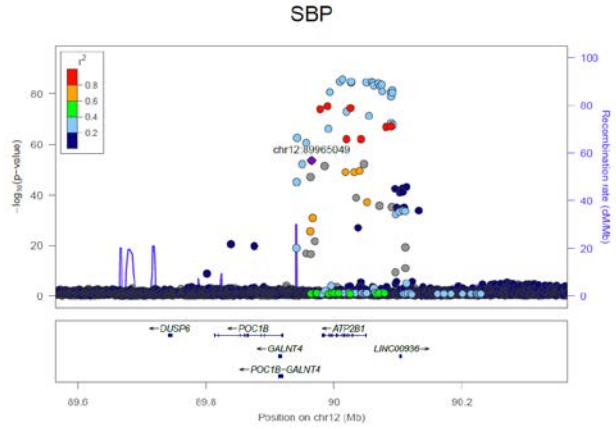
T2 – L11.A



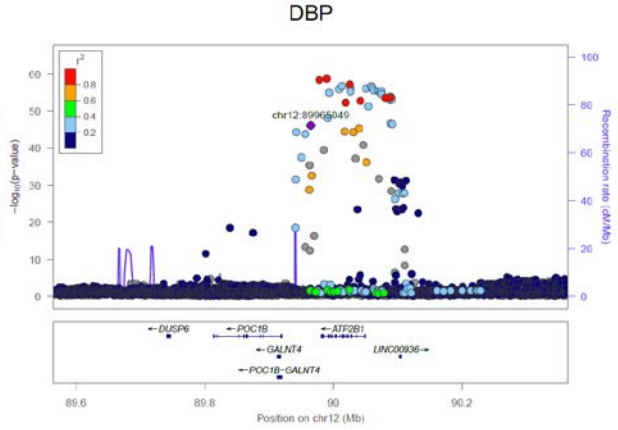
T2 – L11.B



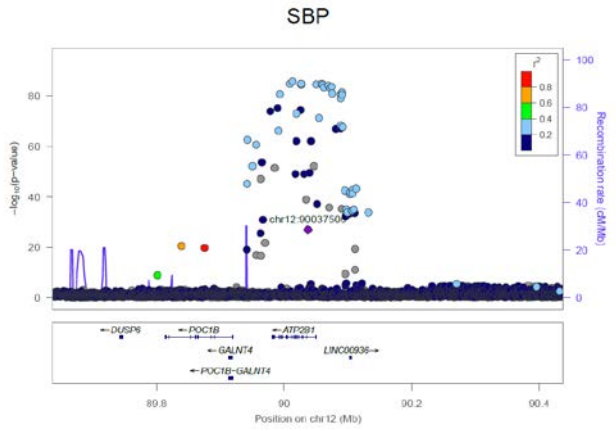
T2 - L12.A



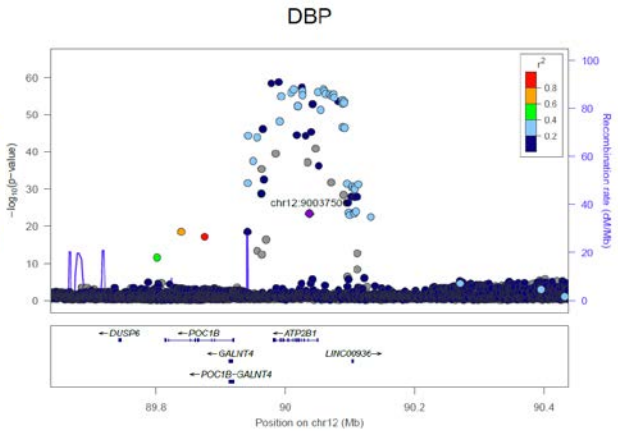
T2 - L12.B



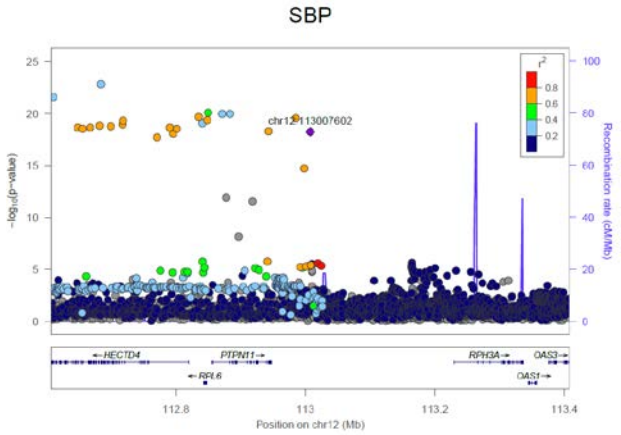
T2 - L13.A



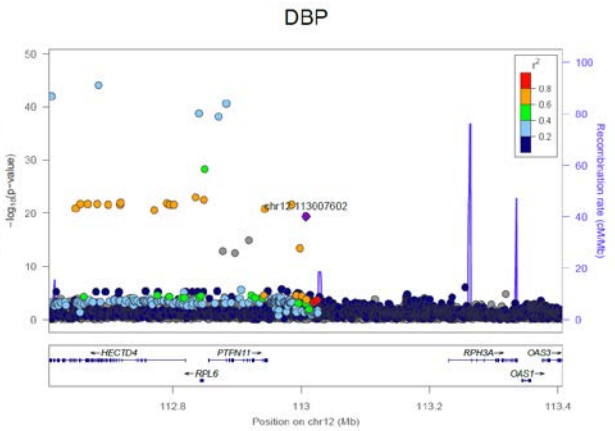
T2 - L13.B



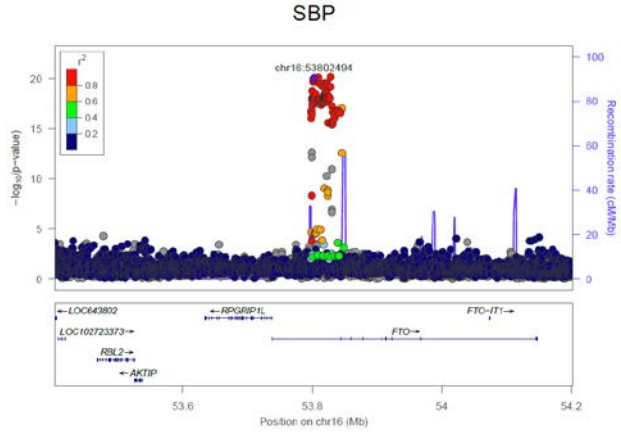
T2 - L14.A



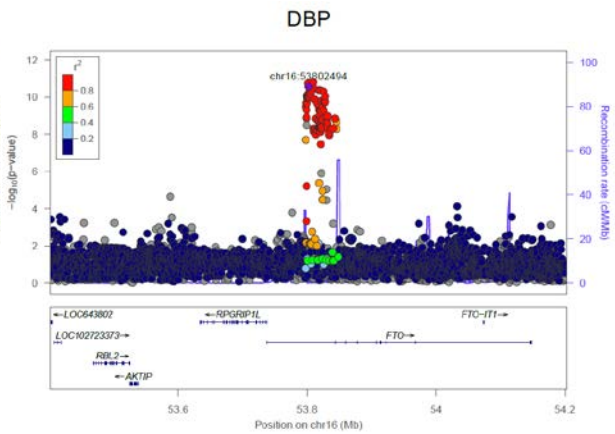
T2 - L14.B



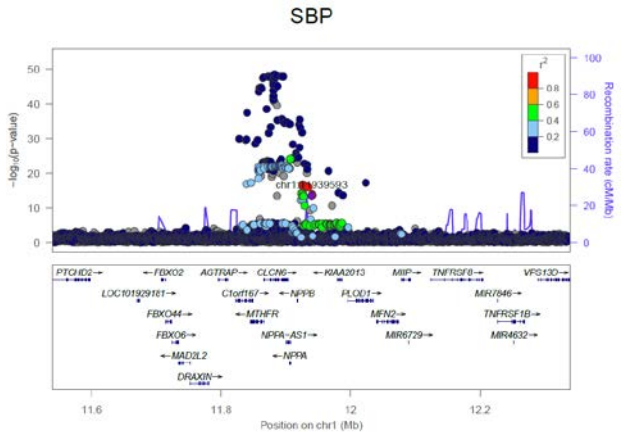
T2 - L15.A



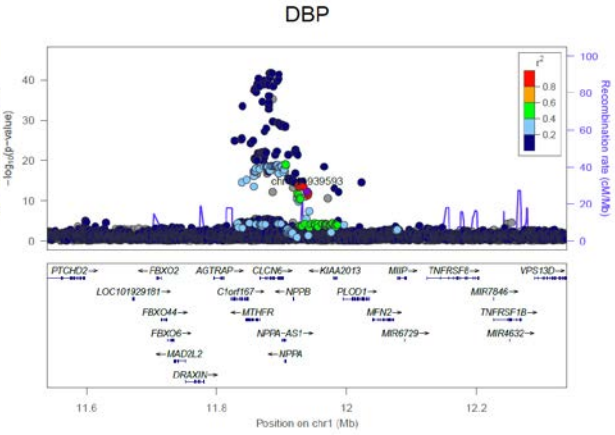
T2 - L15.B



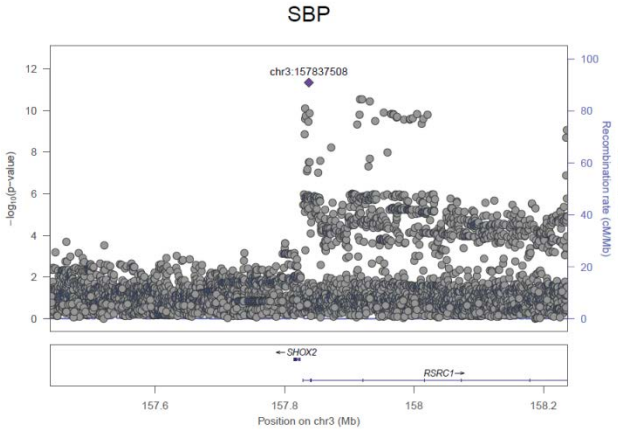
T3 - L1.A



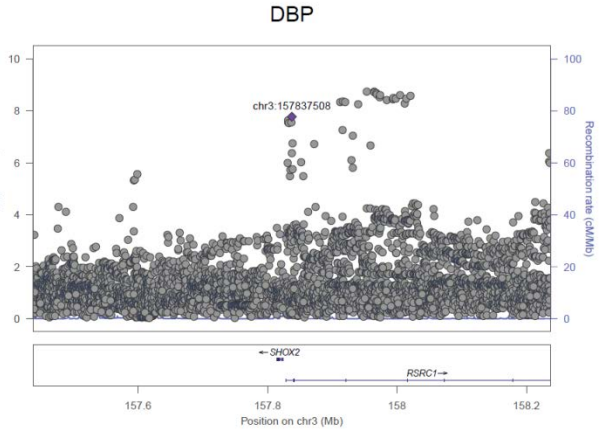
T3 - L1.B



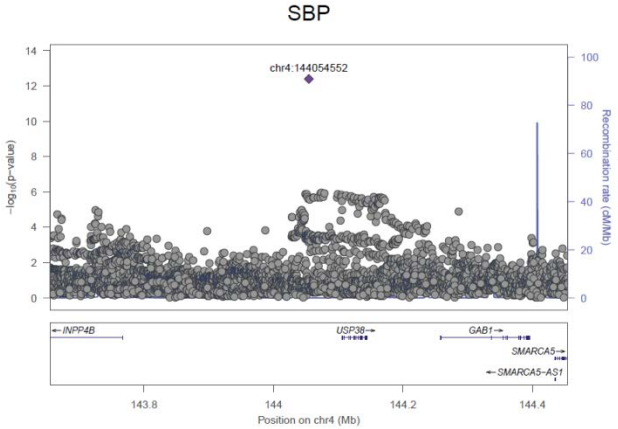
T3 – L2.A



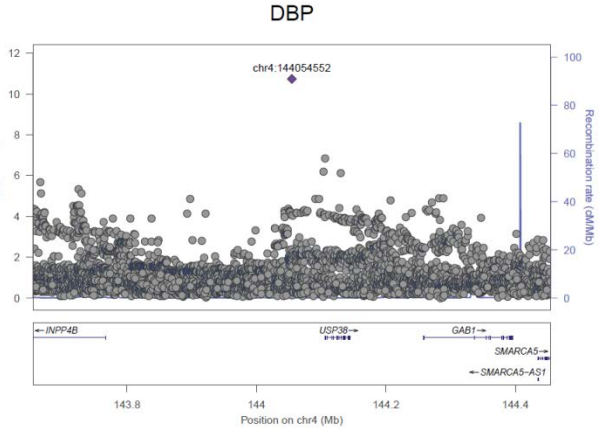
T3 – L2.B



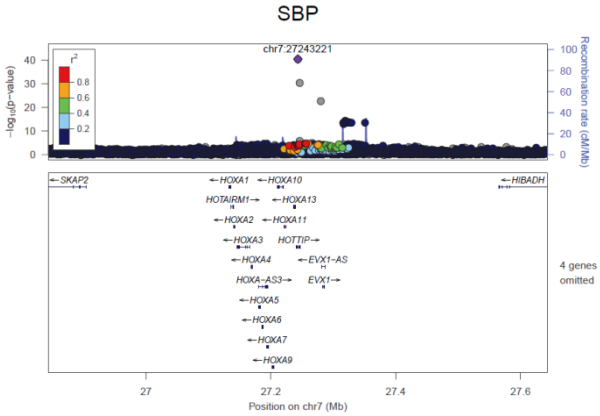
T3 – L3.A



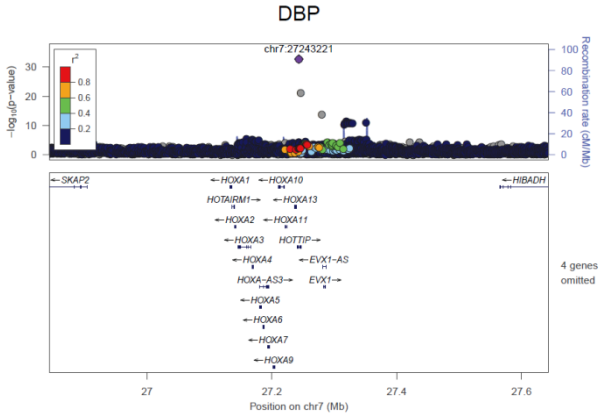
T3 – L3.B



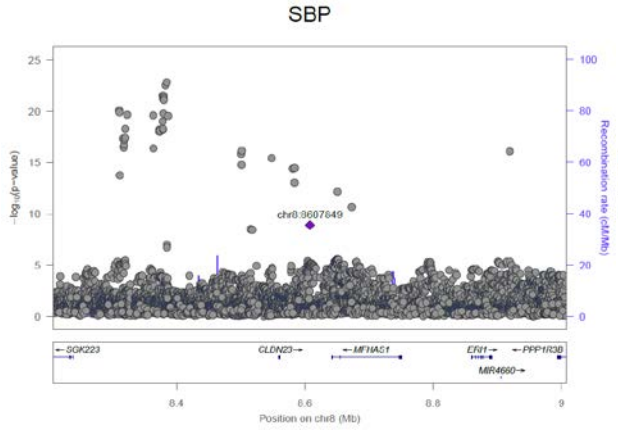
T3 – L4.A



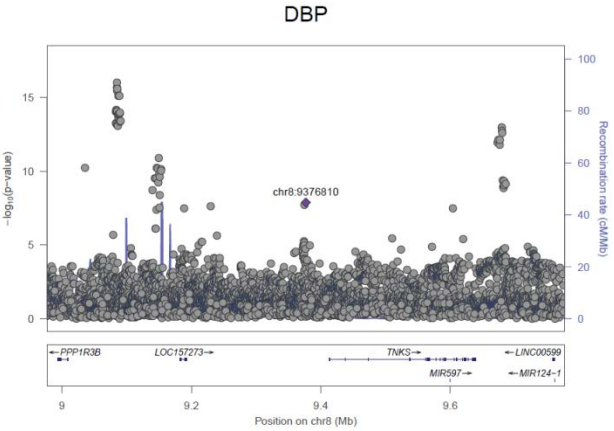
T3 – L4.B



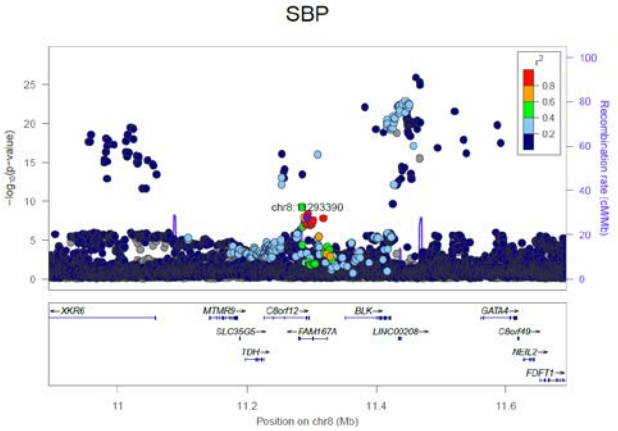
T3 – L5



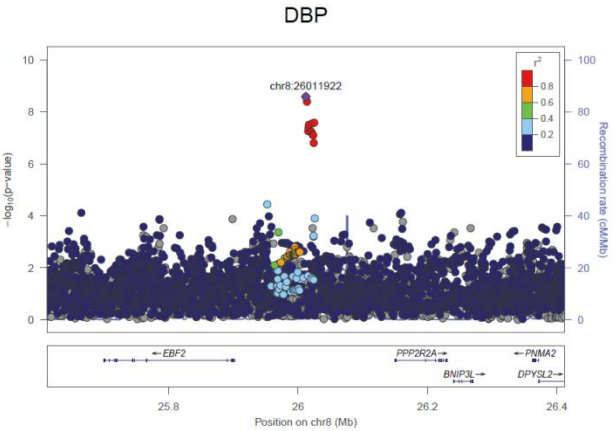
T3 – L6



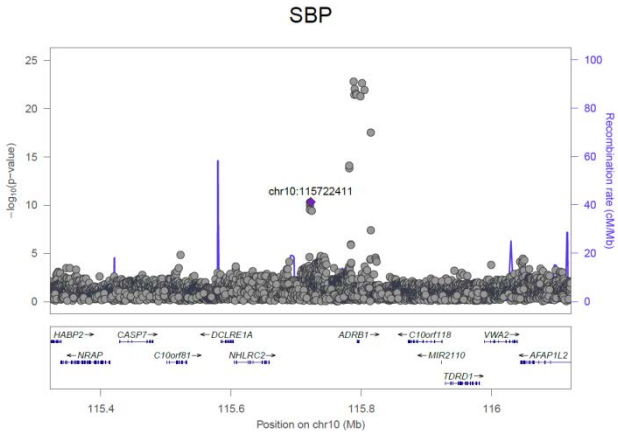
T3 – L7



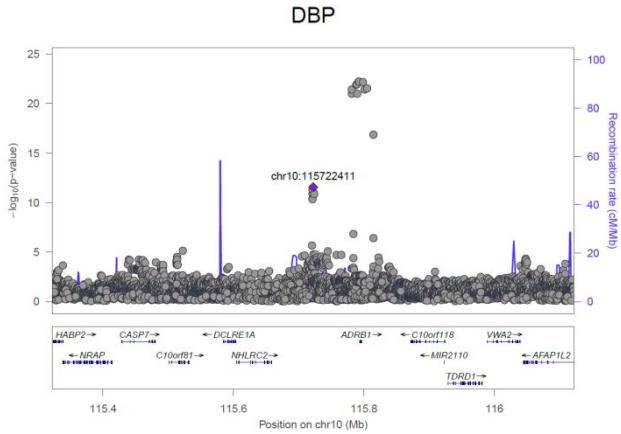
T3 - L8



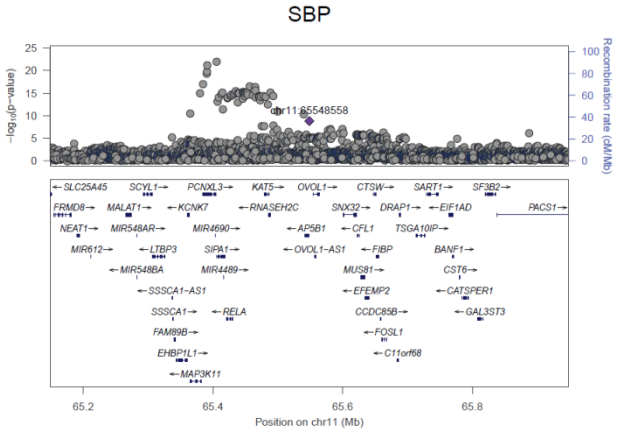
T3 - L9.A



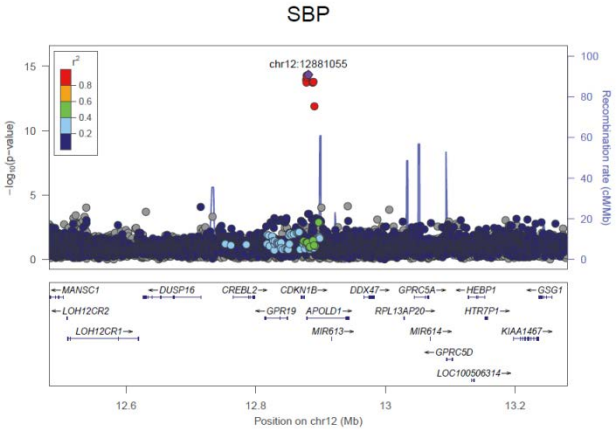
T3 – L9.B



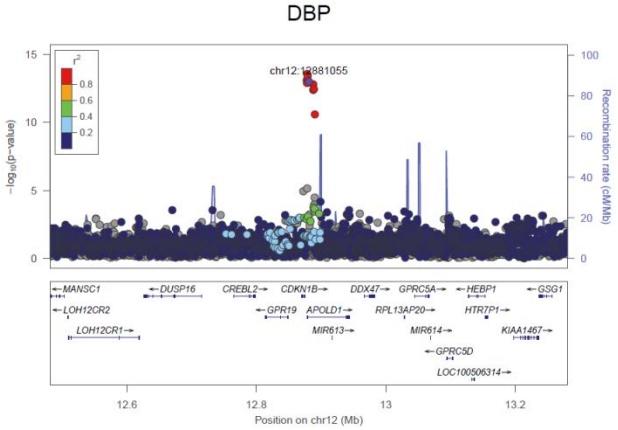
T3 – L10



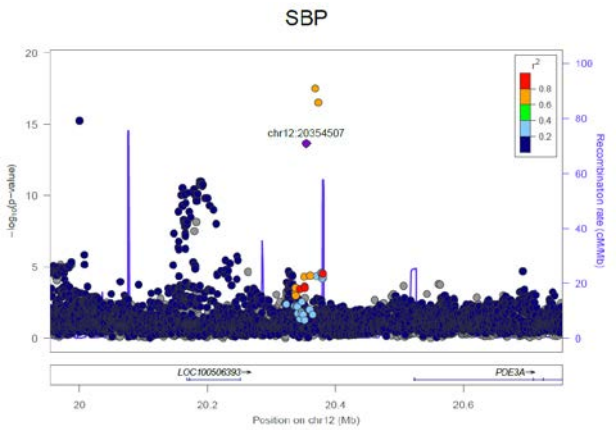
T3 – L11.A



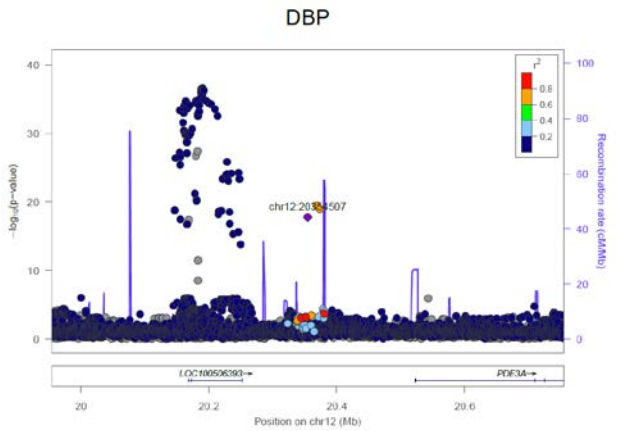
T3 - L11.B



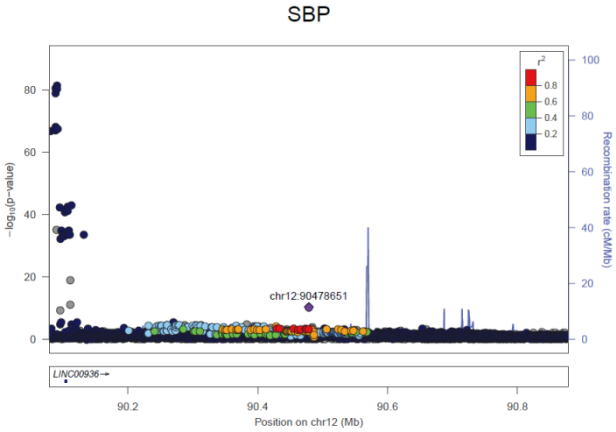
T3 - L12.A



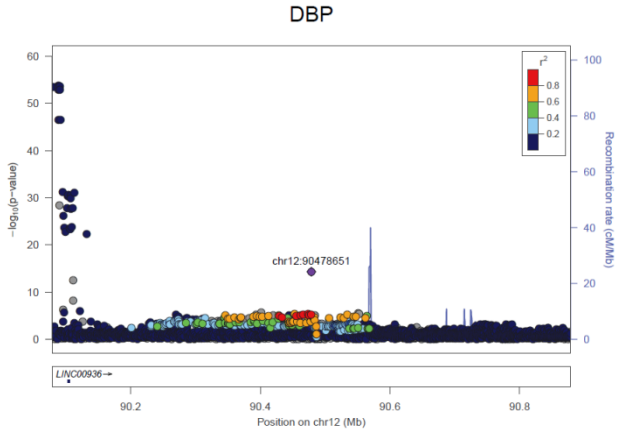
T3 - L12.B



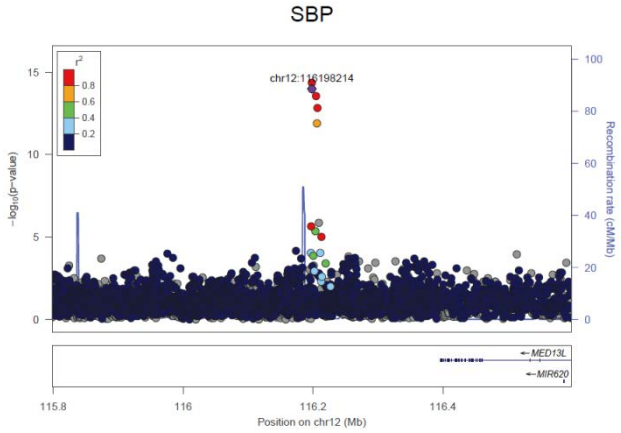
T3 - L13.A



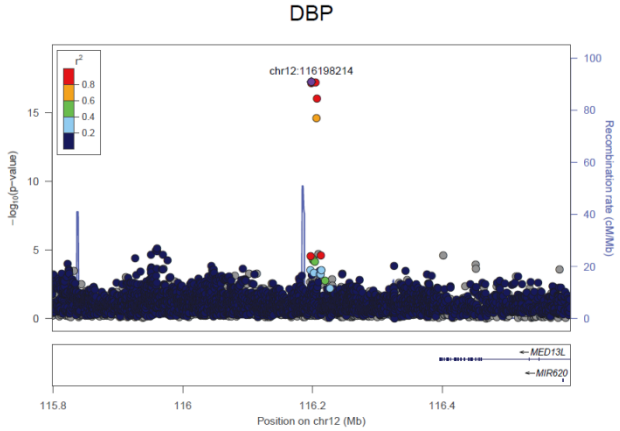
T3 - L13.B



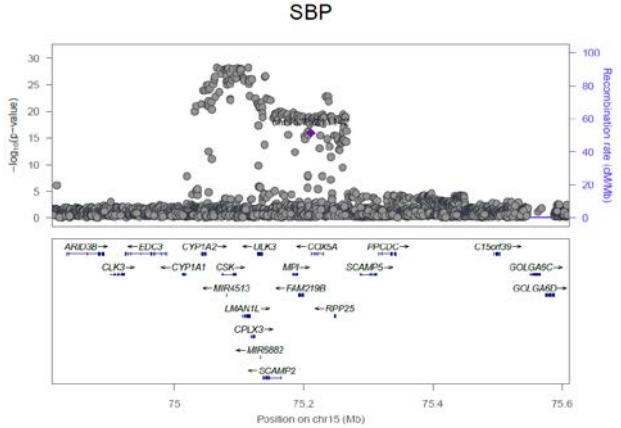
T3 - L14.A



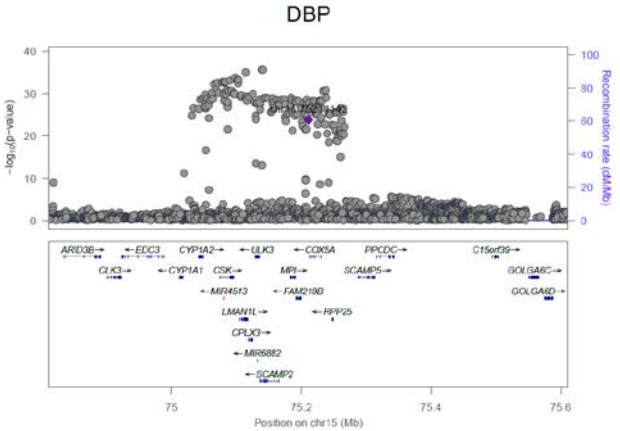
T3 – L14.B



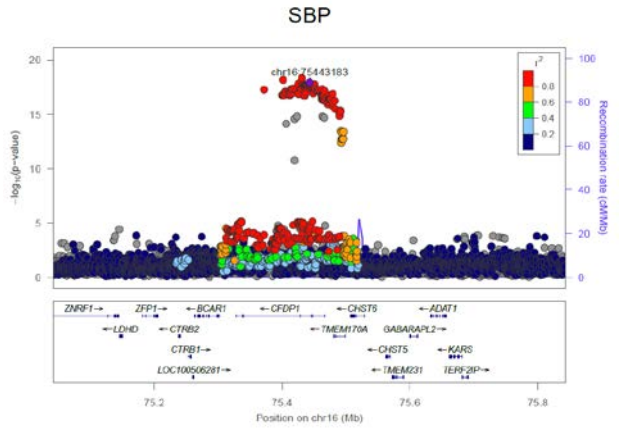
T3 – L15.A



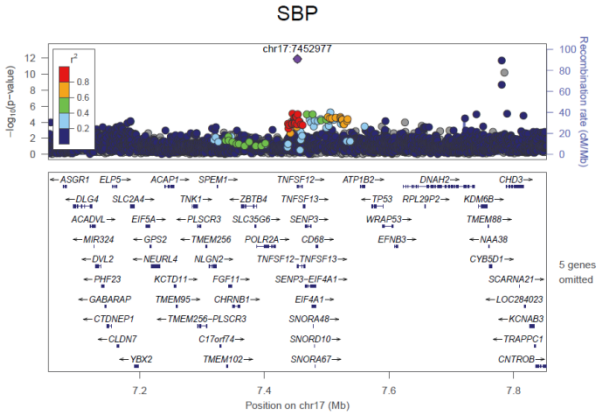
T3 – L15.B



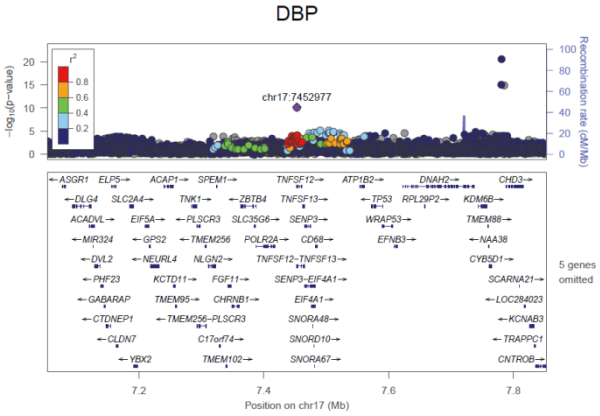
T3 – L16



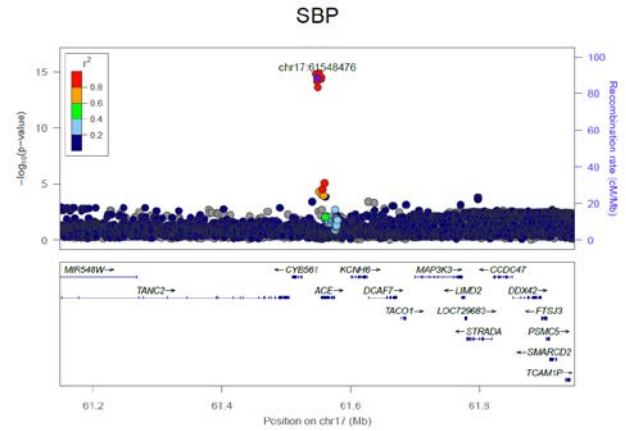
T3 – L17.A



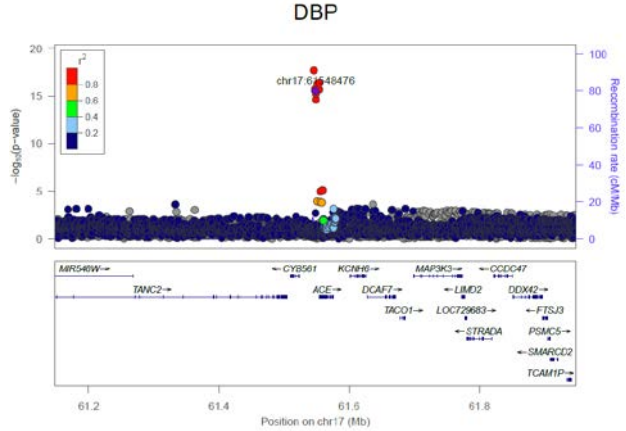
T3 – L17.B



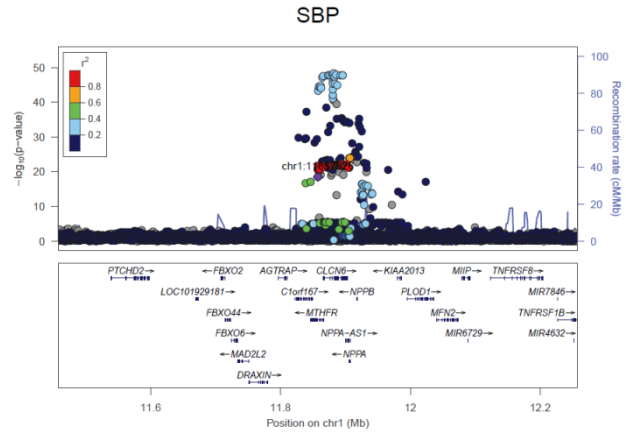
T3 – L18.A



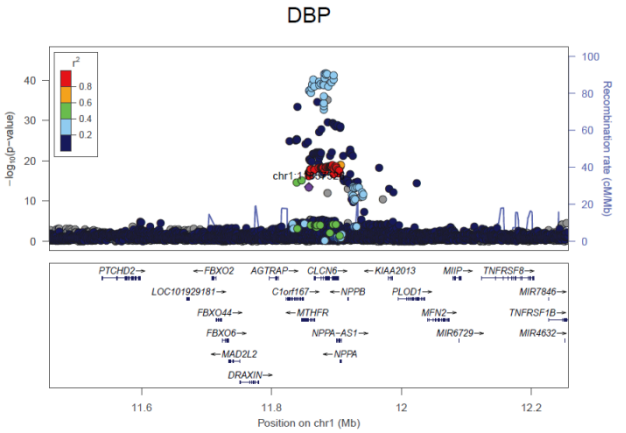
T3 – L18.B



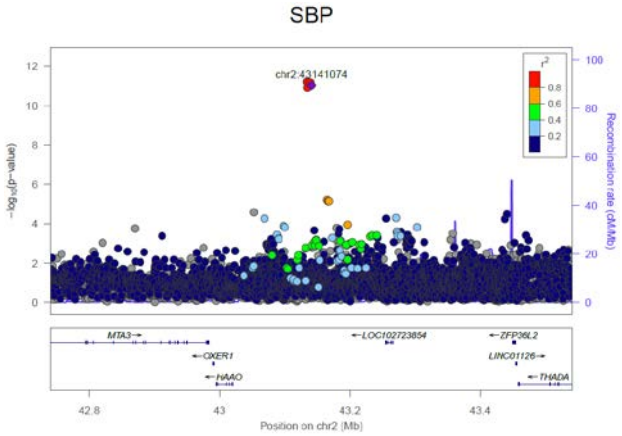
T4 – L1.A



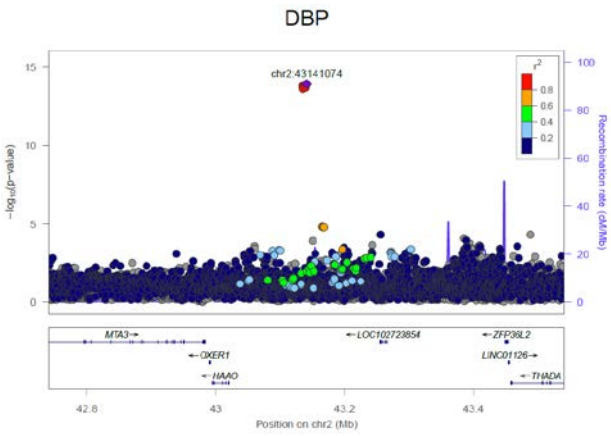
T4 – L1.B



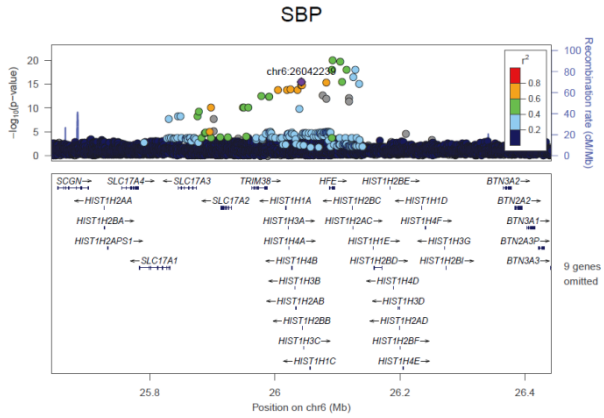
T4 – L2.A



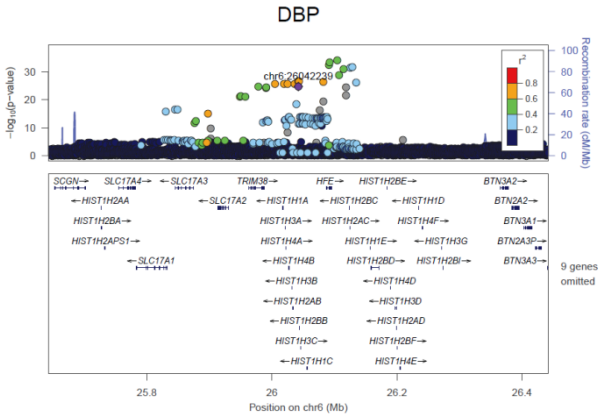
T4 – L2.B



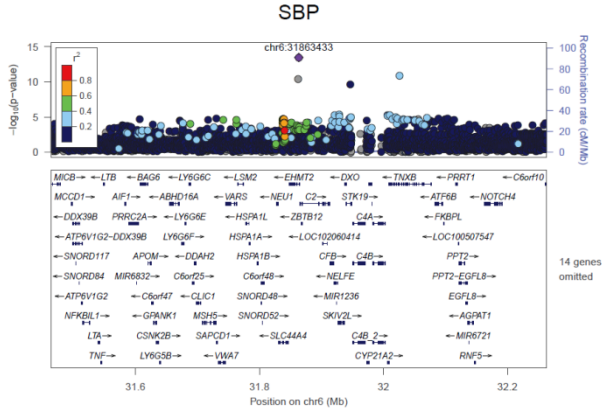
T4 – L3.A



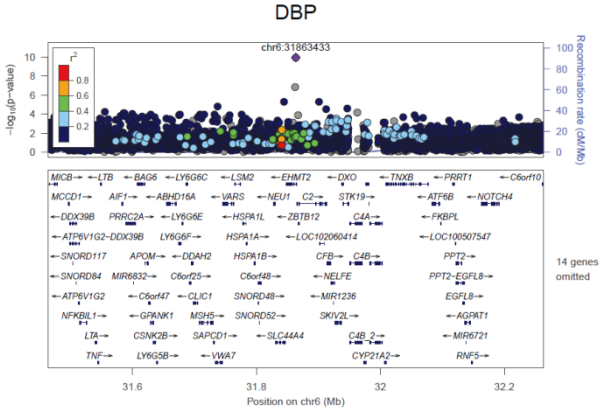
T4 – L3.B



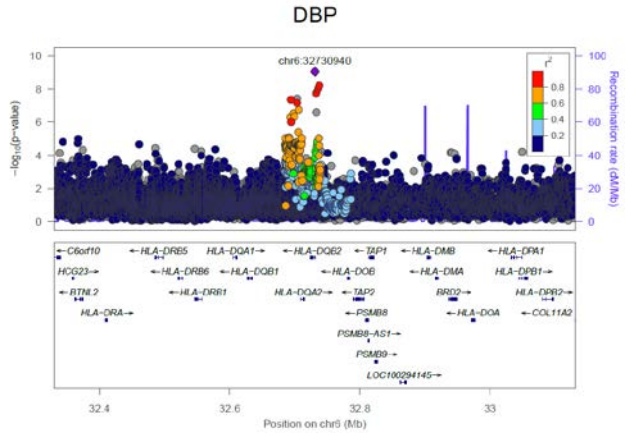
T4 – L4.A



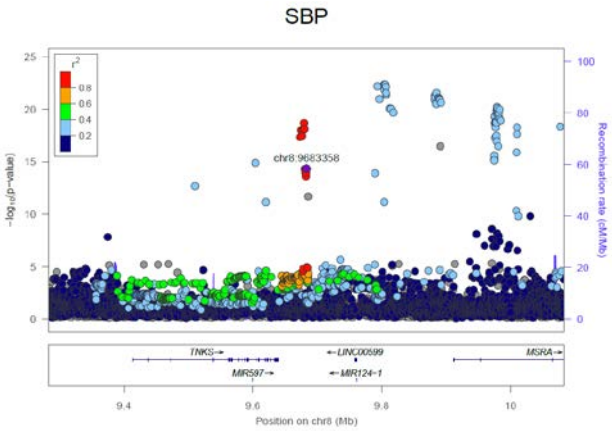
T4 – L4.B



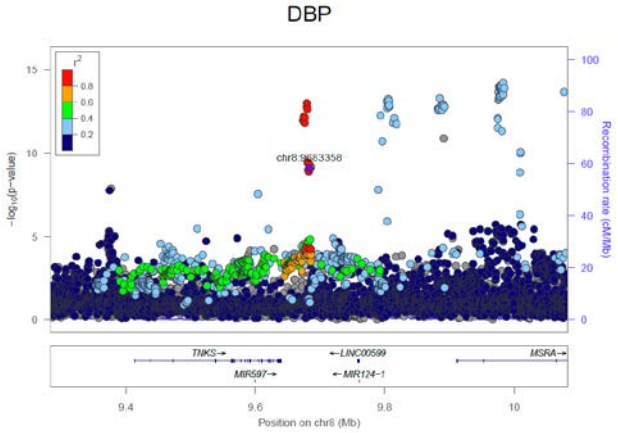
T4 – L5



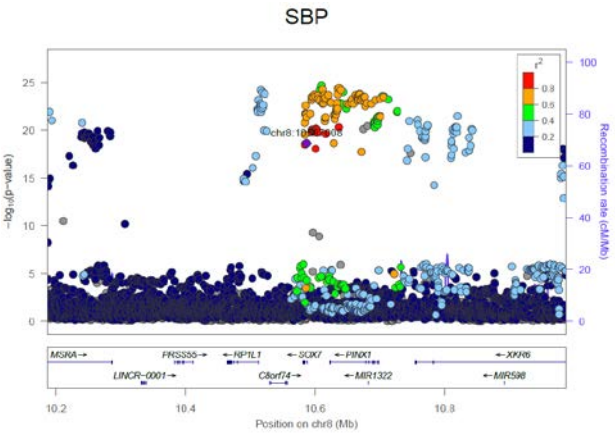
T4 – L6.A



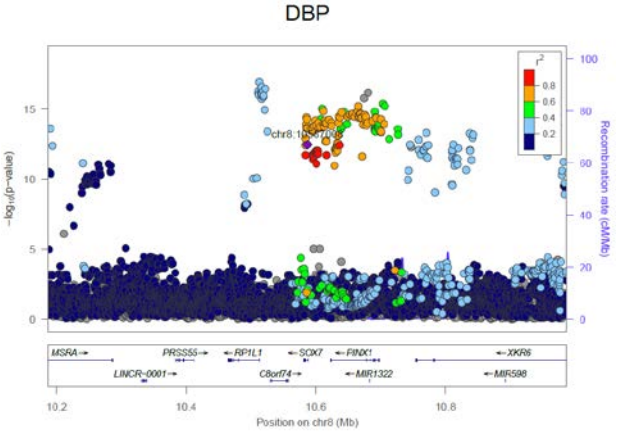
T4 – L6.B



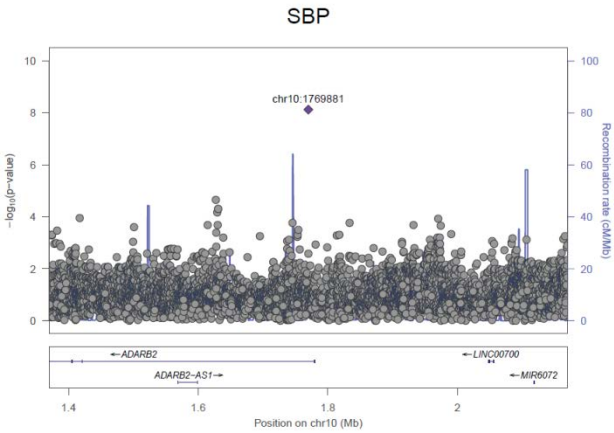
T4 – L7.A



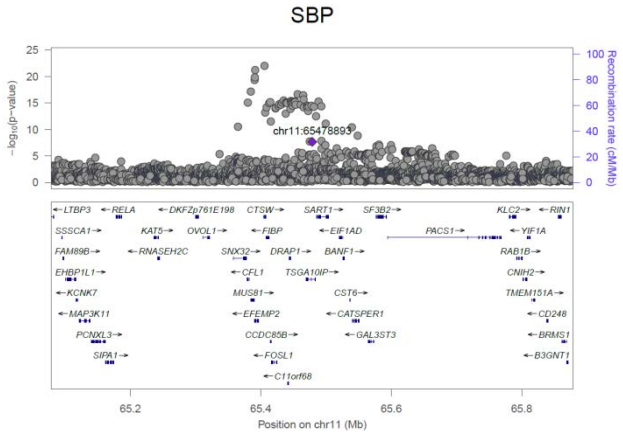
T4 – L7.B



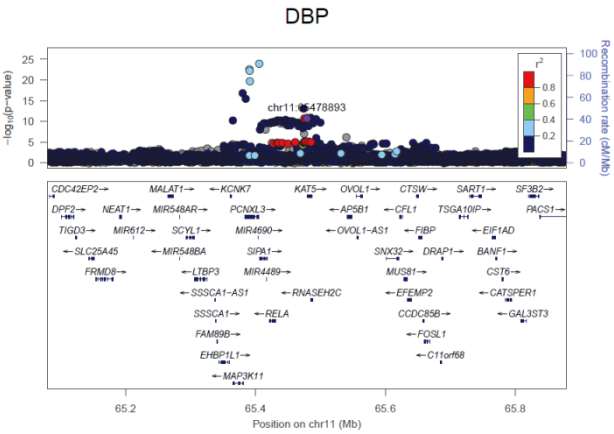
T4 – L8



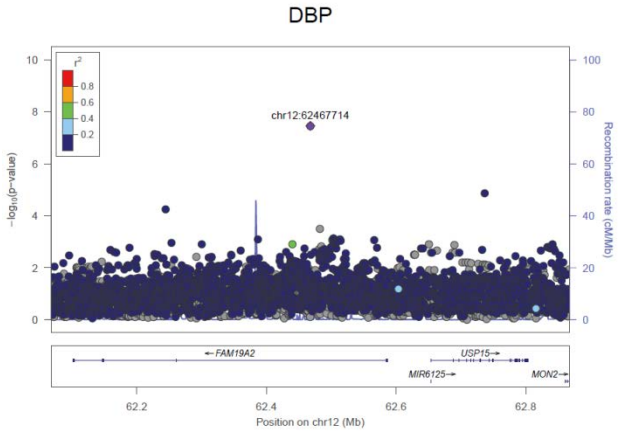
T4 – L8



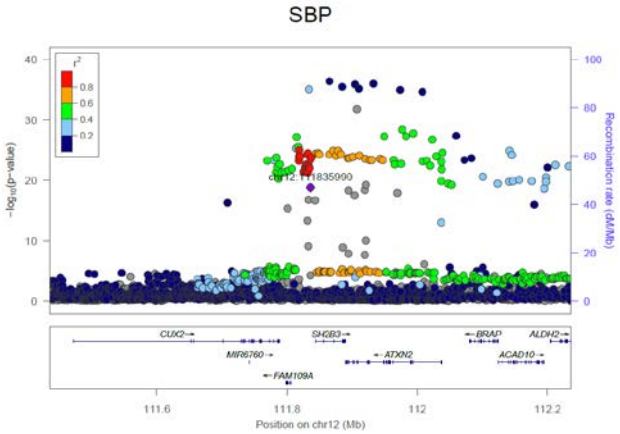
T4 – L9



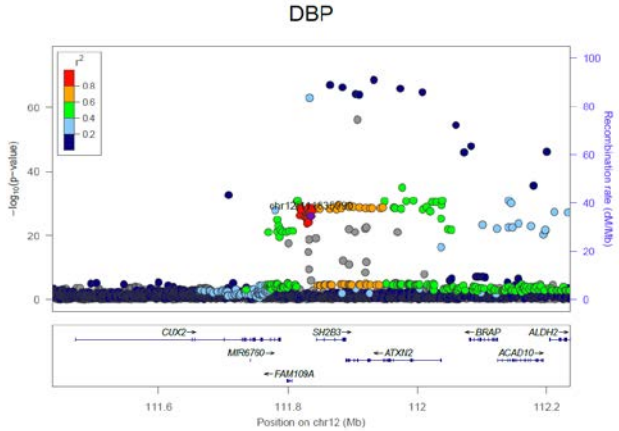
T4 – L10



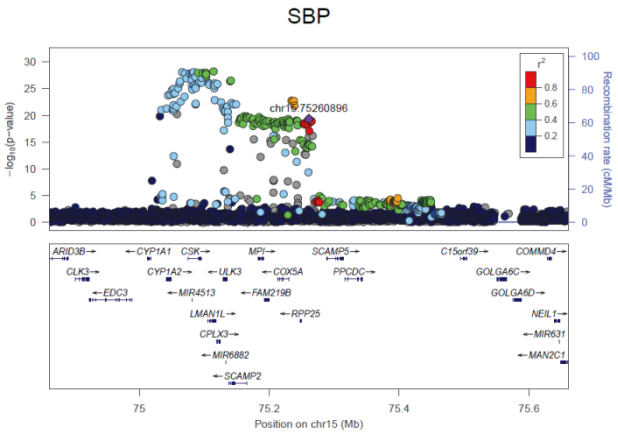
T4 – L11.A



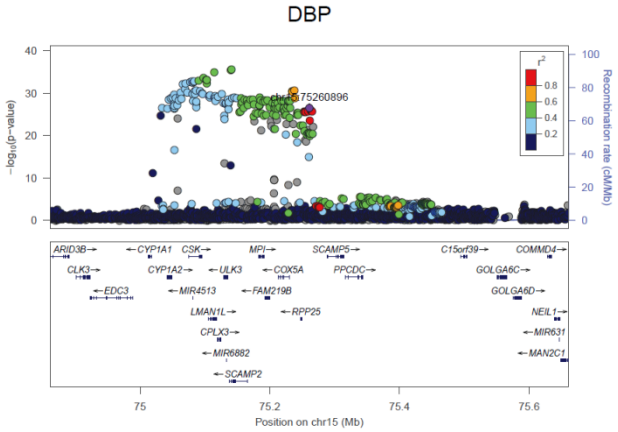
T4 – L11.B



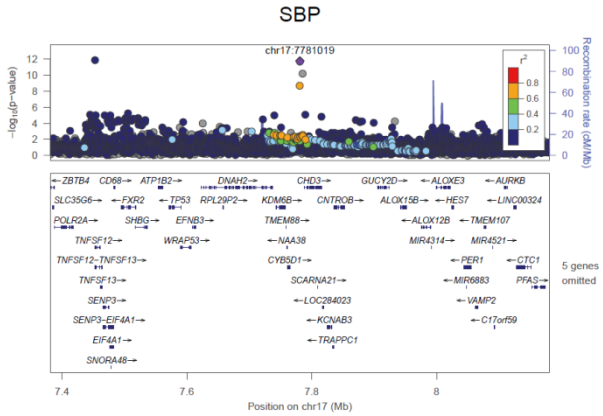
T4 – L12.A



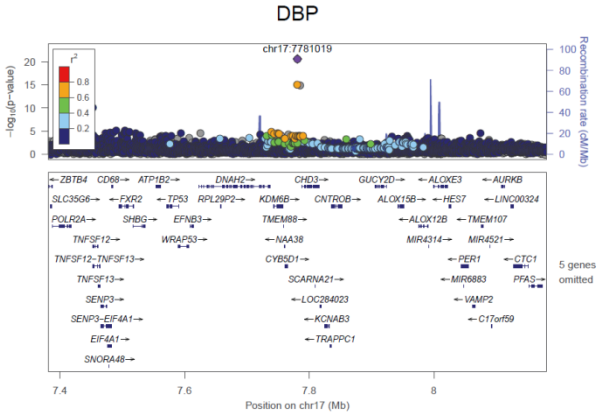
T4 – L12.B



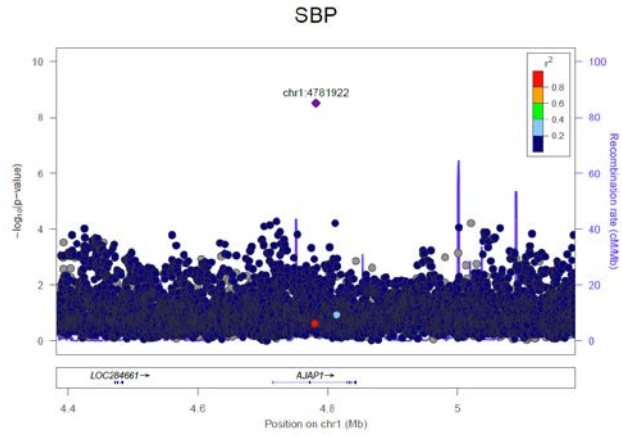
T4 – L13.A



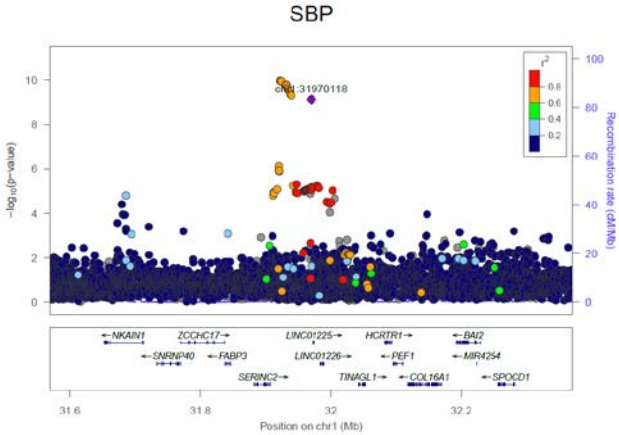
T4 – L13.B



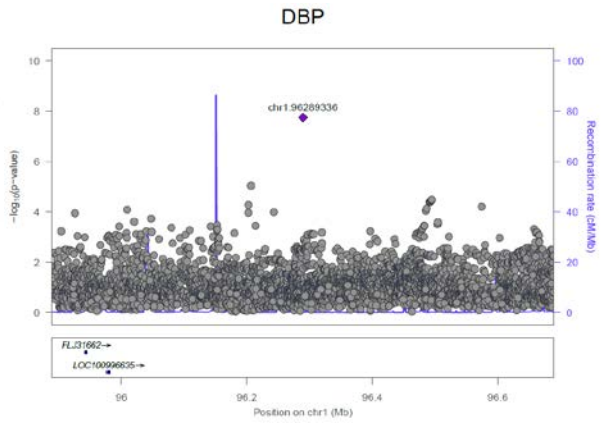
T5 – L1



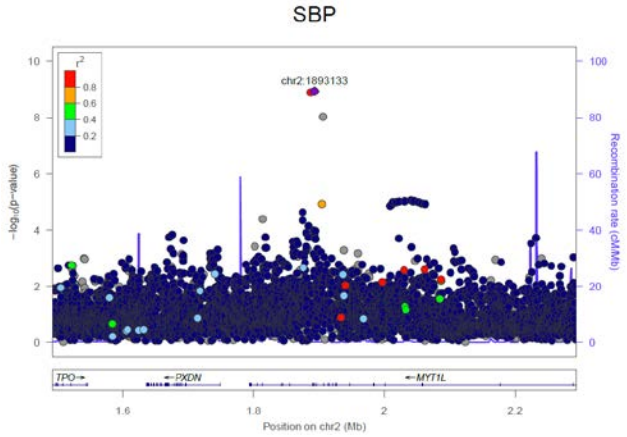
T5 - L2



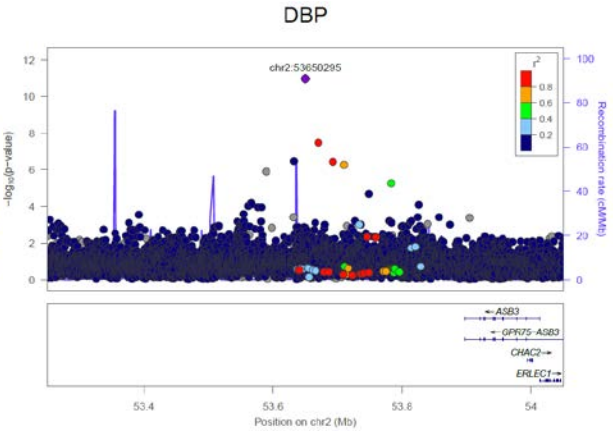
T5 - L3



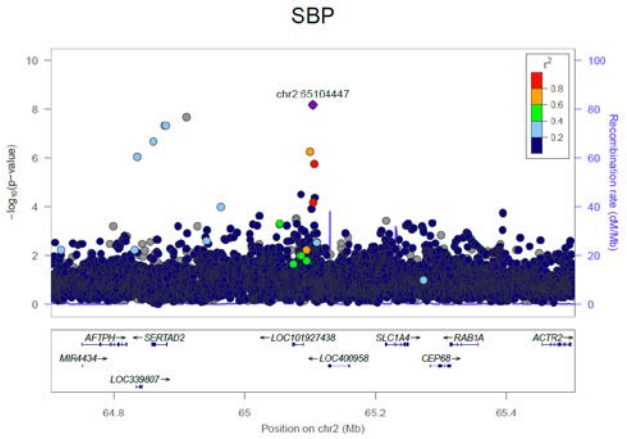
T5 - L4



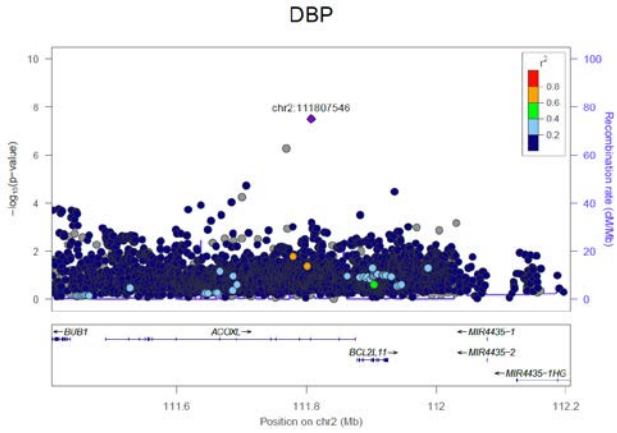
T5 - L5



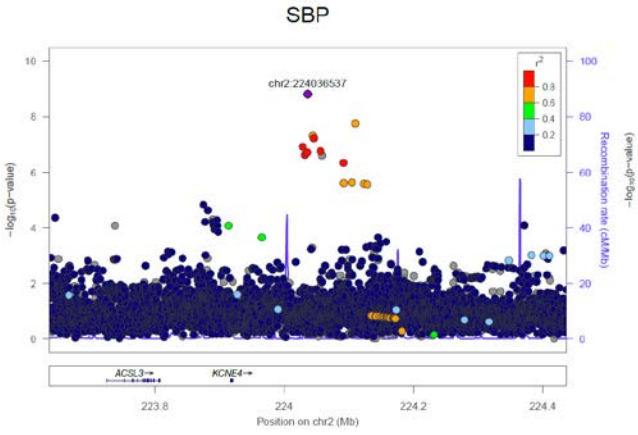
T5 - L6



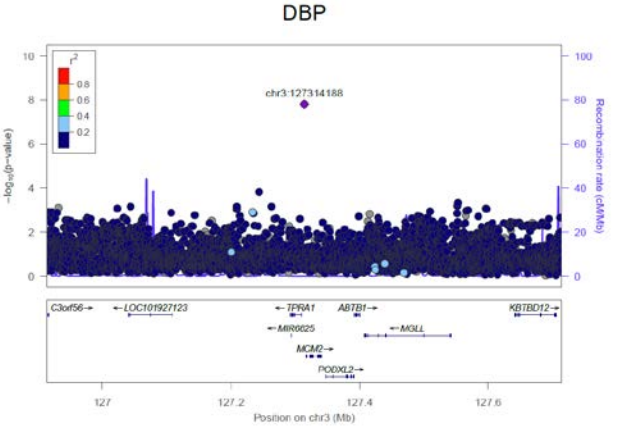
T5 - L7



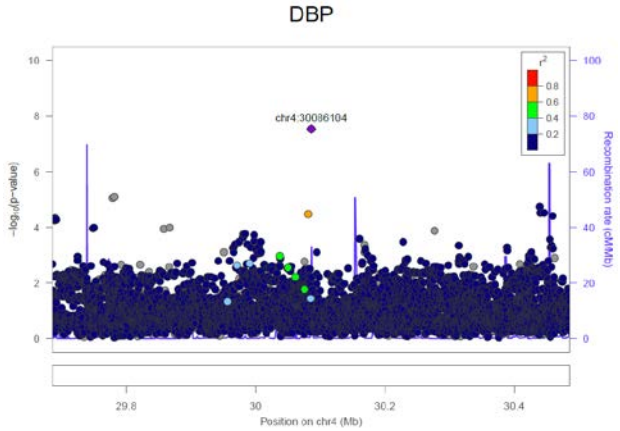
T5 - L8



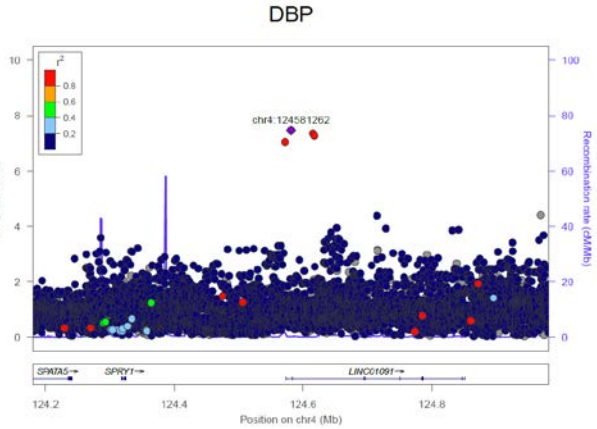
T5 - L9



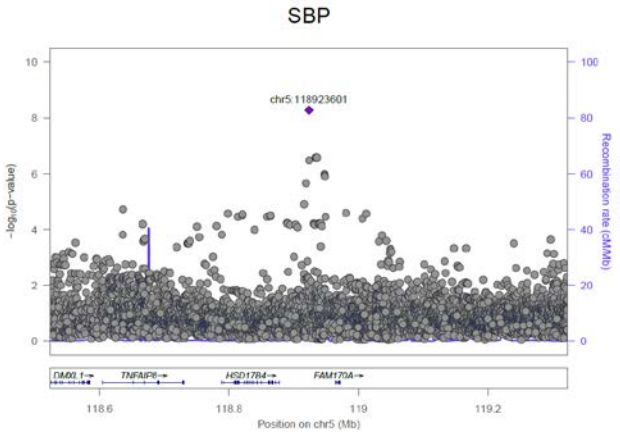
T5 - L10



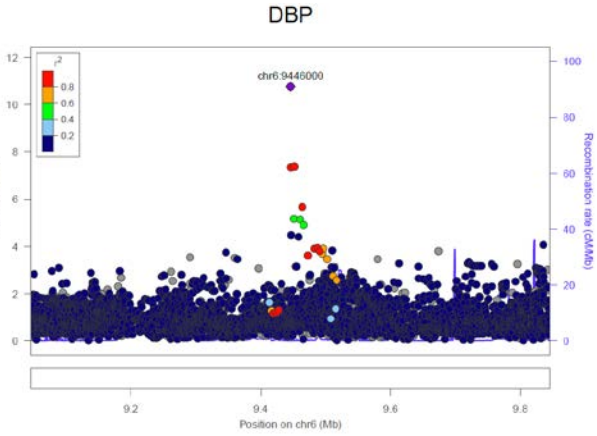
T5 - L11



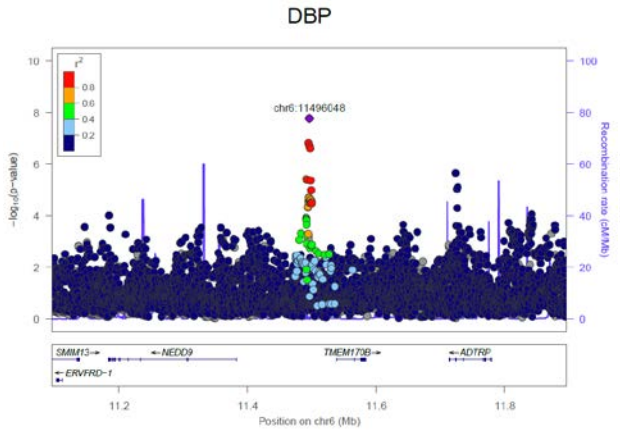
T5 - L12



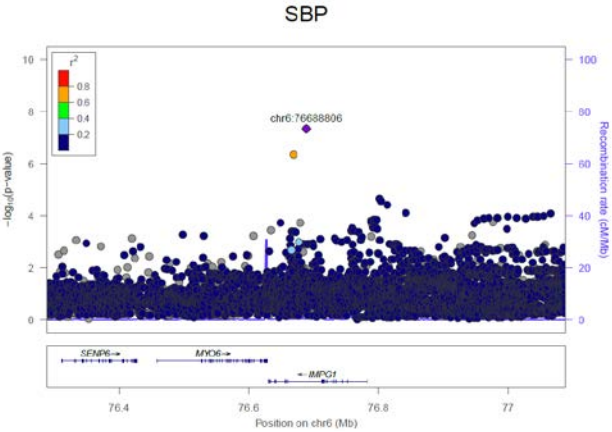
T5 - L13



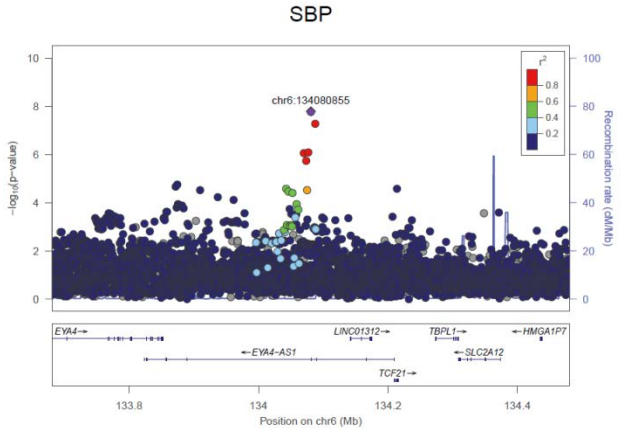
T5 - L14



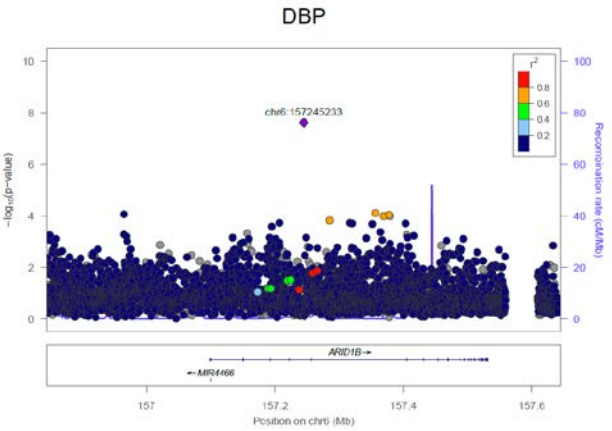
T5 - L15



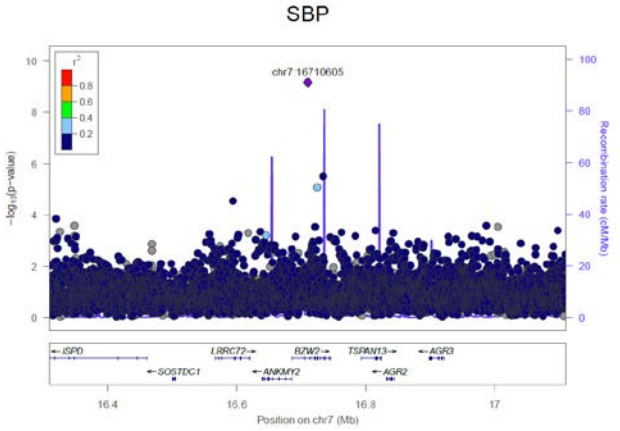
T5 - L16



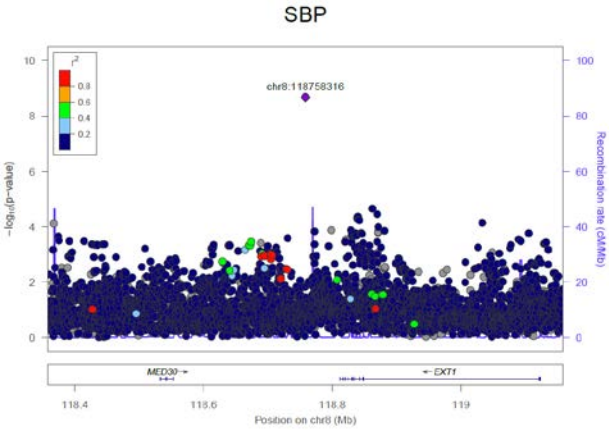
T5 - L17



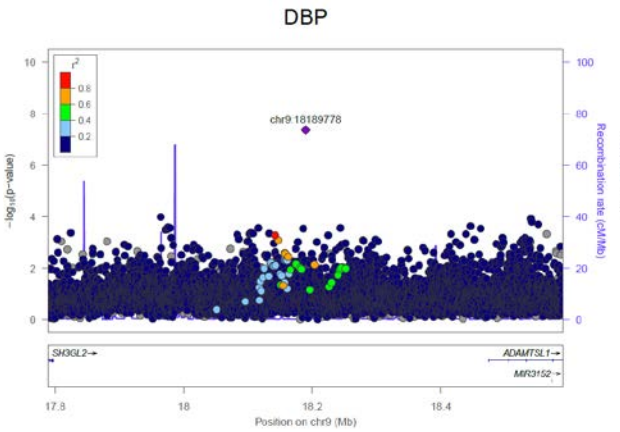
T5 - L18



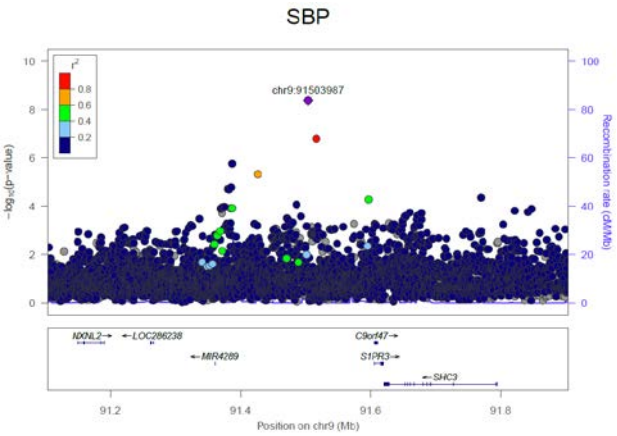
T5 - L19



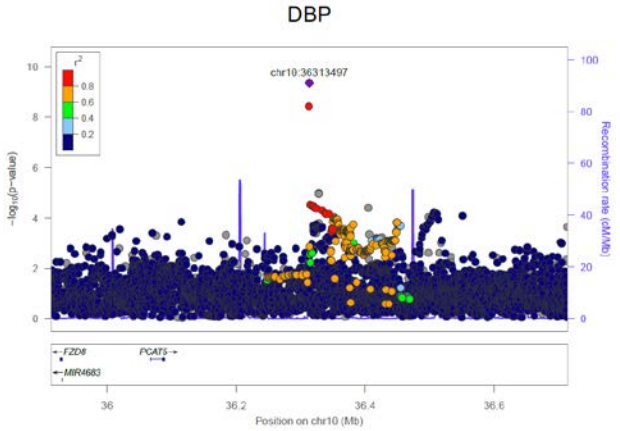
T5 – L20



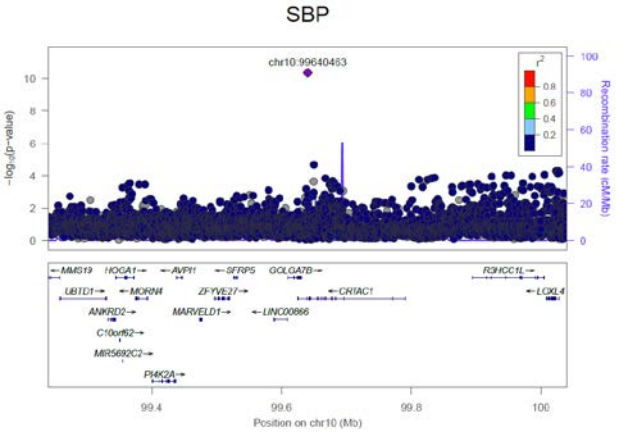
T5 – L21



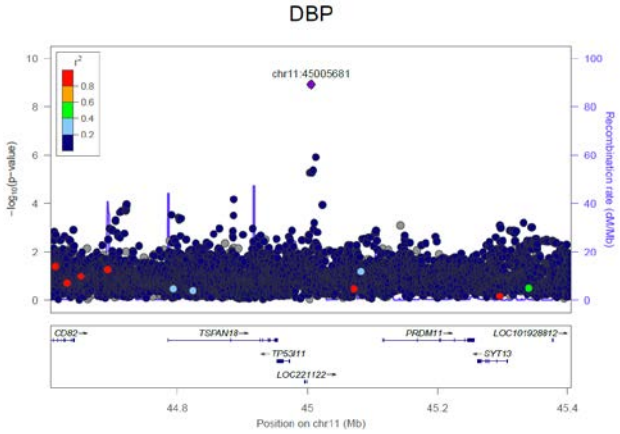
T5 - L22



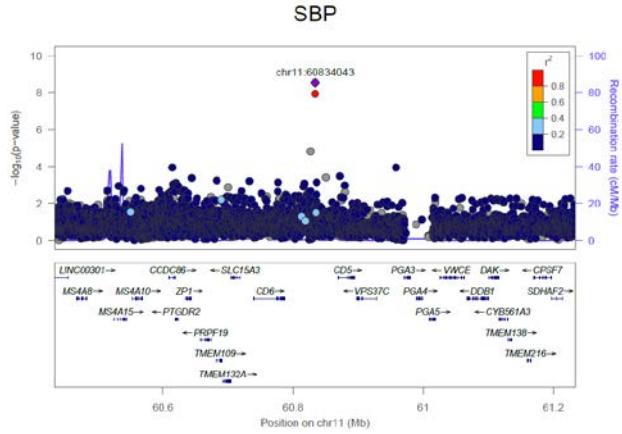
T5 - L23



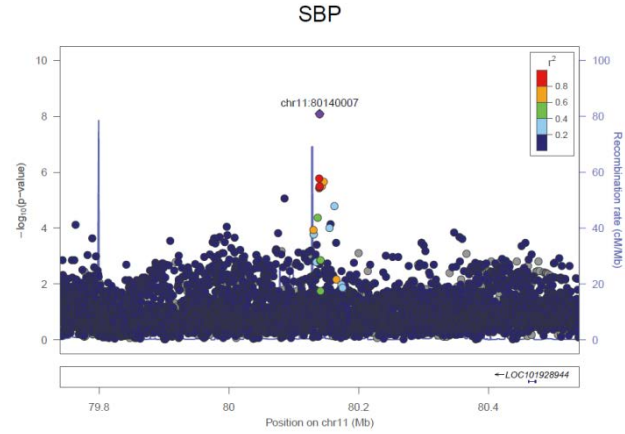
T5 - L24



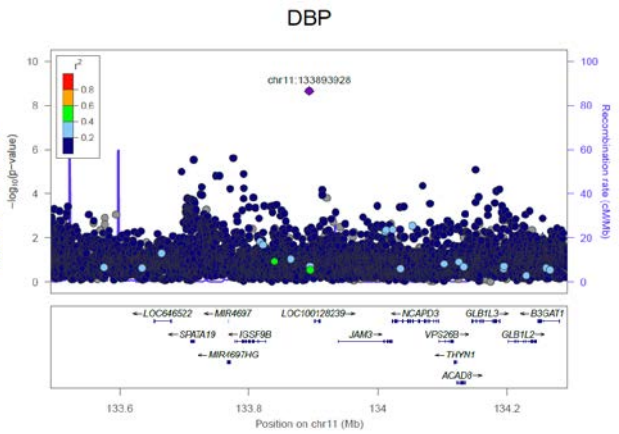
T5 - L25



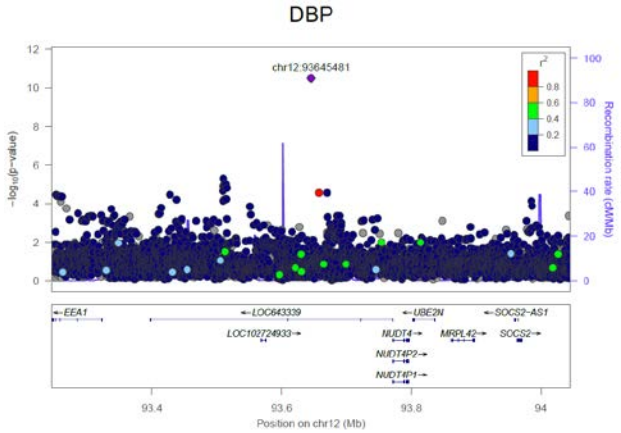
T5 - L26



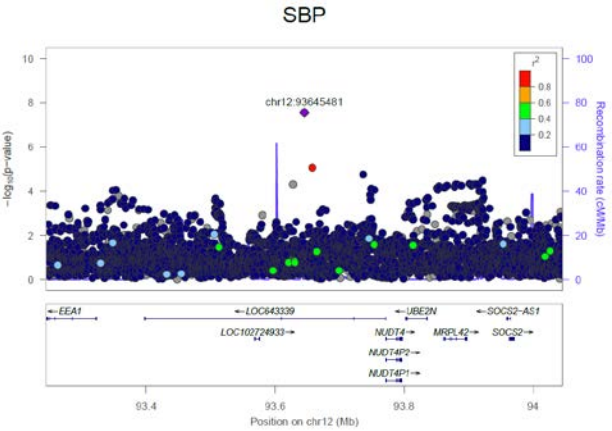
T5 - L27



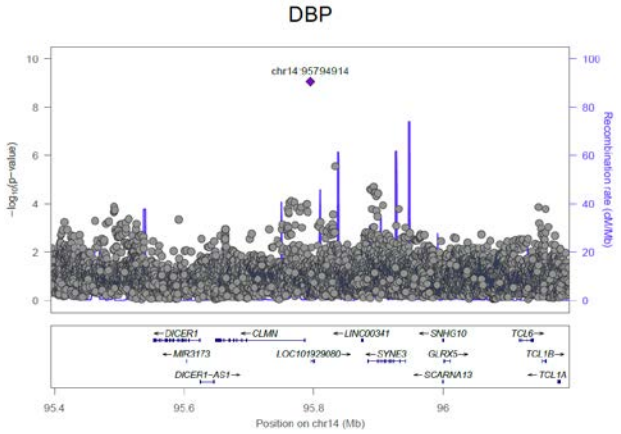
T5 - L28.A



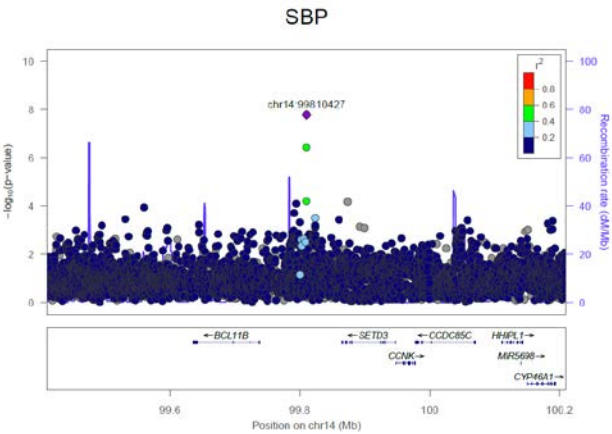
T5 - L28.B



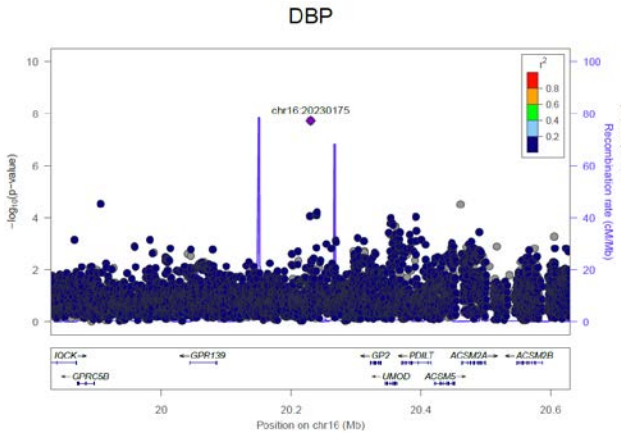
T5 - L29



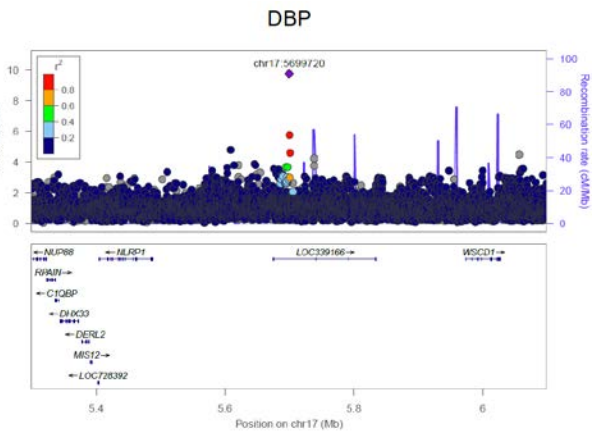
T5 – L30



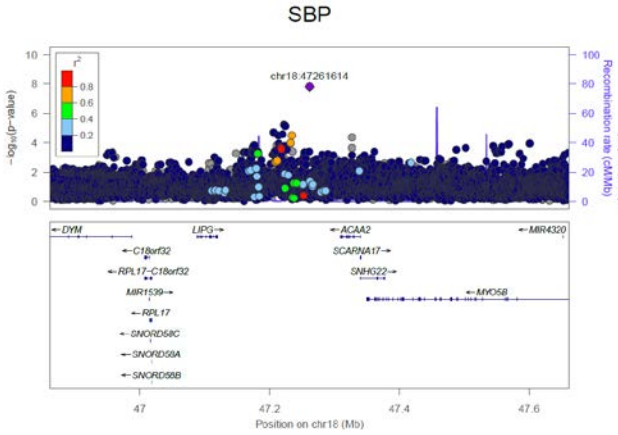
T5 – L31



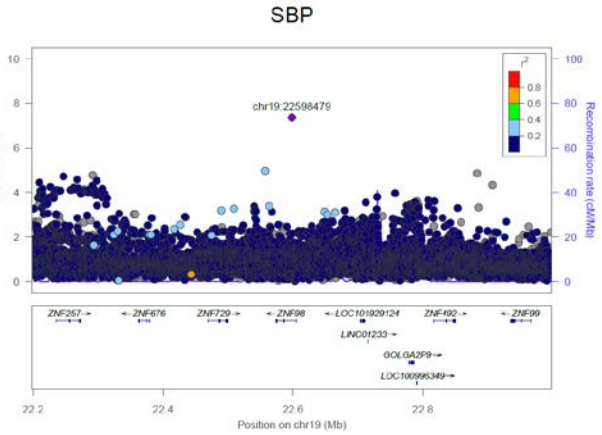
T5 - L32



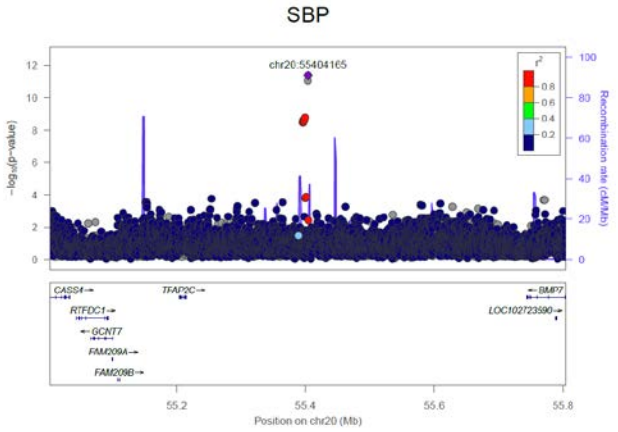
T5 - L33



T5 - L34



T5 - L35



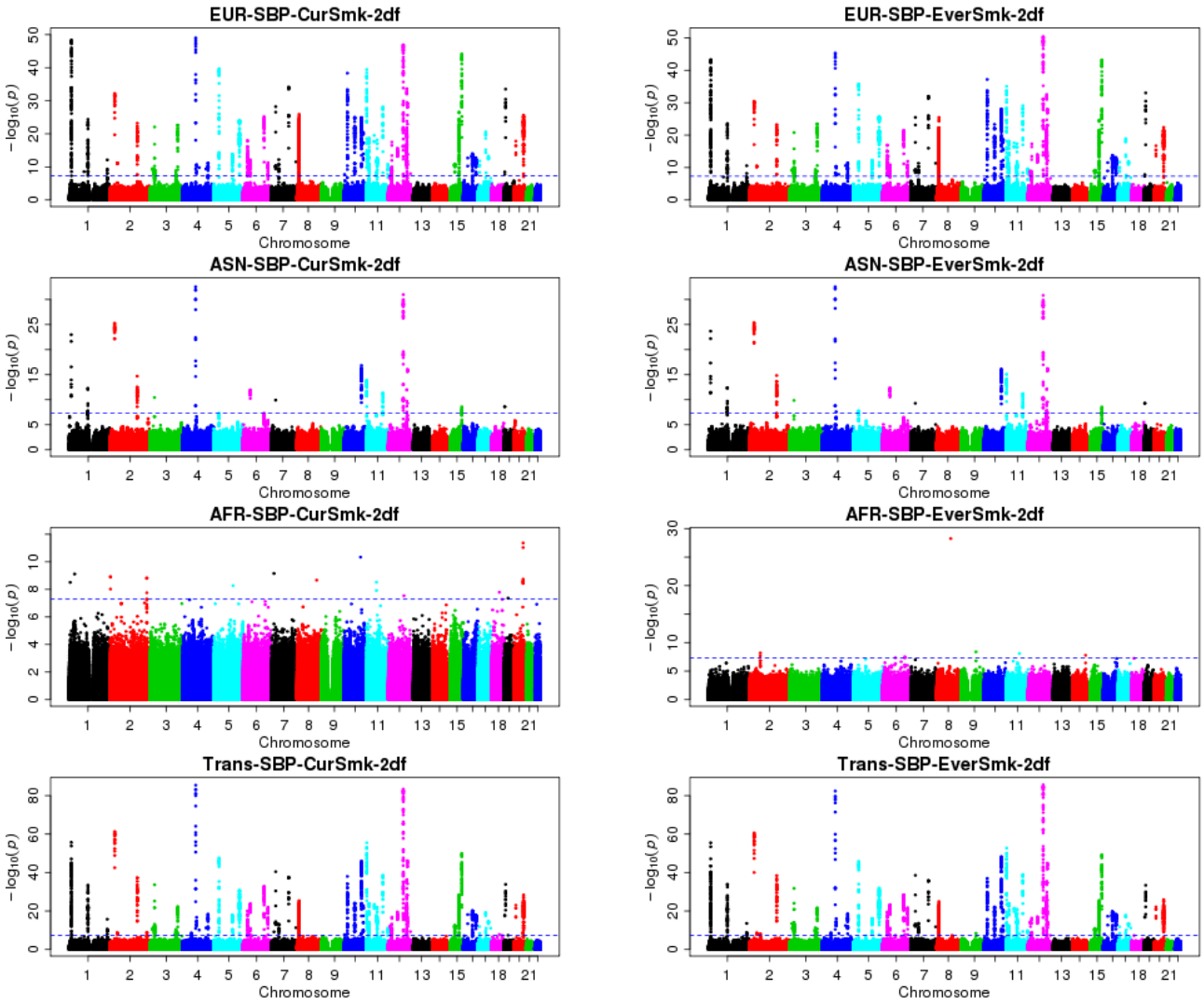


Figure S5: Manhattan plots of SBP using the 2 DF joint test.

The $-\log_{10}(p)$ of each SNP was plotted at the chromosomal location of each variant. The p-values are based on the combined discovery and replication analysis for the select 4,459 variants and the discovery analysis for the remaining 18.8 million variants.

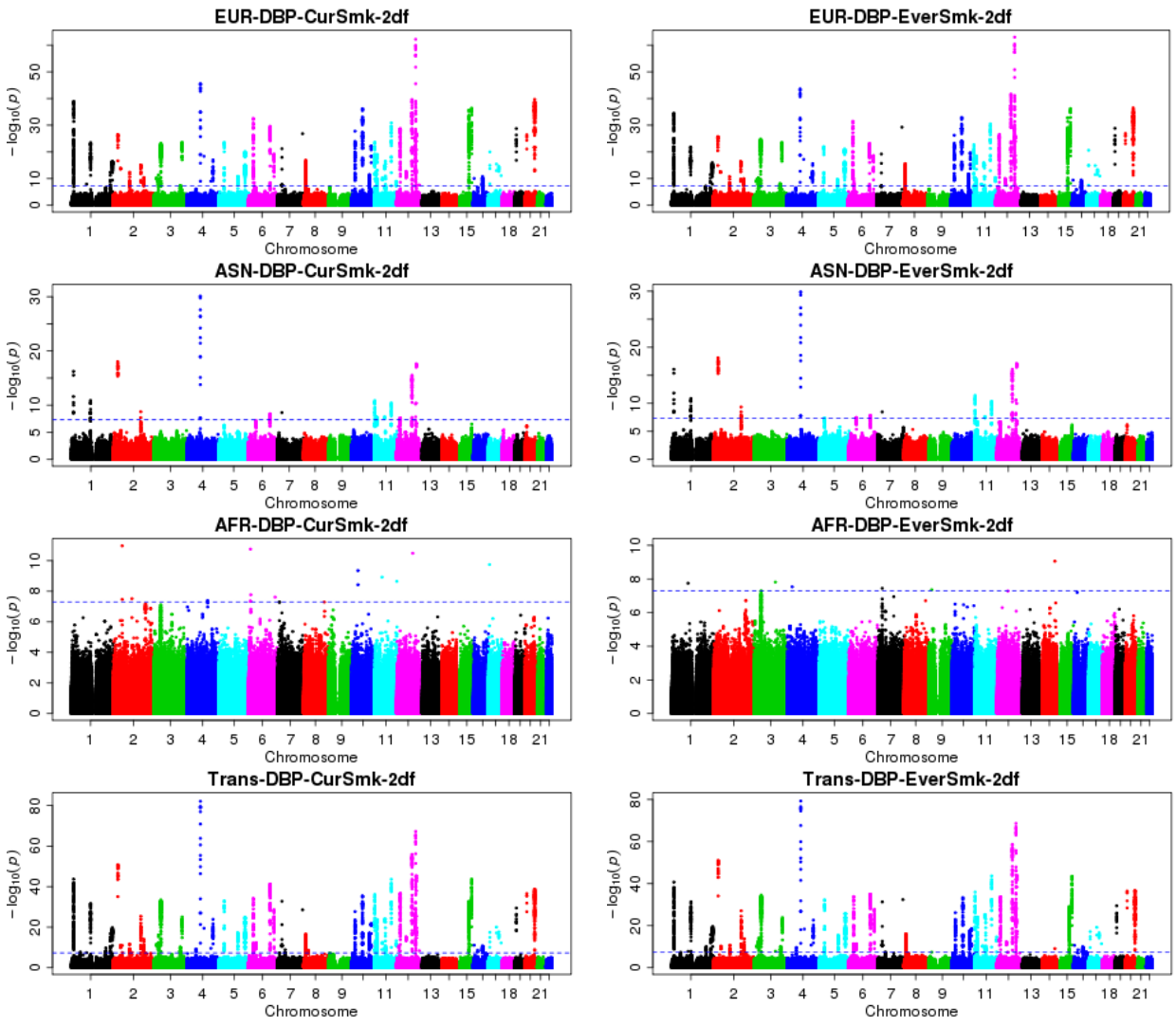


Figure S6: Manhattan plots of DBP using the 2 DF joint test.

The $-\log_{10}(p)$ of each SNP was plotted at the chromosomal location of each variant. The p-values are based on the combined discovery and replication analysis for the select 4,459 variants and the discovery analysis for the remaining 18.8 million variants.

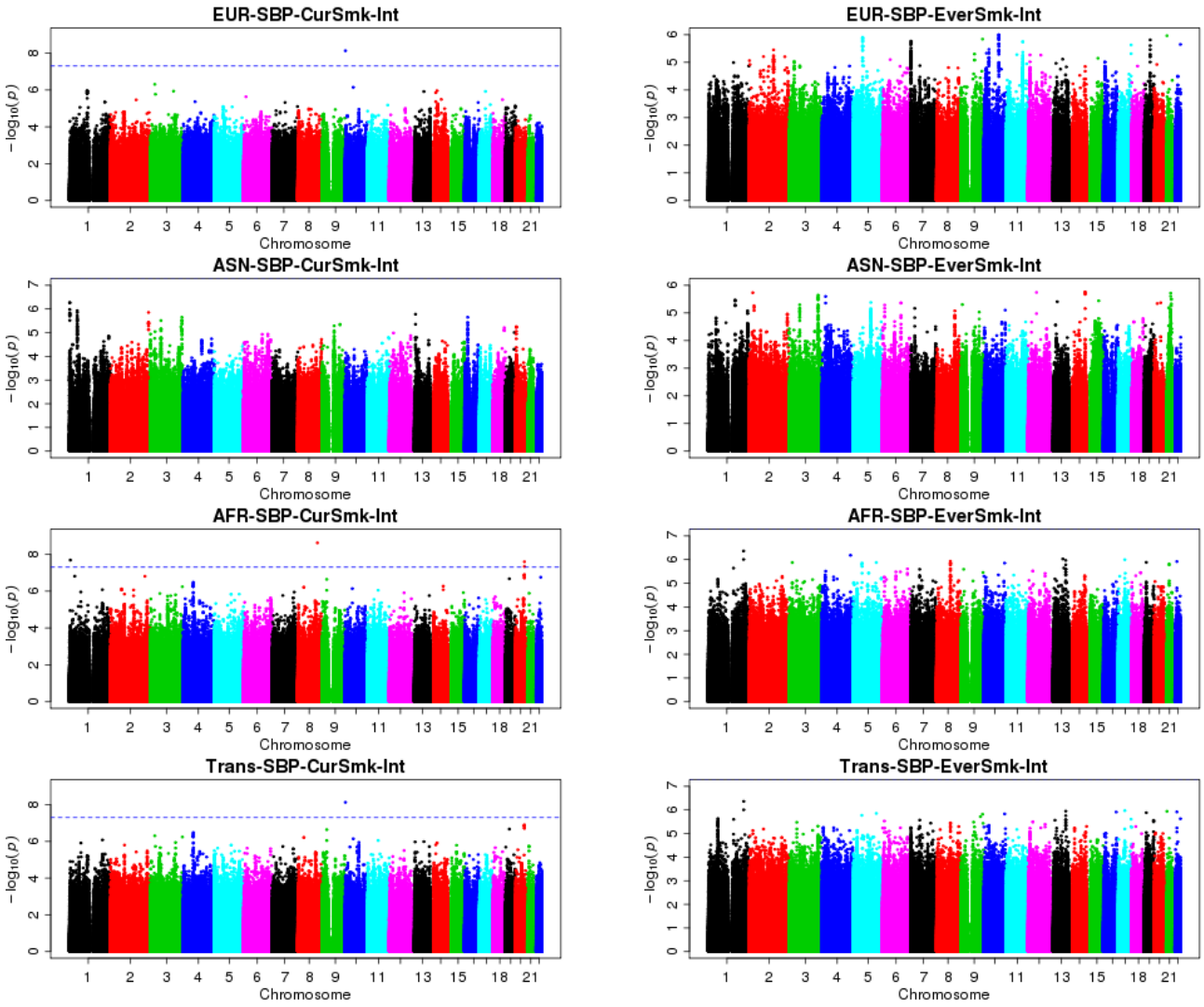


Figure S7: Manhattan plots of SBP using the 1 DF interaction test.

The $-\log_{10}(p)$ of each SNP was plotted at the chromosomal location of each variant. The p-values are based on the combined discovery and replication analysis for the select 4,459 variants and the discovery analysis for the remaining 18.8 million variants.

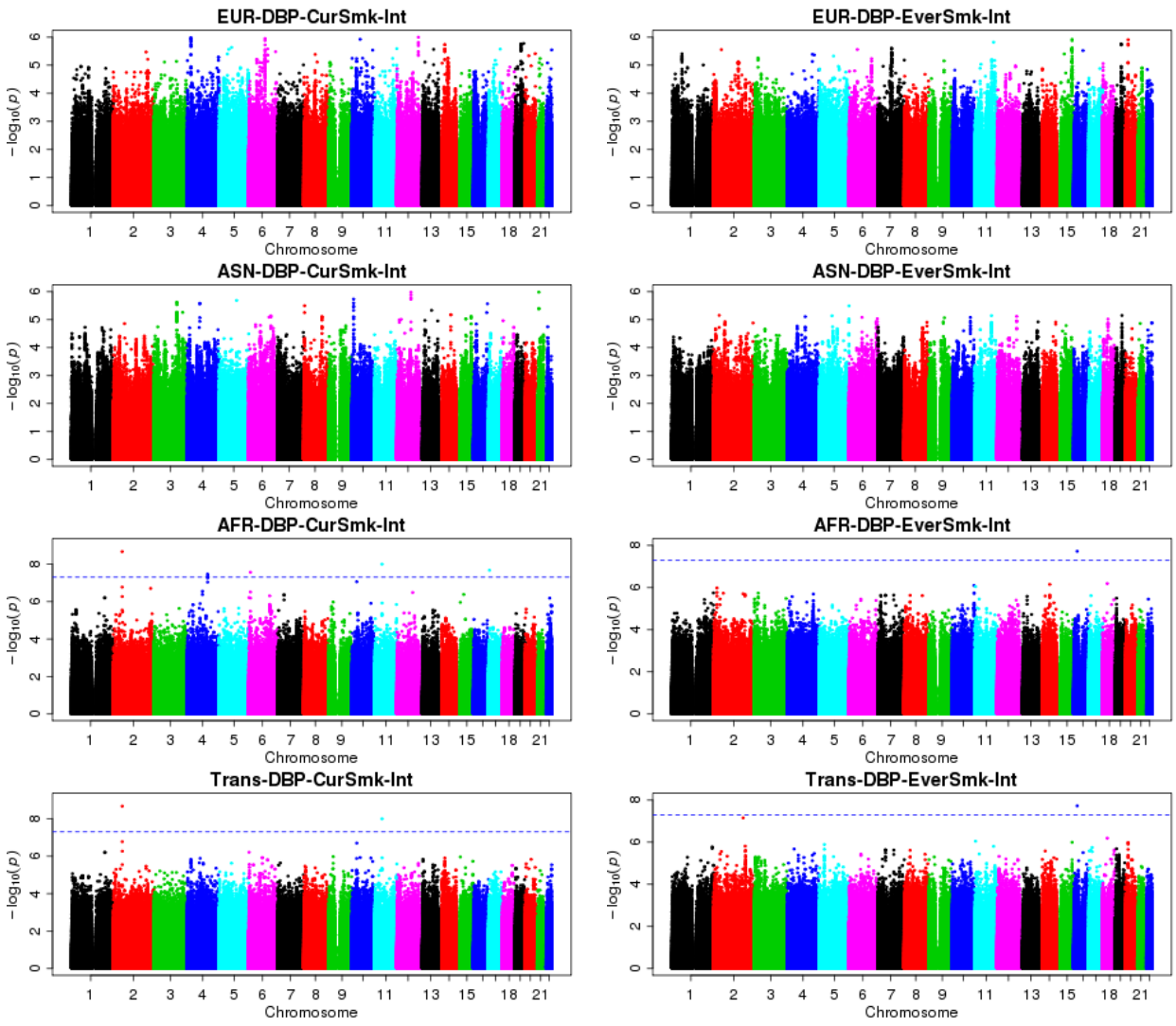
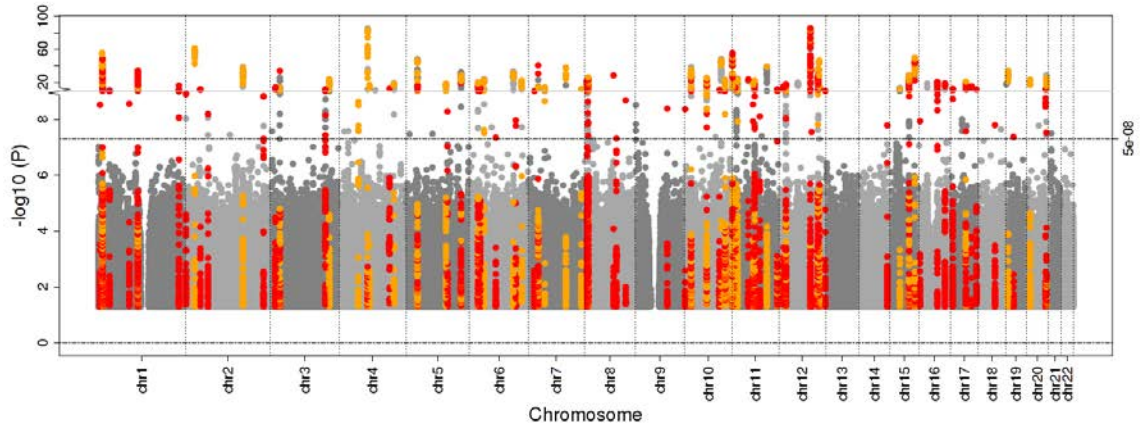


Figure S8: Manhattan plots of DBP using the 1 DF interaction test.

The $-\log_{10}(p)$ of each SNP was plotted at the chromosomal location of each variant. The p-values are based on the combined discovery and replication analysis for the select 4,459 variants and the discovery analysis for the remaining 18.8 million variants.

A SBP



B DBP

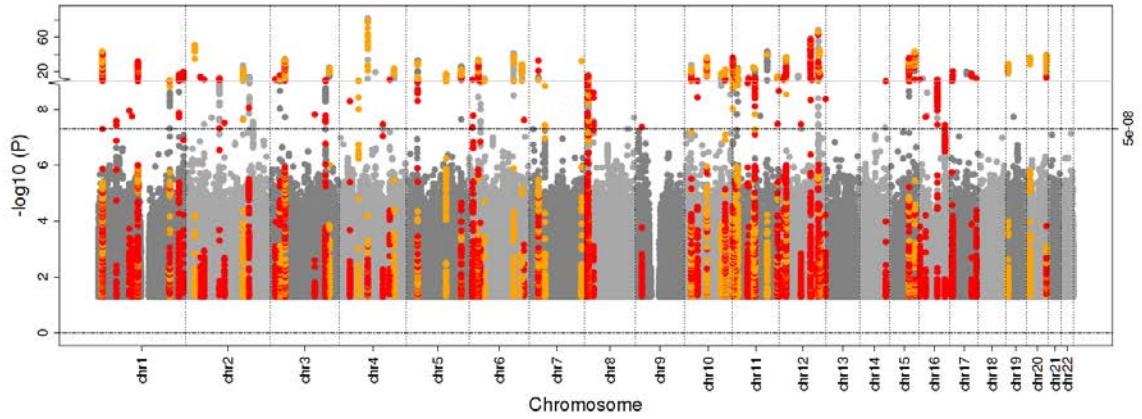


Figure S9. Manhattan plots for SBP and DBP using the 2 DF joint test.

The orange points correspond to known BP loci that were identified, and red points correspond to the newly identified BP loci. The result is based on the combined analysis of the genome-wide discovery analysis in Stage 1 cohorts (18.8 million variants) and focused/replication analysis in the Stage 2 cohorts for the selected 4,459 variants. The $-\log_{10}(p)$ of each SNP was plotted at the chromosomal location of each variant. The minimum p-values across smoking exposures, across tests, and across ancestry-specific and trans-ancestry results were used. **Figures S5-S8** show Manhattan plots separately for each smoking exposure and for three ancestry-specific results and trans-ancestry results.

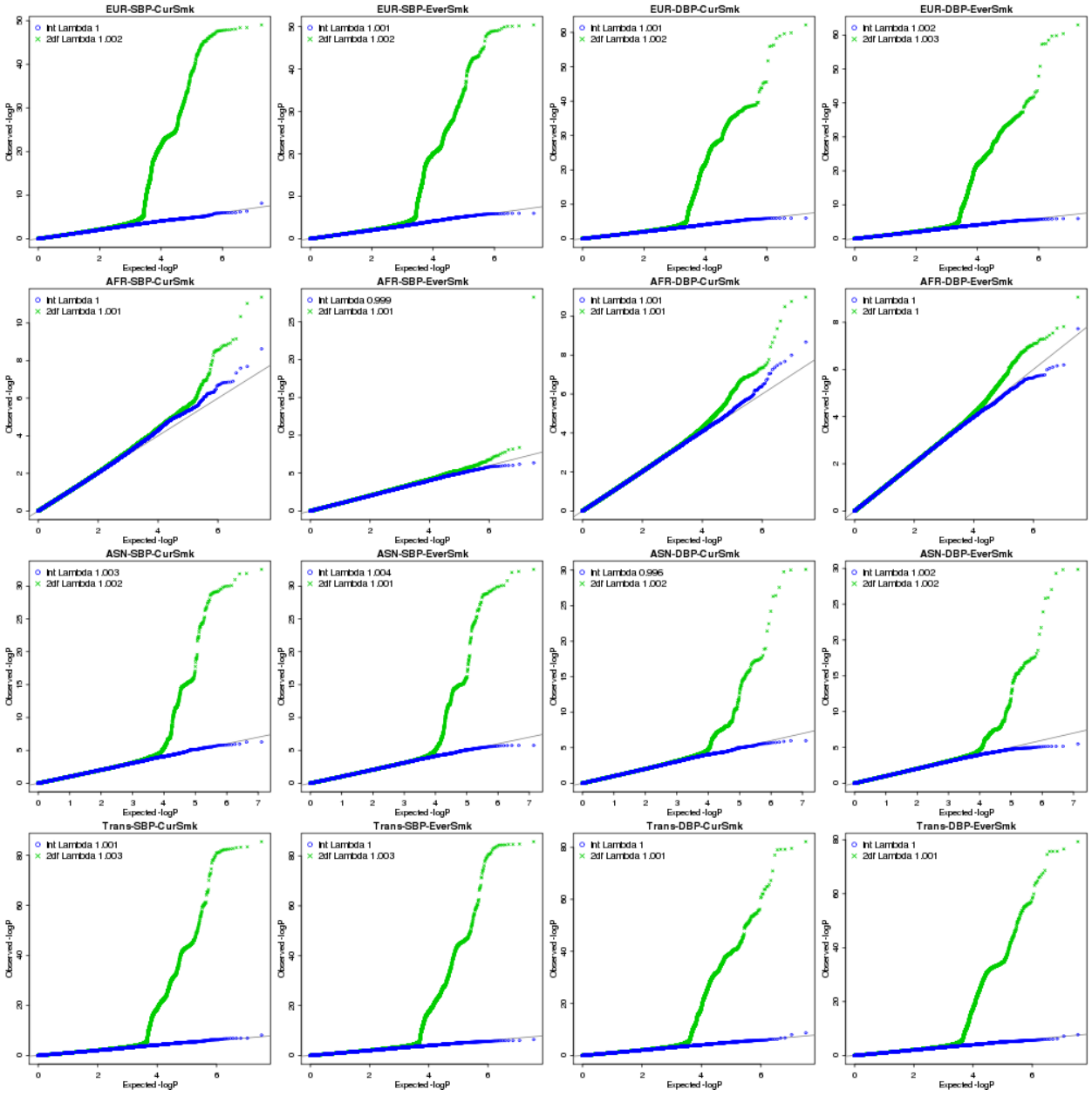


Figure S10: QQ plots of the combined (Stages 1 and 2) meta-analyses.

The combination of BP traits and smoking exposures were used: SBP-CurSmk (1st column), SBP-EverSmk (2nd column), DBP-CurSmk (3rd column), and DBP-EverSmk (4th column). Each plot displays p-values (blue circles for the 1 DF test of interaction effect; green crosses for the 2 DF joint test) and their genomic inflation factor. The p-values are based on the meta-analysis result to combine results from the genome-wide discovery analysis in Stage 1 cohorts (18.8 million variants) and focused/replication analysis in the Stage 2 cohorts for the selected 4,459 variants.

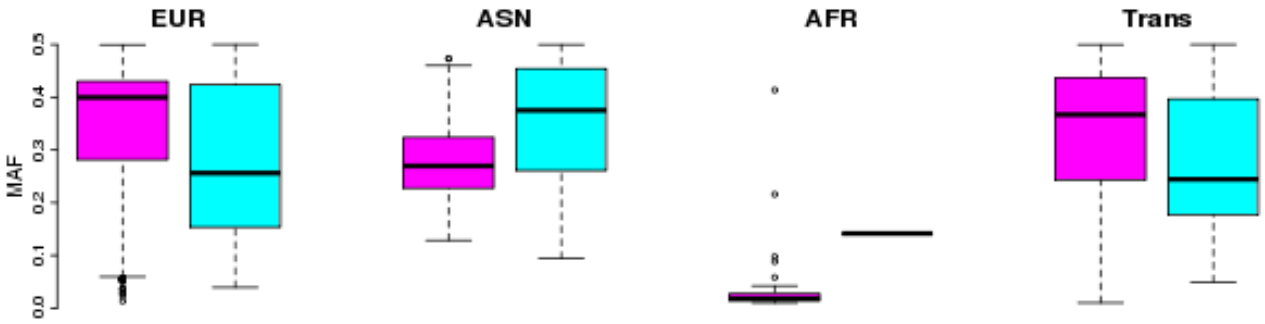


Figure S11: MAF distribution at genome-wide significant variants.

The magenta box is for variants at novel loci and the cyan box is for variants at known loci. There were only two variants at known loci (therefore, no cyan box) in AFR. MAF: minor allele frequency

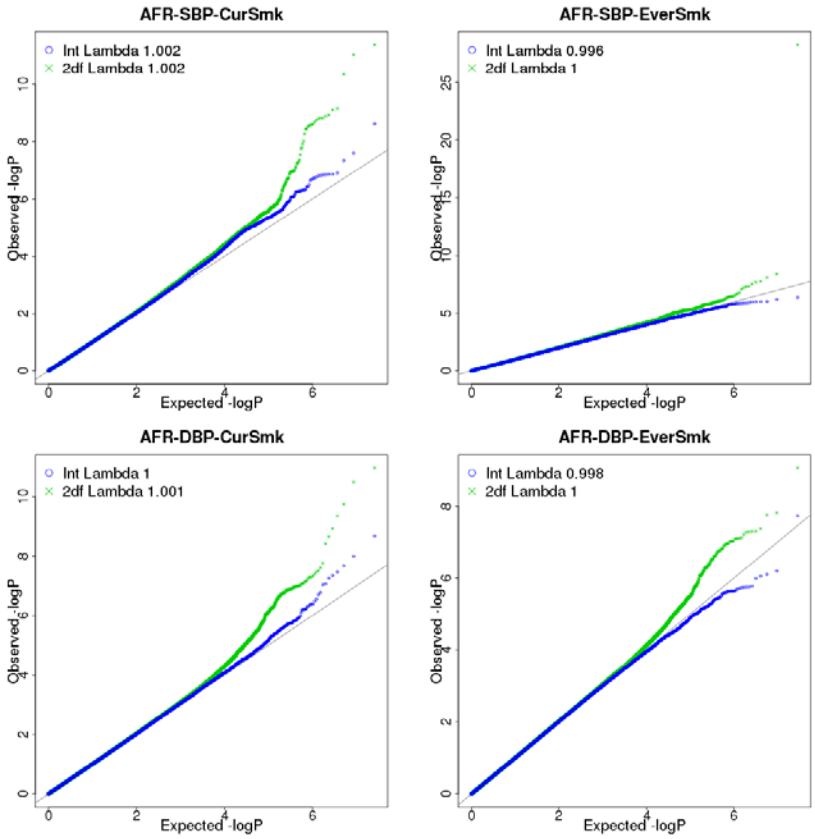
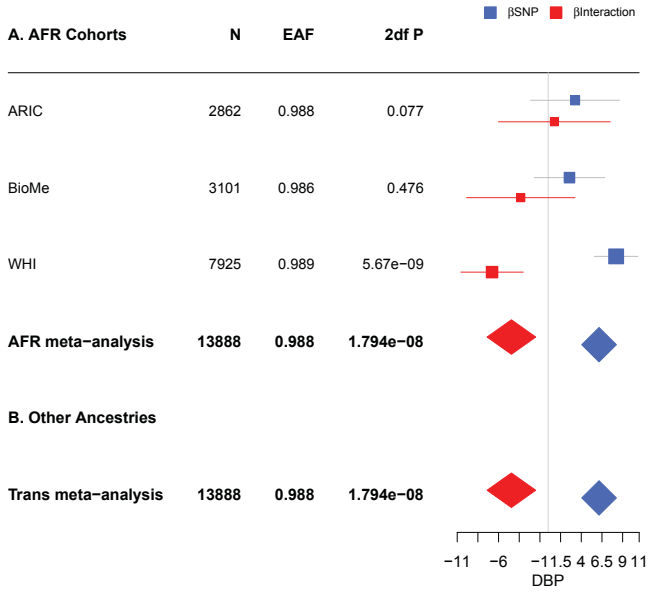


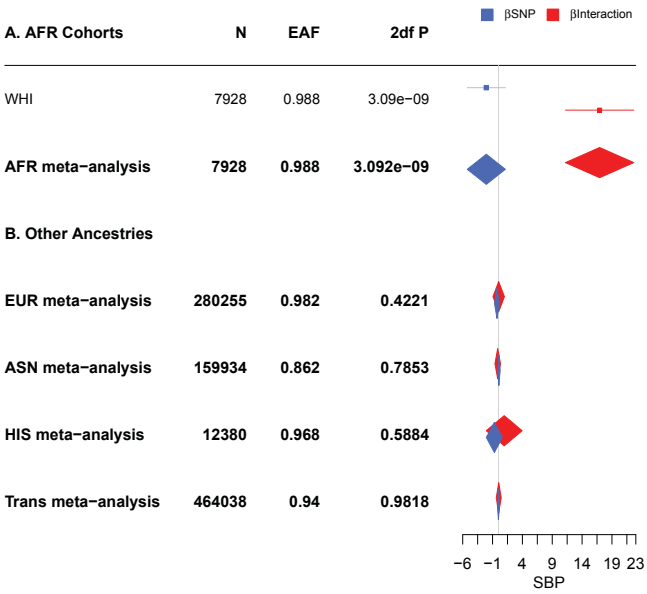
Figure S12: QQ plots of the combined (Stages 1 and 2) meta-analyses without known BP loci.

Each plot displays p-values (blue circles for the 1 DF test of interaction effect; green crosses for the 2 DF joint test) and their genomic inflation factor. The p-values are based on the meta-analysis result to combine results from the genome-wide discovery analysis in Stage 1 cohorts (18.8 million variants) and focused/replication analysis in the Stage 2 cohorts for the selected 4,459 variants. The variants within 1Mb around the known BP loci are excluded.

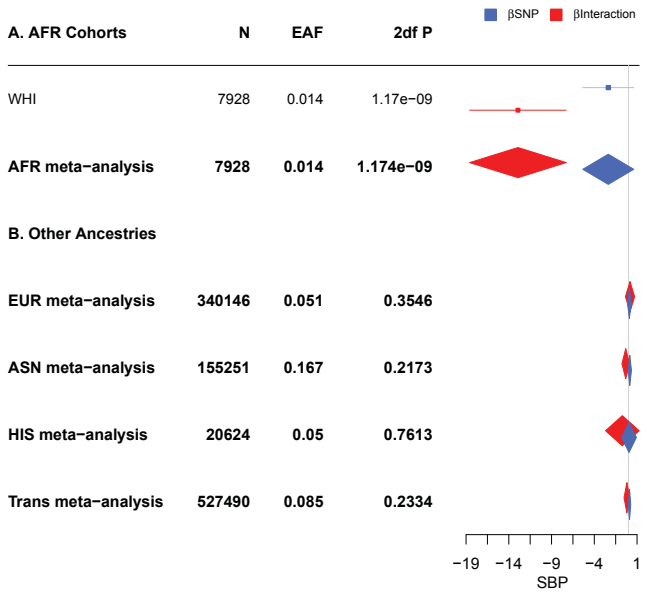
Effect of rs182662555 (T5−L3*) and its interaction with EverSmk on DBP



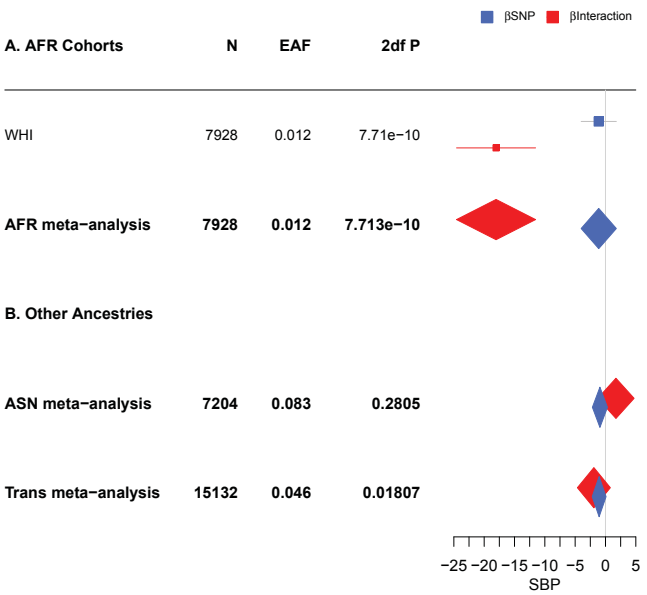
Effect of rs12135881 (T5−L1*) and its interaction with CurSmk on SBP



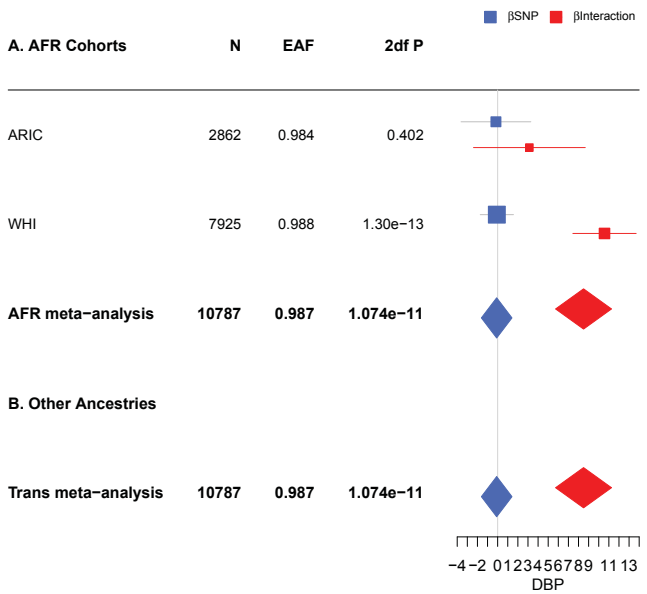
Effect of rs75247762 (T5−L4*) and its interaction with CurSmk on SBP



Effect of rs11809589 (T5−L2*) and its interaction with CurSmk on SBP

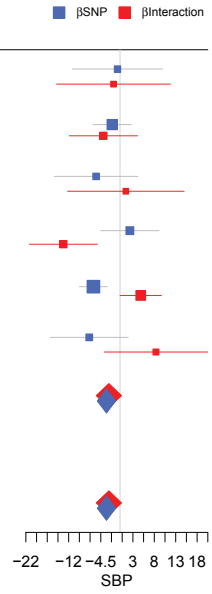


Effect of rs115234772 (T5−L5*) and its interaction with CurSmk on DBP



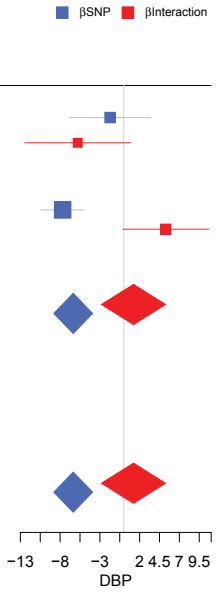
Effect of rs145162854 (T5−L6*) and its interaction with EverSmk on SBP

A. AFR Cohorts	N	EAF	2df P
ARIC	2862	0.014	0.871
JHS	2134	0.018	0.178
MESA	1594	0.018	0.362
BioMe	3101	0.012	6.25e−07
WHI	7928	0.015	7.31e−04
HRS	1993	0.015	0.29
AFR meta−analysis	**20852**	**0.015**	**6.627e−09**
B. Other Ancestries			
Trans meta−analysis	**20852**	**0.015**	**6.627e−09**



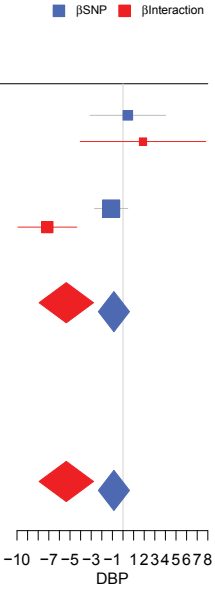
Effect of rs139963642 (T5−L9*) and its interaction with EverSmk on DBP

A. AFR Cohorts	N	EAF	2df P
ARIC	2862	0.012	0.002
BioMe	3101	0.014	1.62e−07
AFR meta−analysis	**5963**	**0.013**	**1.546e−08**
B. Other Ancestries			
Trans meta−analysis	**5963**	**0.013**	**1.546e−08**



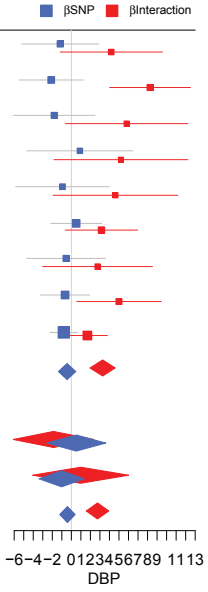
Effect of rs116008367 (T5−L7*) and its interaction with CurSmk on DBP

A. AFR Cohorts	N	EAF	2df P
ARIC	2862	0.015	0.604
WHI	7925	0.014	6.68e−12
AFR meta−analysis	**10787**	**0.014**	**3.09e−08**
B. Other Ancestries			
Trans meta−analysis	**10787**	**0.014**	**3.09e−08**



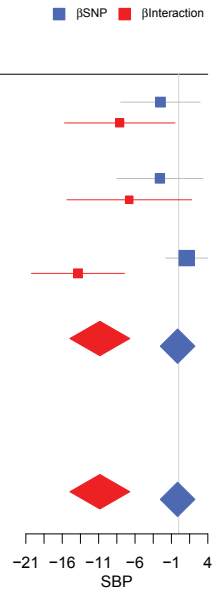
Effect of rs11931572 (T5−L10*) and its interaction with EverSmk on DBP

A. AFR Cohorts	N	EAF	2df P
ARIC	2862	0.979	0.203
FamHS	617	0.947	3.45e−05
GENOA	941	0.967	0.193
HABC	1136	0.967	0.02
HUFS	1686	0.97	0.159
JHS	2134	0.961	0.031
MESA	1594	0.968	0.542
BioMe	3101	0.971	0.054
WHI	7925	0.968	0.294
AFR meta−analysis	**23236**	**0.968**	**2.905e−08**
B. Other Ancestries			
EUR meta−analysis	**3133**	**0.987**	**0.6051**
ASN meta−analysis	**2466**	**0.962**	**0.6416**
Trans meta−analysis	**28835**	**0.969**	**9.235e−07**



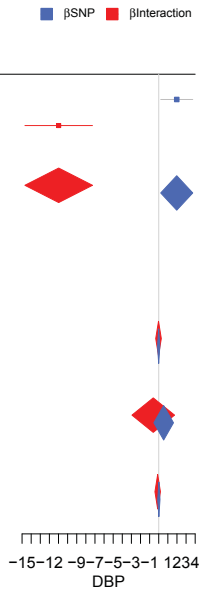
Effect of rs10166552 (T5−L8*) and its interaction with CurSmk on SBP

A. AFR Cohorts	N	EAF	2df P
ARIC	2862	0.016	2.71e−04
HUFS	1686	0.019	0.014
WHI	7928	0.014	5.18e−05
AFR meta−analysis	**12476**	**0.015**	**1.523e−09**
B. Other Ancestries			
Trans meta−analysis	**12476**	**0.015**	**1.523e−09**



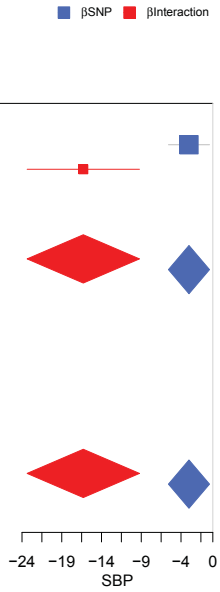
Effect of rs62319742 (T5−L11*) and its interaction with CurSmk on DBP

A. AFR Cohorts	N	EAF	2df P
WHI	7925	0.014	4.09e−08
AFR meta−analysis	**7925**	**0.014**	**4.087e−08**
B. Other Ancestries			
EUR meta−analysis	**355253**	**0.055**	**0.9773**
HIS meta−analysis	**12380**	**0.022**	**0.614**
Trans meta−analysis	**379099**	**0.053**	**0.7524**



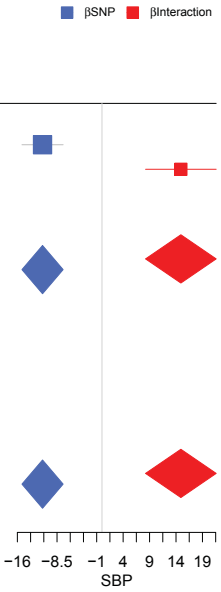
Effect of rs140543491 (T5−L12*) and its interaction with CurSmk on SBP

A. AFR Cohorts	N	EAF	2df P
WHI	7928	0.017	5.34e−09
AFR meta−analysis	**7928**	**0.017**	**5.341e−09**
B. Other Ancestries			
Trans meta−analysis	**7928**	**0.017**	**5.341e−09**



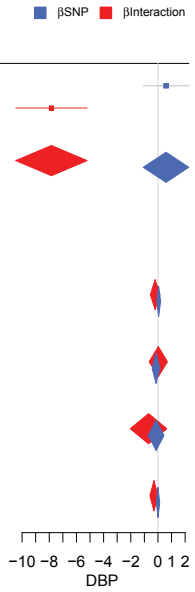
Effect of rs58806982 (T5−L15*) and its interaction with EverSmk on SBP

A. AFR Cohorts	N	EAF	2df P
WHI	7928	0.01	4.57e−08
AFR meta−analysis	**7928**	**0.01**	**4.569e−08**
B. Other Ancestries			
Trans meta−analysis	**7928**	**0.01**	**4.569e−08**



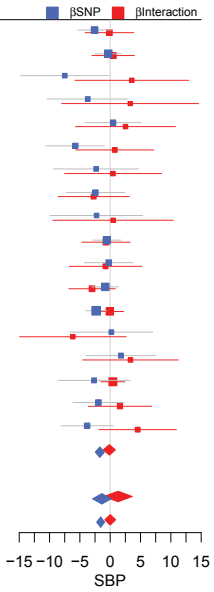
Effect of rs148387718 (T5−L13*) and its interaction with CurSmk on DBP

A. AFR Cohorts	N	EAF	2df P
WHI	7925	0.014	1.77e−11
AFR meta−analysis	**7925**	**0.014**	**1.774e−11**
B. Other Ancestries			
EUR meta−analysis	**351990**	**0.042**	**0.4937**
ASN meta−analysis	**131780**	**0.031**	**0.5709**
HIS meta−analysis	**20624**	**0.045**	**0.3746**
Trans meta−analysis	**515860**	**0.039**	**0.1166**



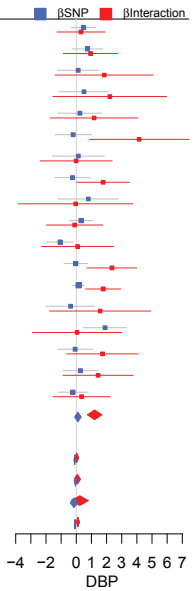
Effect of rs76987554 (T5−L16*) and its interaction with EverSmk on SBP

A. AFR Cohorts	N	EAF	2df P
ARIC	2862	0.099	0.038
CARDIA	945	0.1	0.953
CHS	734	0.077	0.058
FamHS	617	0.08	0.555
GENOA	941	0.094	0.645
HABC	1136	0.082	0.005
HANDLS	903	0.083	0.506
HUFS	1686	0.085	0.005
HyperGEN	418	0.11	0.73
JHS	2134	0.085	0.676
MESA	1594	0.084	0.899
BioMe	3101	0.108	0.047
WHI	7928	0.083	3.67e−04
AADHS	584	0.081	0.113
CFS	562	0.082	0.144
GeneSTAR	1107	0.066	0.552
HRS	1993	0.102	0.65
HyperGEN−AXIOM	1240	0.075	0.207
AFR meta−analysis	**30485**	**0.088**	**2.957e−08**
B. Other Ancestries			
HIS meta−analysis	**21122**	**0.024**	**0.2136**
Trans meta−analysis	**55148**	**0.059**	**1.632e−08**



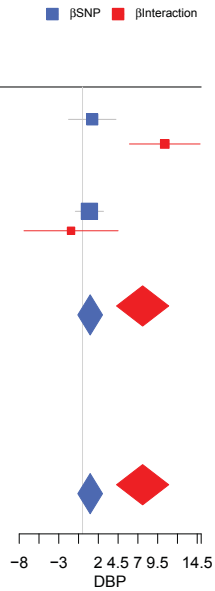
Effect of rs9348895 (T5−L14*) and its interaction with CurSmk on DBP

A. AFR Cohorts	N	EAF	2df P
ARIC	2862	0.587	0.27
CARDIA	945	0.611	0.03
CHS	734	0.579	0.404
FamHS	617	0.612	0.274
GENOA	941	0.585	0.503
HABC	1136	0.583	0.04
HANDLS	903	0.59	0.977
HUFS	1686	0.584	0.068
HyperGEN	418	0.586	0.666
JHS	2134	0.591	0.705
MESA	1594	0.588	0.044
BioMe	3101	0.588	0.008
WHI	7925	0.58	0.002
AADHS	584	0.569	0.651
CFS	561	0.581	0.01
GeneSTAR	1107	0.574	0.254
HyperGEN−AXIOM	1240	0.614	0.233
JUPITER	1606	0.58	0.889
AFR meta−analysis	**30094**	**0.586**	**1.706e−08**
B. Other Ancestries			
EUR meta−analysis	**361601**	**0.419**	**0.0008746**
ASN meta−analysis	**162367**	**0.515**	**0.4906**
HIS meta−analysis	**22079**	**0.377**	**0.4352**
Trans meta−analysis	**580555**	**0.453**	**0.001081**



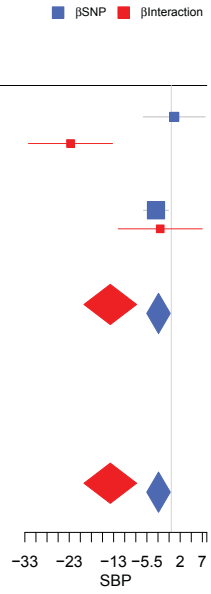
Effect of rs112140754 (T5−L17*) and its interaction with CurSmk on DBP

A. AFR Cohorts	N	EAF	2df P
ARIC	2862	0.987	2.95e−11
WHI	7925	0.988	0.617
AFR meta−analysis	**10787**	**0.988**	**2.436e−08**
B. Other Ancestries			
Trans meta−analysis	**10787**	**0.988**	**2.436e−08**



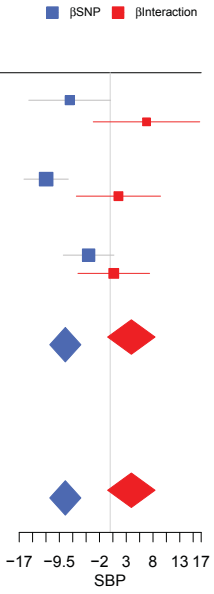
Effect of rs116196735 (T5−L18*) and its interaction with CurSmk on SBP

A. AFR Cohorts	N	EAF	2df P
ARIC	2862	0.016	1.25e−10
WHI	7928	0.018	0.025
AFR meta−analysis	**10790**	**0.018**	**6.982e−10**
B. Other Ancestries			
Trans meta−analysis	**10790**	**0.018**	**6.982e−10**



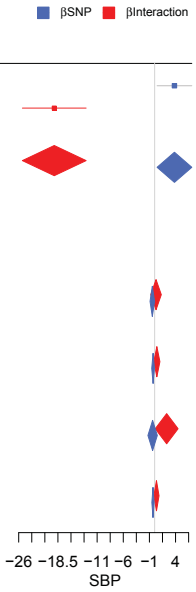
Effect of rs192642798 (T5−L21*) and its interaction with EverSmk on SBP

A. AFR Cohorts	N	EAF	2df P
ARIC	2862	0.012	0.148
BioMe	3101	0.014	9.95e−10
WHI	7928	0.011	0.095
AFR meta−analysis	**13891**	**0.012**	**4.227e−09**
B. Other Ancestries			
Trans meta−analysis	**13891**	**0.012**	**4.227e−09**



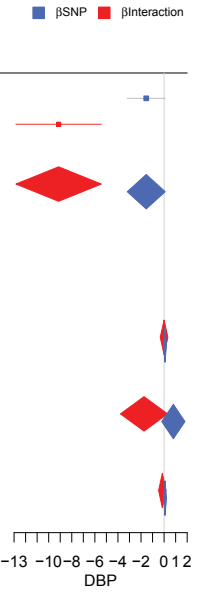
Effect of rs74701635 (T5−L19*) and its interaction with CurSmk on SBP

A. AFR Cohorts	N	EAF	2df P
WHI	7928	0.016	2.13e−09
AFR meta−analysis	**7928**	**0.016**	**2.128e−09**
B. Other Ancestries			
EUR meta−analysis	**290825**	**0.014**	**0.1244**
ASN meta−analysis	**155029**	**0.14**	**0.06466**
HIS meta−analysis	**20624**	**0.044**	**0.1186**
Trans meta−analysis	**477947**	**0.057**	**0.01408**



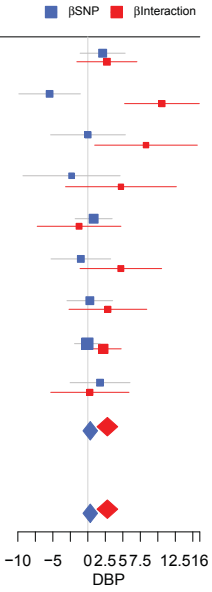
Effect of rs76726877 (T5−L22*) and its interaction with CurSmk on DBP

A. AFR Cohorts	N	EAF	2df P
WHI	7925	0.015	4.47e−10
AFR meta−analysis	**7925**	**0.016**	**4.465e−10**
B. Other Ancestries			
EUR meta−analysis	**347849**	**0.05**	**0.4585**
HIS meta−analysis	**12380**	**0.027**	**0.1699**
Trans meta−analysis	**371695**	**0.049**	**0.4461**



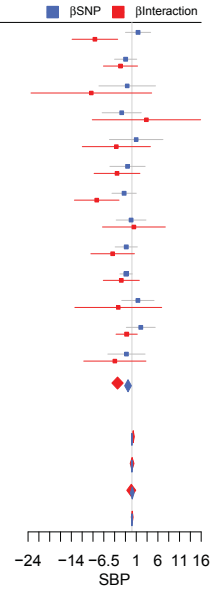
Effect of rs146250839 (T5−L20*) and its interaction with EverSmk on DBP

A. AFR Cohorts	N	EAF	2df P
ARIC	2862	0.973	0.002
GENOA	941	0.967	3.63e−04
HABC	1136	0.978	0.006
HUFS	1686	0.976	0.44
JHS	2134	0.974	0.818
MESA	1594	0.979	0.169
BioMe	3101	0.982	0.383
WHI	7925	0.978	0.054
HRS	1993	0.981	0.396
AFR meta−analysis	**24612**	**0.977**	**4.361e−08**
B. Other Ancestries			
Trans meta−analysis	**24612**	**0.977**	**4.361e−08**

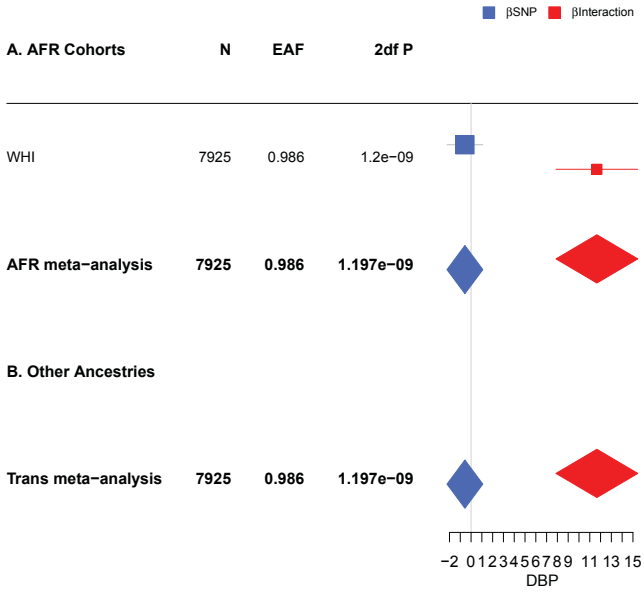


Effect of rs11599481 (T5−L23*) and its interaction with CurSmk on SBP

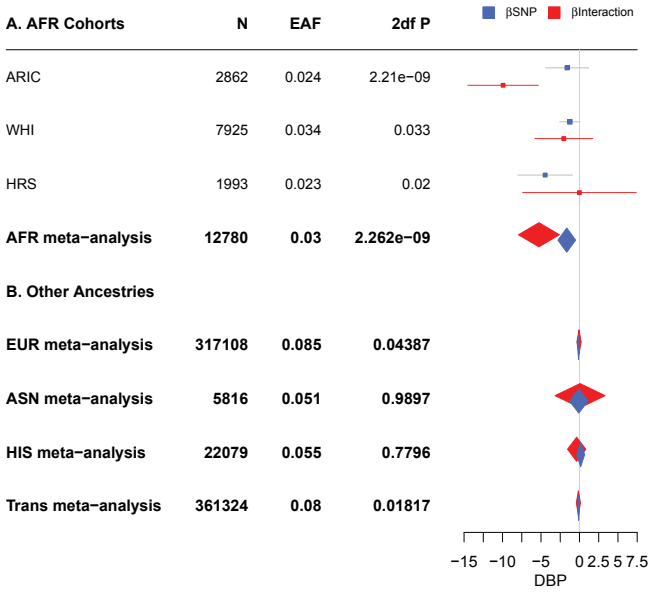
A. AFR Cohorts	N	EAF	2df P
ARIC	2862	0.053	0.004
CARDIA	945	0.048	0.015
FamHS	617	0.058	0.27
HABC	1136	0.065	0.582
HANDLS	903	0.055	0.563
HUFS	1686	0.054	0.042
JHS	2134	0.053	2.62e−05
MESA	1594	0.067	0.991
BioMe	3101	0.055	0.021
WHI	7928	0.066	0.032
CFS	561	0.059	0.724
GeneSTAR	1107	0.053	0.234
HRS	1993	0.044	0.179
AFR meta−analysis	**27807**	**0.058**	**4.545e−11**
B. Other Ancestries			
EUR meta−analysis	**366964**	**0.248**	**0.05075**
ASN meta−analysis	**160833**	**0.163**	**0.9767**
HIS meta−analysis	**22079**	**0.229**	**0.937**
Trans meta−analysis	**582097**	**0.215**	**0.7726**



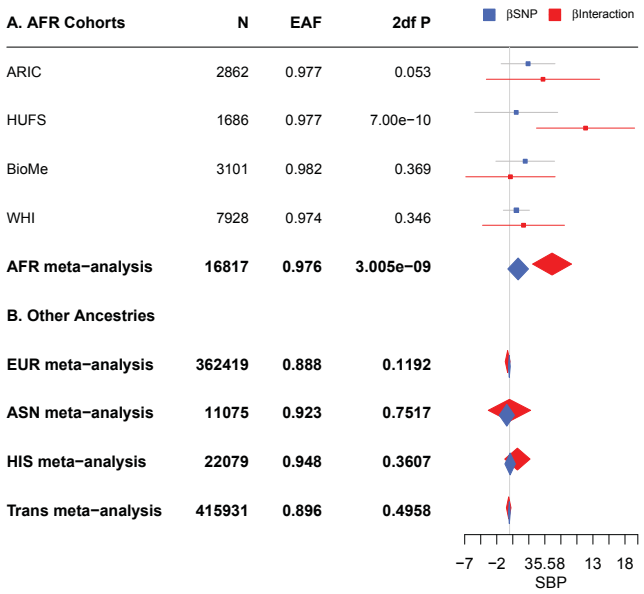
Effect of rs148772934 (T5−L24*) and its interaction with CurSmk on DBP



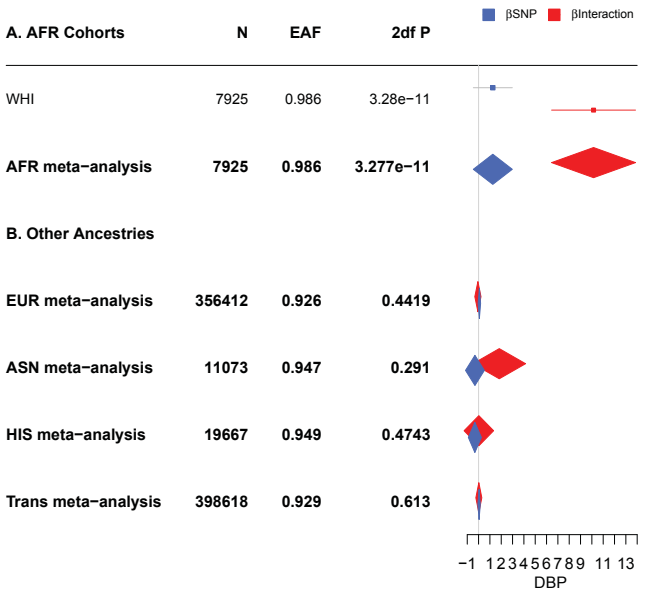
Effect of rs78103586 (T5−L27*) and its interaction with CurSmk on DBP



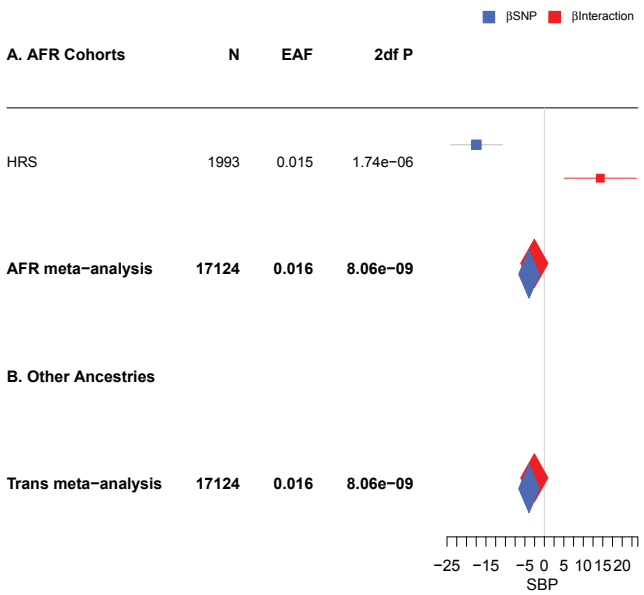
Effect of rs11601370 (T5−L25) and its interaction with CurSmk on SBP



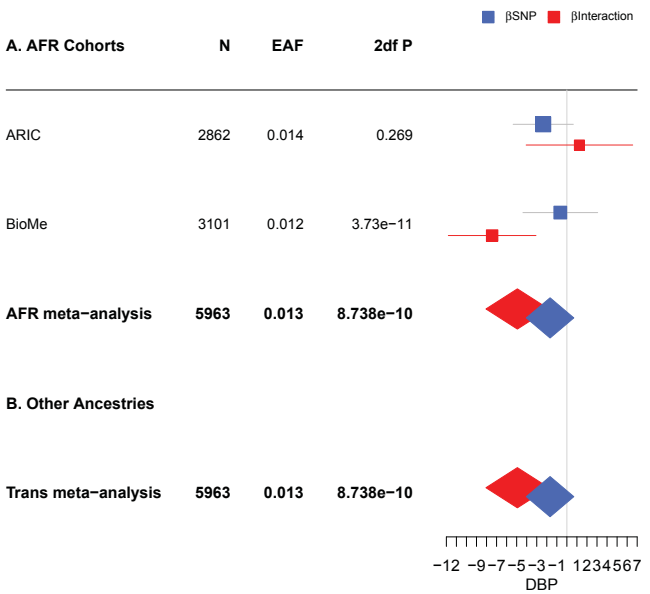
Effect of rs61935525 (T5−L28*) and its interaction with CurSmk on DBP



Effect of rs74601585 (T5−L26*) and its interaction with EverSmk on SBP

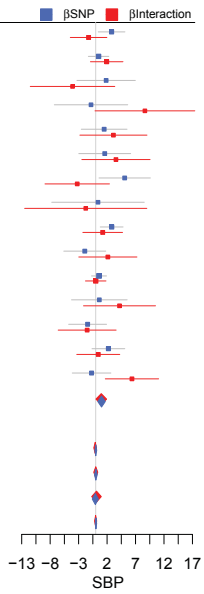


Effect of rs187852559 (T5−L29*) and its interaction with EverSmk on DBP



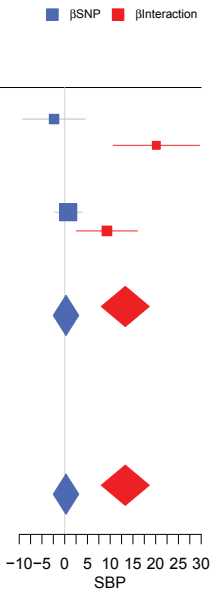
Effect of rs1257310 (T5−L30*) and its interaction with EverSmk on SBP

A. AFR Cohorts	N	EAF	2df P
ARIC	2862	0.793	0.025
CARDIA	945	0.792	0.085
CHS	734	0.809	0.561
FamHS	617	0.818	0.048
GENOA	941	0.788	0.102
HABC	1136	0.8	0.028
HANDLS	903	0.776	0.049
HyperGEN	418	0.809	0.926
JHS	2134	0.791	5.05e−04
MESA	1594	0.802	0.59
WHI	7928	0.787	0.442
AADHS	584	0.807	0.061
GeneSTAR	1107	0.493	0.704
HRS	1993	0.818	0.032
HyperGEN−AXIOM	1240	0.793	0.006
AFR meta−analysis	**25136**	**0.781**	**1.669e−08**
B. Other Ancestries			
EUR meta−analysis	**346712**	**0.809**	**0.6871**
ASN meta−analysis	**154002**	**0.707**	**0.9341**
HIS meta−analysis	**22271**	**0.777**	**0.9465**
Trans meta−analysis	**552535**	**0.778**	**0.7131**



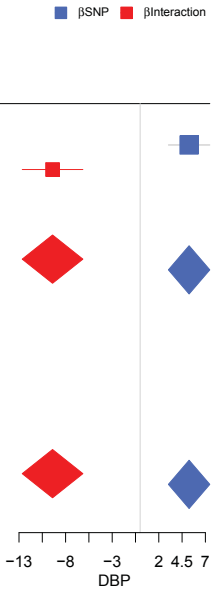
Effect of rs9965695 (T5−L33*) and its interaction with CurSmk on SBP

A. AFR Cohorts	N	EAF	2df P
ARIC	2862	0.984	3.74e−07
WHI	7928	0.982	0.004
AFR meta−analysis	**10790**	**0.983**	**1.627e−08**
B. Other Ancestries			
Trans meta−analysis	**10790**	**0.983**	**1.627e−08**



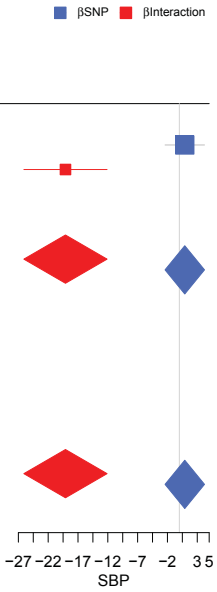
Effect of rs148753653 (T5−L31) and its interaction with EverSmk on DBP

A. AFR Cohorts	N	EAF	2df P
WHI	7925	0.989	6.3e−08
AFR meta−analysis	**7925**	**0.989**	**6.298e−08**
B. Other Ancestries			
Trans meta−analysis	**7925**	**0.989**	**6.298e−08**



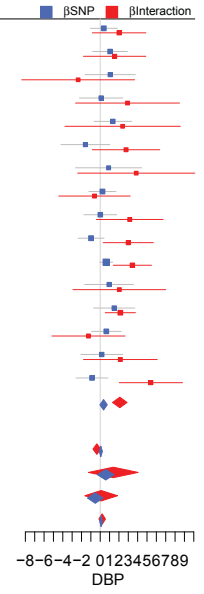
Effect of rs10405764 (T5−L34*) and its interaction with CurSmk on SBP

A. AFR Cohorts	N	EAF	2df P
WHI	7928	0.017	4.3e−08
AFR meta−analysis	**7928**	**0.017**	**4.302e−08**
B. Other Ancestries			
Trans meta−analysis	**7928**	**0.017**	**4.302e−08**



Effect of rs138973557 (T5−L32*) and its interaction with CurSmk on DBP

A. AFR Cohorts	N	EAF	2df P
ARIC	2862	0.9	0.142
CARDIA	945	0.891	0.094
FamHS	617	0.898	0.666
GENOA	941	0.901	0.499
HABC	1136	0.903	0.187
HANDLS	903	0.898	0.321
HyperGEN	418	0.885	0.169
JHS	2134	0.91	0.929
MESA	1594	0.906	0.14
BioMe	3101	0.895	0.087
WHI	7925	0.905	3.73e−05
AADHS	584	0.893	0.296
GeneSTAR	1107	0.912	0.025
HRS	1993	0.899	0.687
HyperGEN−AXIOM	1240	0.911	0.422
JUPITER	1606	0.906	0.006
AFR meta−analysis	**29106**	**0.903**	**1.805e−10**
B. Other Ancestries			
EUR meta−analysis	**337182**	**0.968**	**0.2335**
ASN meta−analysis	**9014**	**0.958**	**0.1445**
HIS meta−analysis	**12380**	**0.961**	**0.3958**
Trans meta−analysis	**391223**	**0.963**	**0.09528**



Effect of rs115893283 (T5−L35*) and its interaction with CurSmk on SBP

A. AFR Cohorts	N	EAF	2df P
ARIC	2862	0.046	0.106
BioMe	3101	0.044	0.001
WHI	7928	0.039	0.677
JUPITER	1606	0.053	0.031
AFR meta−analysis	**16737**	**0.043**	**4.241e−12**
B. Other Ancestries			
EUR meta−analysis	**311102**	**0.056**	**0.1841**
ASN meta−analysis	**2468**	**0.056**	**0.1286**
HIS meta−analysis	**12380**	**0.032**	**0.04962**
Trans meta−analysis	**346228**	**0.054**	**0.629**

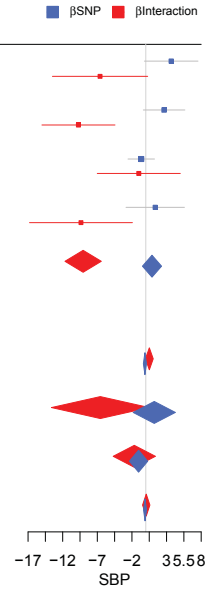
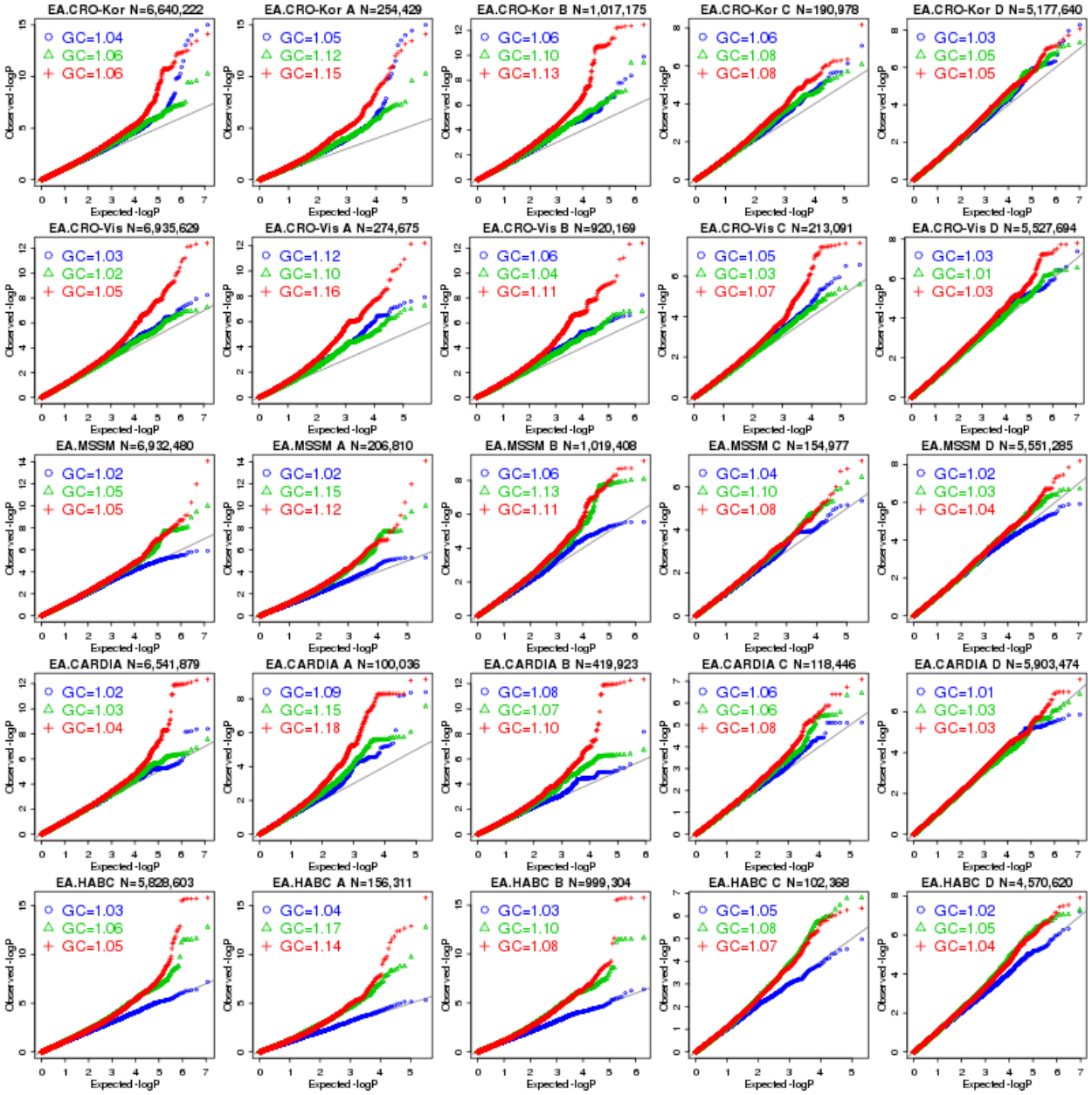
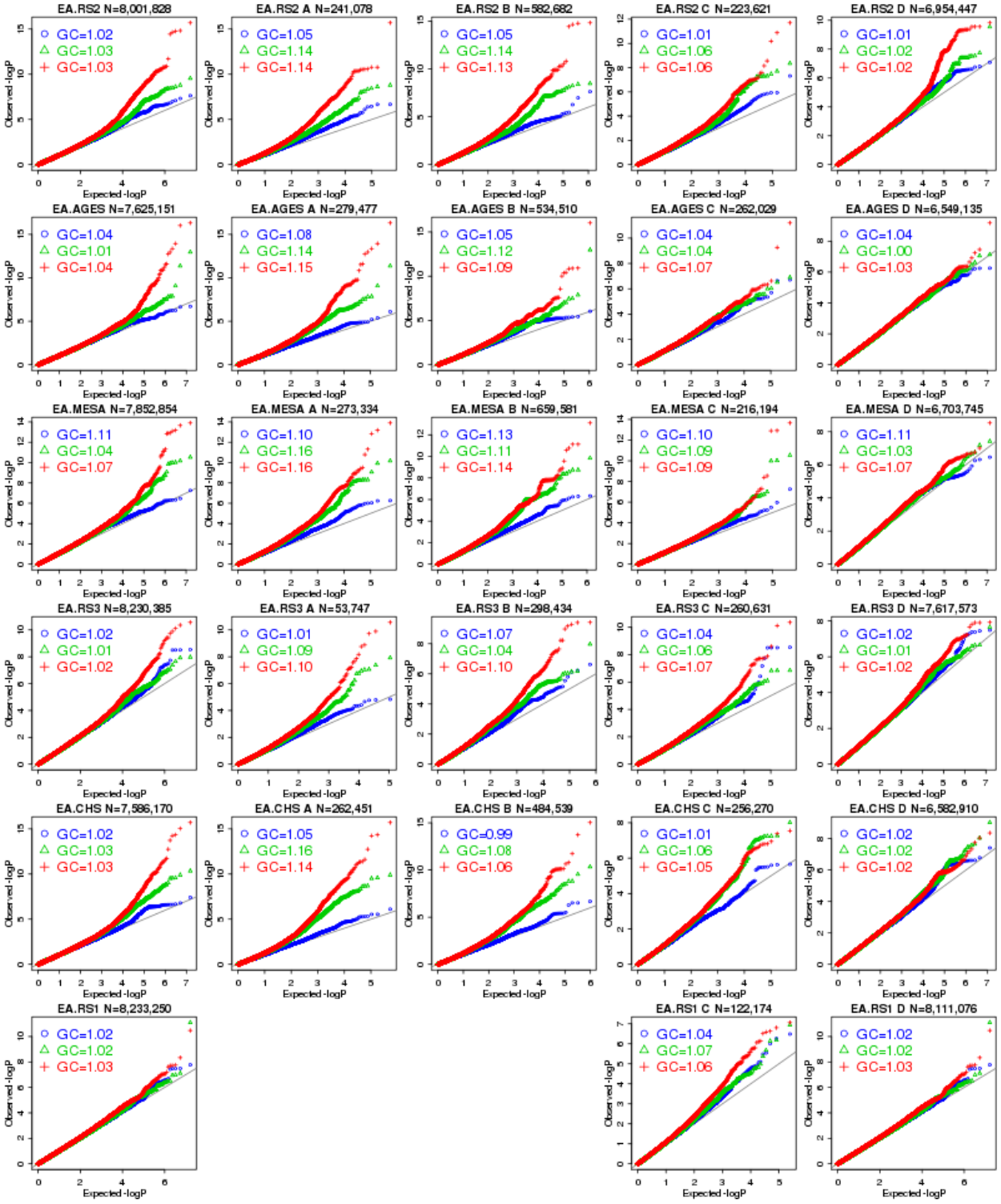
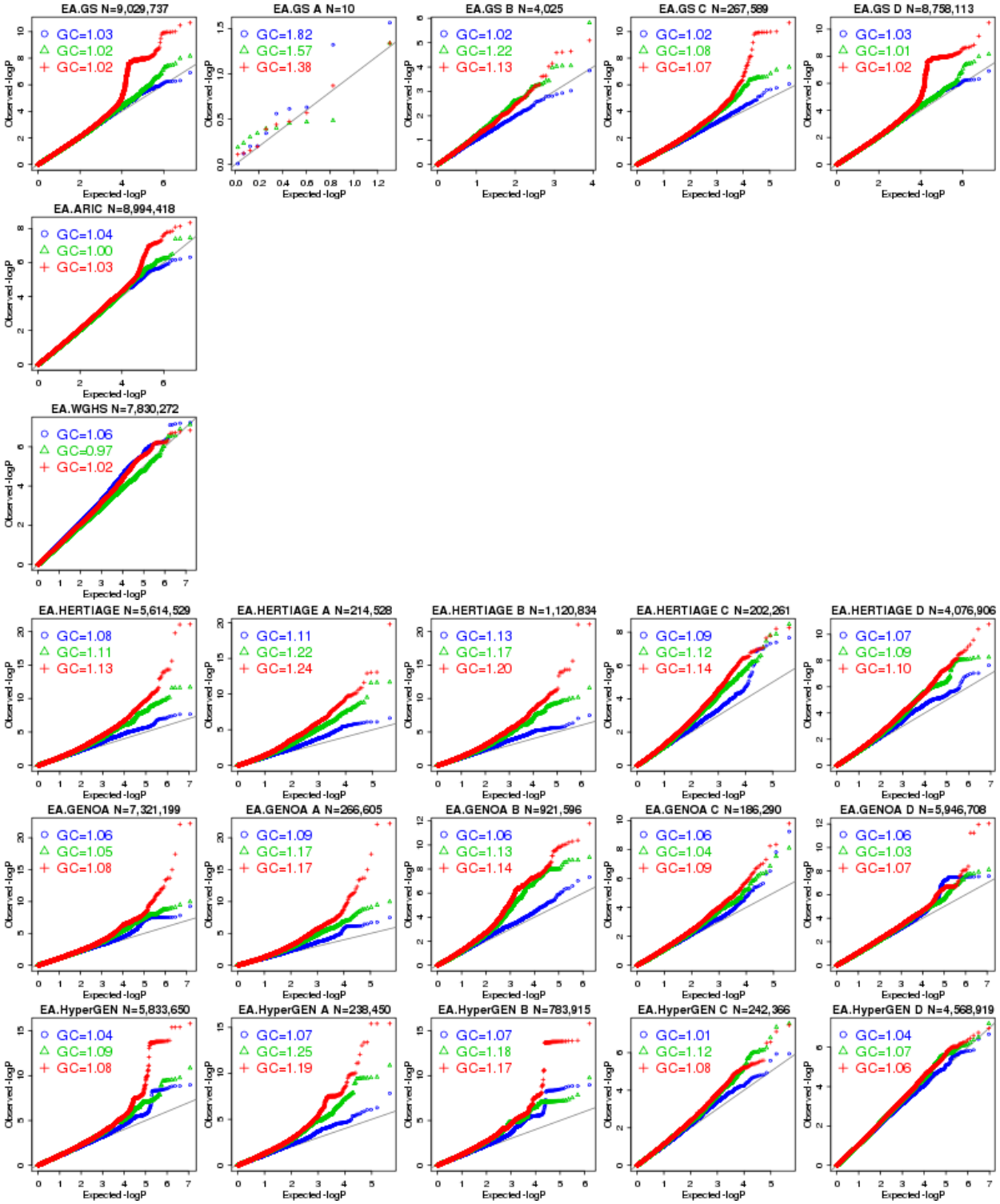


Figure S14: Cohort-specific QQ Plots in European ancestry

The first column in the QQ plots below is based on the first filter min(MAC0, MAC1) ≥ 10 (area A+B+C+D, see Supplemental Notes: More details on the Quality Control). The next 4 columns show QQ plots for the variants that fall into each section (A-D). Within each QQ plot, blue circles are based on the test of main effect, green triangles are based on the test of interaction effect, and the red crosses are based on the joint test of main and interaction effects.




EA.RS2 N=8,001,828
GC=1.02
GC=1.03
GC=1.03
EA.RS2 A N=241,078
GC=1.05
GC=1.14
GC=1.14
EA.RS2 B N=582,682
GC=1.05
GC=1.14
GC=1.13
EA.RS2 C N=223,621
GC=1.01
GC=1.06
GC=1.06
EA.RS2 D N=6,954,447
GC=1.01
GC=1.02
GC=1.02
EA.AGES N=7,625,151
GC=1.04
GC=1.01
GC=1.04
EA.AGES A N=279,477
GC=1.08
GC=1.14
GC=1.15
EA.AGES B N=534,510
GC=1.05
GC=1.12
GC=1.09
EA.AGES C N=262,029
GC=1.04
GC=1.04
GC=1.07
EA.AGES D N=6,549,135
GC=1.04
GC=1.00
GC=1.03
EA.MESA N=7,852,854
GC=1.11
GC=1.04
GC=1.07
EA.MESA A N=273,334
GC=1.10
GC=1.16
GC=1.16
EA.MESA B N=659,581
GC=1.13
GC=1.11
GC=1.14
EA.MESA C N=216,194
GC=1.10
GC=1.09
GC=1.09
EA.MESA D N=6,703,745
GC=1.11
GC=1.03
GC=1.07
EA.RS3 N=8,230,385
GC=1.02
GC=1.01
GC=1.02
EA.RS3 A N=53,747
GC=1.01
GC=1.09
GC=1.10
EA.RS3 B N=298,434
GC=1.07
GC=1.04
GC=1.10
EA.RS3 C N=260,631
GC=1.04
GC=1.06
GC=1.07
EA.RS3 D N=7,617,573
GC=1.02
GC=1.01
GC=1.02
EA.CHS N=7,586,170
GC=1.02
GC=1.03
GC=1.03
EA.CHS A N=262,451
GC=1.05
GC=1.16
GC=1.14
EA.CHS B N=484,539
GC=0.99
GC=1.08
GC=1.06
EA.CHS C N=256,270
GC=1.01
GC=1.06
GC=1.05
EA.CHS D N=6,582,910
GC=1.02
GC=1.02
GC=1.02
EA.RS1 N=8,233,250
GC=1.02
GC=1.02
GC=1.03
EA.RS1 C N=122,174
GC=1.04
GC=1.07
GC=1.06
EA.RS1 D N=8,111,076
GC=1.02
GC=1.02
GC=1.03
Observed -logP
Expected -logP


EA.GS N=9,029,737
GC=1.03
GC=1.02
GC=1.02
EA.GS A N=10
GC=1.82
GC=1.57
GC=1.38
EA.GS B N=4,025
GC=1.02
GC=1.22
GC=1.13
EA.GS C N=267,589
GC=1.02
GC=1.08
GC=1.07
EA.GS D N=8,758,113
GC=1.03
GC=1.01
GC=1.02
EA.ARIC N=8,994,418
GC=1.04
GC=1.00
GC=1.03
EA.WGHS N=7,830,272
GC=1.06
GC=0.97
GC=1.02
EA.HERTIAGE N=5,614,529
GC=1.08
GC=1.11
GC=1.13
EA.HERTIAGE A N=214,528
GC=1.11
GC=1.22
GC=1.24
EA.HERTIAGE B N=1,120,834
GC=1.13
GC=1.17
GC=1.20
EA.HERTIAGE C N=202,261
GC=1.09
GC=1.12
GC=1.14
EA.HERTIAGE D N=4,076,906
GC=1.07
GC=1.09
GC=1.10
EA.GENOA N=7,321,199
GC=1.06
GC=1.05
GC=1.08
EA.GENOA A N=266,605
GC=1.09
GC=1.17
GC=1.17
EA.GENOA B N=921,596
GC=1.06
GC=1.13
GC=1.14
EA.GENOA C N=186,290
GC=1.06
GC=1.04
GC=1.09
EA.GENOA D N=5,946,708
GC=1.06
GC=1.03
GC=1.07
EA.HyperGEN N=5,833,650
GC=1.04
GC=1.09
GC=1.08
EA.HyperGEN A N=238,450
GC=1.07
GC=1.25
GC=1.19
EA.HyperGEN B N=783,915
GC=1.07
GC=1.18
GC=1.17
EA.HyperGEN C N=242,366
GC=1.01
GC=1.12
GC=1.08
EA.HyperGEN D N=4,568,919
GC=1.04
GC=1.07
GC=1.06
Observed -logP
Expected -logP

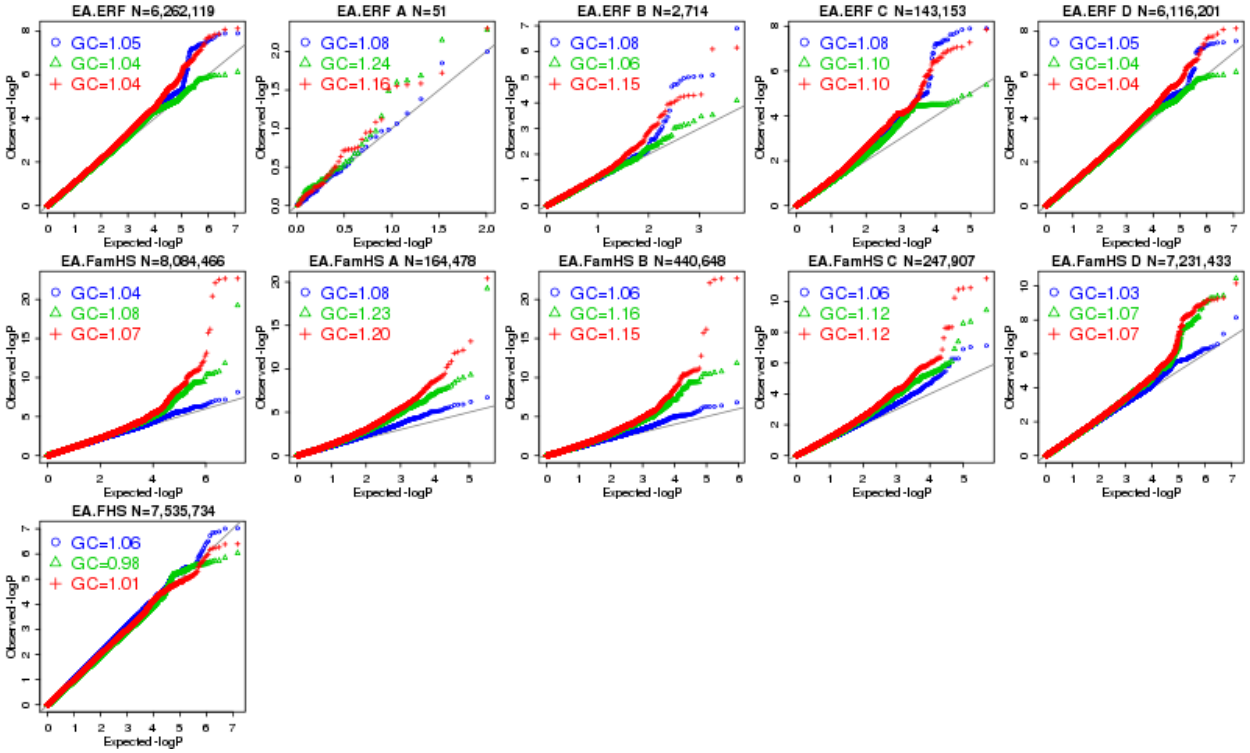
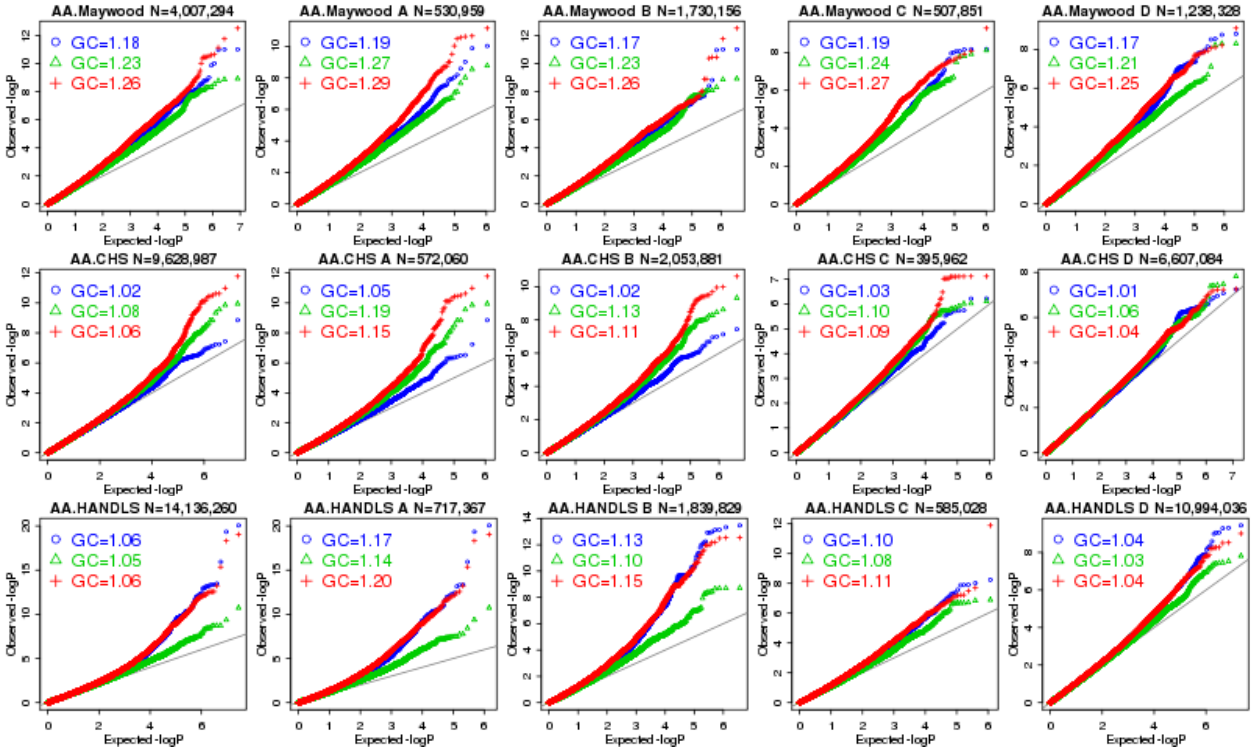
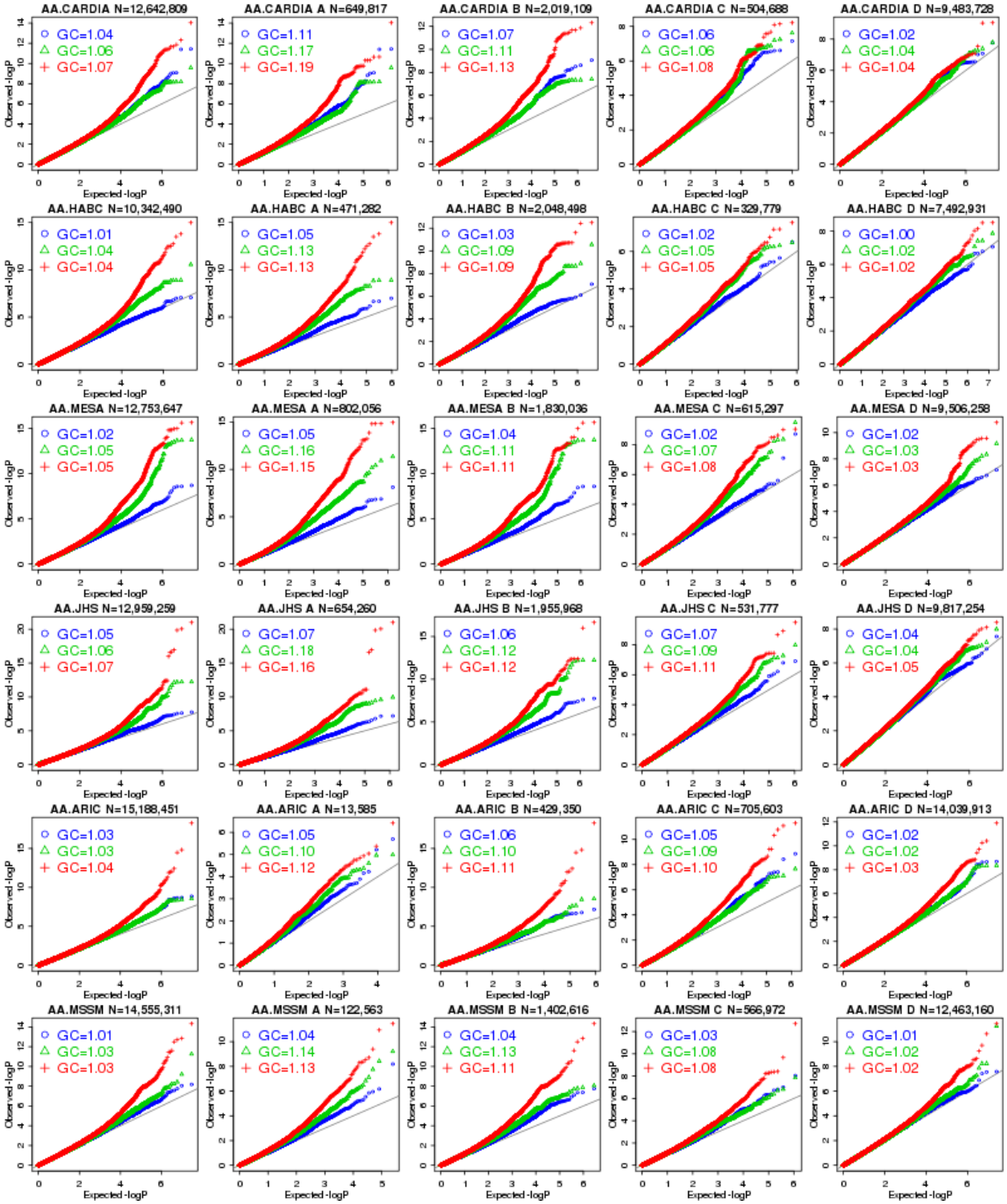


Figure S15: Cohort-specific QQ Plots in African ancestry




AA.CARDIA N=12,642,809
GC=1.04
GC=1.06
GC=1.07
AA.CARDIA A N=649,817
GC=1.11
GC=1.17
GC=1.19
AA.CARDIA B N=2,019,109
GC=1.07
GC=1.11
GC=1.13
AA.CARDIA C N=504,688
GC=1.06
GC=1.06
GC=1.08
AA.CARDIA D N=9,483,728
GC=1.02
GC=1.04
GC=1.04
AA.HABC N=10,342,490
GC=1.01
GC=1.04
GC=1.04
AA.HABC A N=471,282
GC=1.05
GC=1.13
GC=1.13
AA.HABC B N=2,048,498
GC=1.03
GC=1.09
GC=1.09
AA.HABC C N=329,779
GC=1.02
GC=1.05
GC=1.05
AA.HABC D N=7,492,931
GC=1.00
GC=1.02
GC=1.02
AA.MESA N=12,753,647
GC=1.02
GC=1.05
GC=1.05
AA.MESA A N=802,056
GC=1.05
GC=1.16
GC=1.15
AA.MESA B N=1,830,036
GC=1.04
GC=1.11
GC=1.11
AA.MESA C N=615,297
GC=1.02
GC=1.07
GC=1.08
AA.MESA D N=9,506,258
GC=1.02
GC=1.03
GC=1.03
AA.JHS N=12,959,259
GC=1.05
GC=1.06
GC=1.07
AA.JHS A N=654,260
GC=1.07
GC=1.18
GC=1.16
AA.JHS B N=1,955,968
GC=1.06
GC=1.12
GC=1.12
AA.JHS C N=531,777
GC=1.07
GC=1.09
GC=1.11
AA.JHS D N=9,817,254
GC=1.04
GC=1.04
GC=1.05
AA.ARIC N=15,188,451
GC=1.03
GC=1.03
GC=1.04
AA.ARIC A N=13,585
GC=1.05
GC=1.10
GC=1.12
AA.ARIC B N=429,350
GC=1.06
GC=1.10
GC=1.11
AA.ARIC C N=705,603
GC=1.05
GC=1.09
GC=1.10
AA.ARIC D N=14,039,913
GC=1.02
GC=1.02
GC=1.03
AA.MSSM N=14,555,311
GC=1.01
GC=1.03
GC=1.03
AA.MSSM A N=122,563
GC=1.04
GC=1.14
GC=1.13
AA.MSSM B N=1,402,616
GC=1.04
GC=1.13
GC=1.11
AA.MSSM C N=566,972
GC=1.03
GC=1.08
GC=1.08
AA.MSSM D N=12,463,160
GC=1.01
GC=1.02
GC=1.02
Observed -logP
Expected -logP

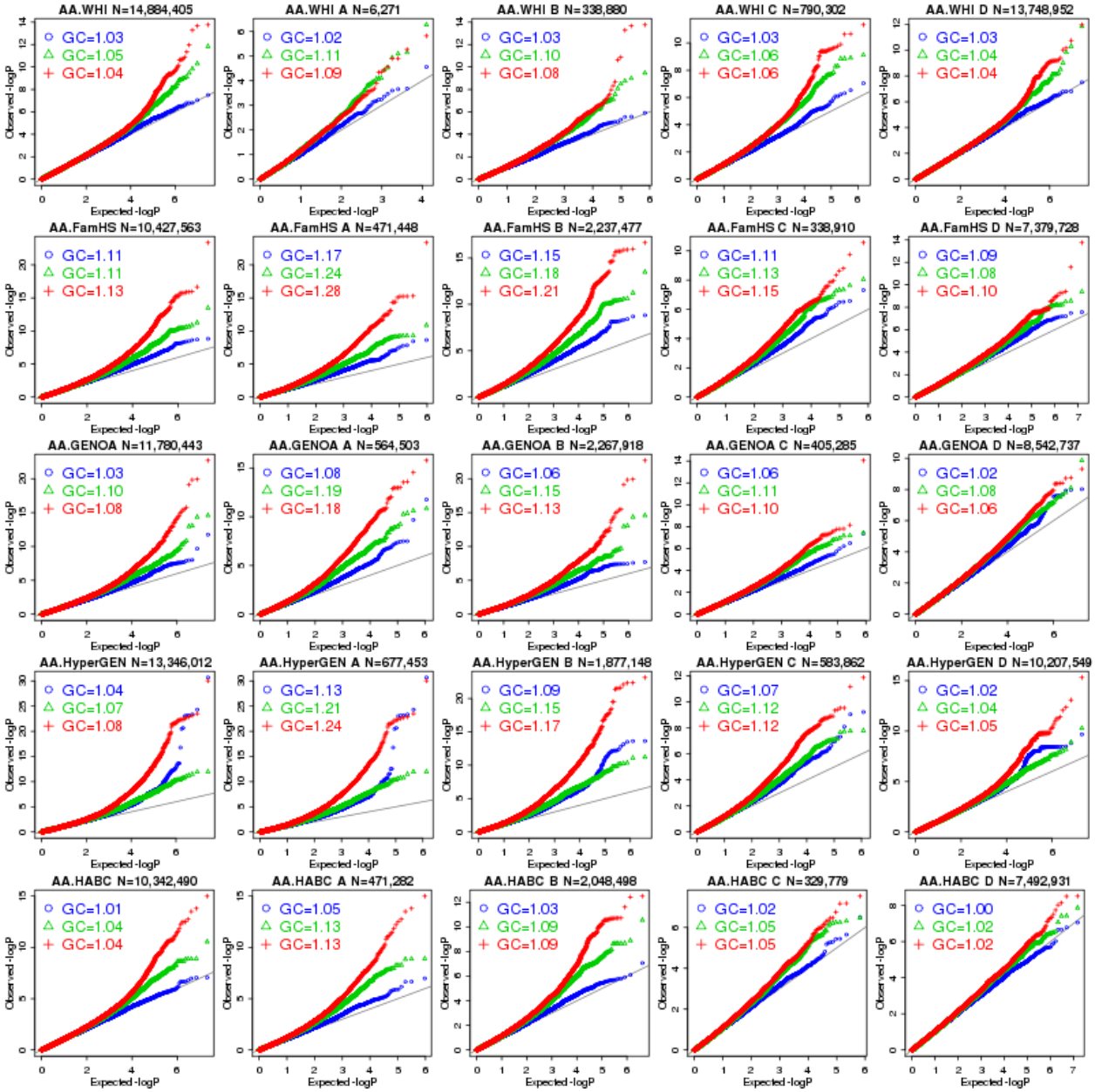
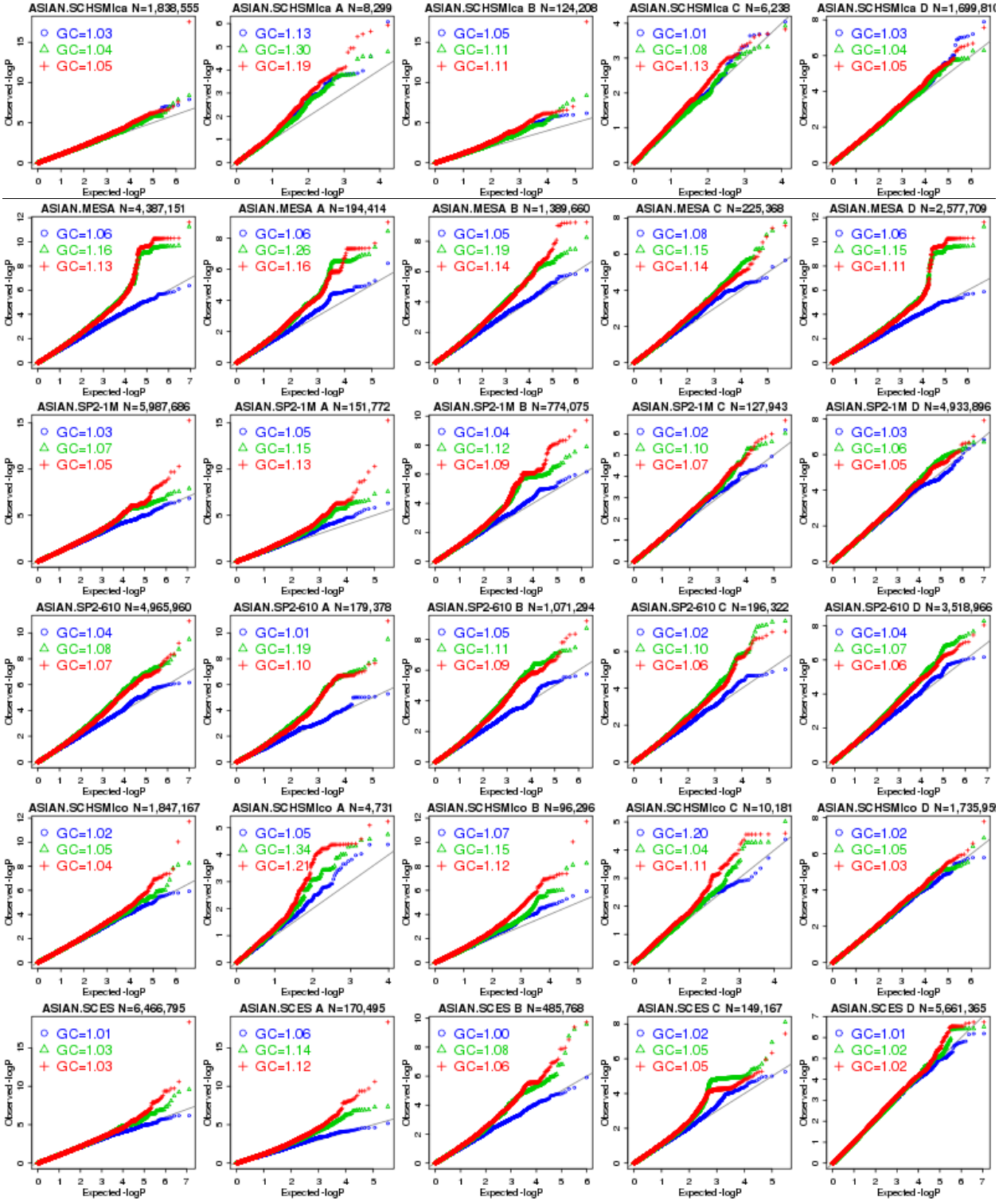

AA.WHI N=14,884,405
GC=1.03
GC=1.05
GC=1.04
AA.WHI A N=6,271
GC=1.02
GC=1.11
GC=1.09
AA.WHI B N=338,880
GC=1.03
GC=1.10
GC=1.08
AA.WHI C N=790,302
GC=1.03
GC=1.06
GC=1.06
AA.WHI D N=13,748,952
GC=1.03
GC=1.04
GC=1.04
AA.FamHS N=10,427,563
GC=1.11
GC=1.11
GC=1.13
AA.FamHS A N=471,448
GC=1.17
GC=1.24
GC=1.28
AA.FamHS B N=2,237,477
GC=1.15
GC=1.18
GC=1.21
AA.FamHS C N=338,910
GC=1.11
GC=1.13
GC=1.15
AA.FamHS D N=7,379,728
GC=1.09
GC=1.08
GC=1.10
AA.GENOA N=11,780,443
GC=1.03
GC=1.10
GC=1.08
AA.GENOA A N=564,503
GC=1.08
GC=1.19
GC=1.18
AA.GENOA B N=2,267,918
GC=1.06
GC=1.15
GC=1.13
AA.GENOA C N=405,285
GC=1.06
GC=1.11
GC=1.10
AA.GENOA D N=8,542,737
GC=1.02
GC=1.08
GC=1.06
AA.HyperGEN N=13,346,012
GC=1.04
GC=1.07
GC=1.08
AA.HyperGEN A N=677,453
GC=1.13
GC=1.21
GC=1.24
AA.HyperGEN B N=1,877,148
GC=1.09
GC=1.15
GC=1.17
AA.HyperGEN C N=583,862
GC=1.07
GC=1.12
GC=1.12
AA.HyperGEN D N=10,207,549
GC=1.02
GC=1.04
GC=1.05
AA.HABC N=10,342,490
GC=1.01
GC=1.04
GC=1.04
AA.HABC A N=471,282
GC=1.05
GC=1.13
GC=1.13
AA.HABC B N=2,048,498
GC=1.03
GC=1.09
GC=1.09
AA.HABC C N=329,779
GC=1.02
GC=1.05
GC=1.05
AA.HABC D N=7,492,931
GC=1.00
GC=1.02
GC=1.02
Observed -logP
Expected -logP

Figure S16: Cohort-specific QQ Plots in Asian ancestry



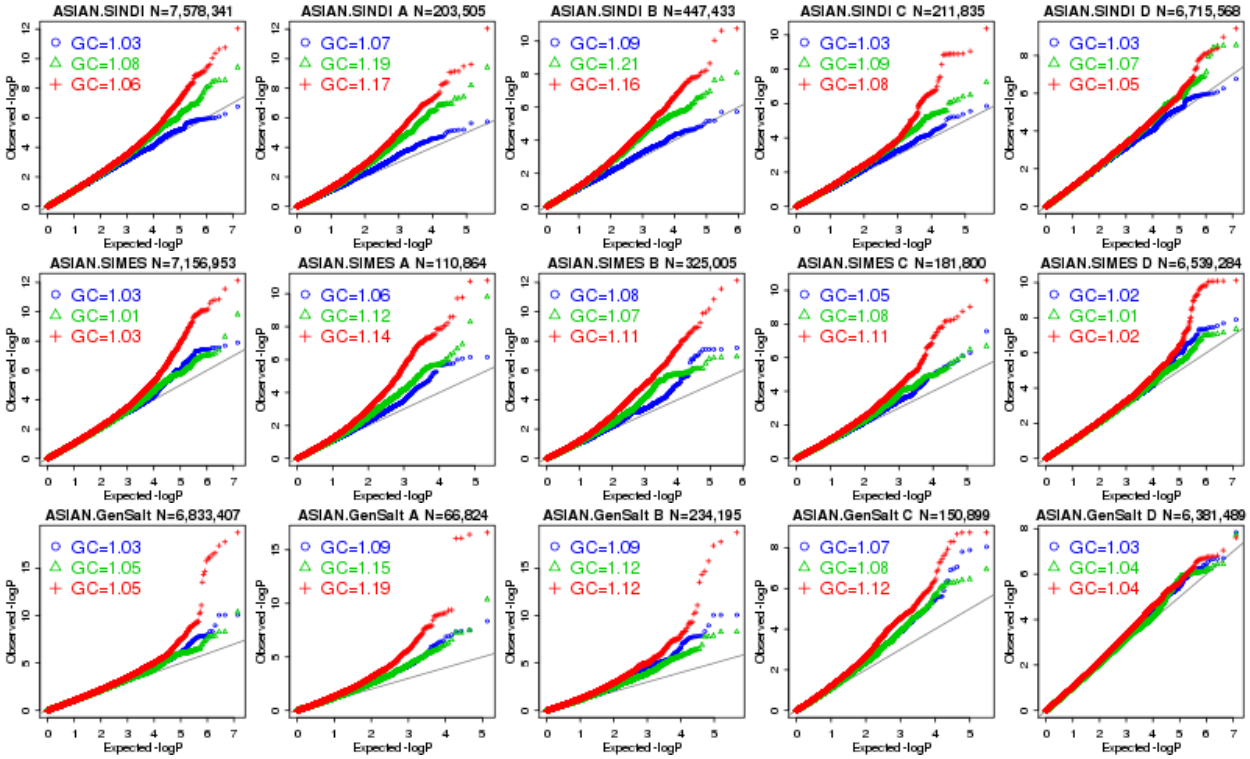


Figure S17: Cohort-specific QQ Plots in Hispanic ancestry