The power of genetic diversity in genome-wide association studies of lipids

VA Million Veteran Program; Global Lipids Genetics Consortium; Graham, Sarah E; Clarke, Shoa L; Wu, Kuan-Han H; Kanoni, Stavroula

Published in:
Nature

DOI:
10.1038/s41586-021-04064-3

Publication date:
2021

Document Version
Peer reviewed version

Title: The power of genetic diversity in genome-wide association studies of lipids

Authors:
Sarah E Graham[1], Shoa L Clarke[2,3‡], Kuan-Han H Wu[4‡], Stavroula Kanoni[5‡], Greg JM Zajac[6‡], Shweta Ramdas[7‡], Ida Surakka[1], Ioanna Ntalla[8], Sailaja Vedantam[9,10], Thomas W Winkler[11], Adam E Locke[12], Eirini Marouli[5], Mi Yeong Hwang[13], Sohee Han[13], Akira Narita[14], Ananyo Choudhury[15], Amy R Bentley[16], Kenneth Ekoru[16], Anurag Verma[17], Bhavi Trivedi[18], Hilary C Martin[19], Karen A Hunt[18], Qin Hui[20,21], Derek Klarin[22,23,24], Xiang Zhu[25,26,27,28], Gudmar Thorleifsson[29], Anna Helgadottir[29], Daniel F Gudbjartsson[29,30], Hilma Holm[29], Isleifur Olafsson[31], Masato Akiyama[32,33], Saori Sakaue[34,32,35], Chikashi Terao[36], Masahiro Kanai[37,38,39], Wei Zhou[40,41,42], Ben M Brumpton[43,44,45], Humaira Rasheed[43,44], Sanni E Ruotsalainen[46], Aki S Havulinna[46,47], Yogasudha Veturi[48], QiPing Feng[49], Elisabeth A Rosenthal[50], Todd Lingren[51], Jennifer Allen Pacheco[52], Sarah A Pendergrass[53], Jeffrey Haessler[54], Franco Giulianini[55], Yuki Bradford[48], Jason E Miller[48], Archie Campbell[56,57], Kuang Lin[58], Iona Y Millwood[58,59], George Hindy[60], Asif Rasheed[61], Jessica D Faul[62], Wei Zhao[63], David R Weir[62], Constance Turman[64], Hongyan Huang[64], Mariaelisa Graff[65], Anubha Mahajan[66#], Michael R Brown[67], Weihua Zhang[68,69,70], Ketian Yu[71], Ellen M Schmidt[71], Anita Pandit[71], Stefan Gustafsson[72], Xianyong Yin[71], Jian'an Luan[73], Jing-Hua Zhao[73], Fumihiko Matsuda[74], Hye-Mi Jang[13], Kyungheon Yoon[13], Carolina Medina-Gomez[75,76], Achilleas Pitsillides[77], Jouke Jan Hottenga[78,79], Gonneke Willemsen[78,80], Andrew R Wood[81], Yingji Ji[81], Zishan Gao[82,83,84], Simon Haworth[85,86], Ruth E Mitchell[85,87], Jin Fang Chai[88], Mette Aadahl[89], Jie Yao[90], Ani Manichaikul[91], Helen R Warren[92,93], Julia Ramirez[92], Jette Bork-Jensen[94], Line L Kårhus[95], Anuj Goel[96,97], Maria Sabater-Lleal[98,99], Raymond Noordam[100], Carlo Sidore[101], Edoardo Fiorillo[102], Aaron F McDaid[103,104], Pedro Marques-Vidal[105], Matthias Wielscher[106], Stella Trompet[107,108], Naveed Sattar[109], Line T Møllehave[89], Betina H Thuesen[89], Matthias Munz[110], Lingyao Zeng[111,112], Jianfeng Huang[113], Bin Yang[113], Alaitz Poveda[114], Azra Kurbasic[114], Claudia Lamina[115], Lukas Forer[115], Markus Scholz[116,117], Tessel E. Galesloot[118], Jonathan P. Bradfield[119], E Warwick Daw[120], Joseph M Zmuda[121], Jonathan S Mitchell[122], Christian Fuchsberger[122], Henry Christensen[123], Jennifer A Brody[124], Mary F Feitosa[120], Mary K Wojczynski[120], Michael Preuss[125], Massimo Mangino[126,127], Paraskevi Christofidou[126], Niek Verweij[128], Jan W Benjamins[128], Jorgen Engmann[129,130], Rachel L Kember[131], Roderick C Slieker[132,133], Ken Sin Lo[134], Nuno R Zilhao[135], Phuong Le[136], Marcus E Kleber[137,138], Graciela E Delgado[137], Shaofeng Huo[139], Daisuke D Ikeda[140], Hiroyuki Iha[140], Jian Yang[141,142], Jun Liu[143], Hampton L Leonard[144,145], Jonathan Marten[146], Börge Schmidt[147], Marina Arendt[147,148], Laura J Smyth[149], Marisa Cañadas-Garre[149], Chaolong Wang[150,151], Masahiro Nakatochi[152], Andrew Wong[153], Nina Hutri-Kähönen[154,155], Xueling Sim[88], Rui Xia[156], Alicia Huerta-Chagoya[157], Juan Carlos Fernandez-Lopez[158], Valeriya Lyssenko[159,160], Meraj Ahmed[161], Anne U Jackson[6], Marguerite R Irvin[162], Christopher Oldmeadow[163], Han-Na Kim[164,165], Seungho Ryu[166,167], Paul RHJ Timmers[168,146], Liubov Arbeeva[169], Rajkumar Dorajoo[170], Leslie A Lange[171], Xiaoran Chai[172,173], Gauri Prasad[174,175], Laura Lorés-Motta[176], Marc Pauper[176], Jirong Long[177], Xiaohui Li[90], Elizabeth Theusch[178], Fumihiko Takeuchi[179], Cassandra N Spracklen[180,181], Anu Loukola[46], Sailalitha Bollepalli[46], Sophie C Warner[182,183], Ya Xing Wang[184], Wen B. Wei[185], Teresa

Nutile[186], Daniela Ruggiero[186,187], Yun Ju Sung[188], Yi-Jen Hung[189], Shufeng Chen[113], Fangchao Liu[113], Jingyun Yang[190,191], Katherine A Kentistou[168], Mathias Gorski[11,192], Marco Brumat[193], Karina Meidtner[194,195], Lawrence F Bielak[196], Jennifer A Smith[196,197], Prashantha Hebbar[198], Aliki-Eleni Farmaki[199,200], Edith Hofer[201,202], Maoxuan Lin[203], Chao Xue[204], Jifeng Zhang[1], Maria Pina Concas[205], Simona Vaccargiu[206], Peter J van der Most[207], Niina Pitkänen[208,209], Brian E Cade[210,211], Jiwon Lee[210], Sander W. van der Laan[212], Kumaraswamy Naidu Chitrala[213], Stefan Weiss[214], Martina E Zimmermann[11], Jong Young Lee[215], Hyeok Sun Choi[216], Maria Nethander[217,218], Sandra Freitag-Wolf[219], Lorraine Southam[220,221], Nigel W Rayner[222,223,224,220], Carol A Wang[225], Shih-Yi Lin[226,227,228], Jun-Sing Wang[229,230], Christian Couture[231], Leo-Pekka Lyytikäinen[232,233], Kjell Nikus[234,235], Gabriel Cuellar-Partida[236], Henrik Vestergaard[237], Bertha Hildalgo[238], Olga Giannakopoulou[5], Qiuyin Cai[177], Morgan O Obura[132], Jessica van Setten[239], Xiaoyin Li[240], Karen Schwander[241], Natalie Terzikhan[242], Jae Hun Shin[216], Rebecca D Jackson[243], Alexander P Reiner[244], Lisa Warsinger Martin[245], Zhengming Chen[58,59], Liming Li[246], Heather M Highland[65], Kristin L Young[65], Takahisa Kawaguchi[74], Joachim Thiery[247,117], Joshua C Bis[124], Girish N. Nadkarni[125], Lenore J Launer[248], Huaixing Li[139], Mike A Nalls[144,145], Olli T Raitakari[249,250,251], Sahoko Ichihara[252], Sarah H Wild[253], Christopher P Nelson[182,183], Harry Campbell[168], Susanne Jäger[194,195], Toru Nabika[254], Fahd Al-Mulla[198], Harri Niinikoski[255,256], Peter S Braund[182,183], Ivana Kolcic[257], Peter Kovacs[258], Tota Giardoglou[259], Tomohiro Katsuya[260,261], Konain Fatima Bhatti[5], Dominique de Kleijn[262], Gert J. de Borst[262], Eung Kweon Kim[263], Hieab H.H. Adams[242,264], M. Arfan Ikram[242], Xiaofeng Zhu[240], Folkert W Asselbergs[239], Adriaan O Kraaijeveld[239], Joline WJ Beulens[132,265], Xiao-Ou Shu[177], Loukianos S Rallidis[266], Oluf Pedersen[94], Torben Hansen[94], Paul Mitchell[267], Alex W Hewitt[268,269], Mika Kähönen[270,271], Louis Pérusse[231,272], Claude Bouchard[273], Anke Tönjes[274], Yii-Der Ida Chen[90], Craig E Pennell[225], Trevor A Mori[275], Wolfgang Lieb[276], Andre Franke[277], Claes Ohlsson[217,278], Dan Mellström[217,279], Yoon Shin Cho[216], Hyejin Lee[280], Jian-Min Yuan[281,282], Woon-Puay Koh[283,284], Sang Youl Rhee[285], Jeong-Taek Woo[285], Iris M Heid[11], Klaus J Stark[11], Henry Völzke[286], Georg Homuth[214], Michele K Evans[287], Alan B Zonderman[287], Ozren Polasek[257], Gerard Pasterkamp[212], Imo E Hoefer[212], Susan Redline[210,211], Katja Pahkala[208,209,288], Albertine J Oldehinkel[289], Harold Snieder[207], Ginevra Biino[290], Reinhold Schmidt[201], Helena Schmidt[291], Y Eugene Chen[1], Stefania Bandinelli[292], George Dedoussis[199], Thangavel Alphonse Thanaraj[293], Sharon LR Kardia[196], Norihiro Kato[179], Matthias B Schulze[194,195,294], Giorgia Girotto[193,295], Bettina Jung[296], Carsten A Böger[296,297,298], Peter K Joshi[168], David A Bennett[190,191], Philip L De Jager[299,300], Xiangfeng Lu[113], Vasiliki Mamakou[301,302], Morris Brown[303,93], Mark J Caulfield[92,93], Patricia B Munroe[92,93], Xiuqing Guo[90], Marina Ciullo[186,187], Jost B. Jonas[304,184,305], Nilesh J Samani[182,183], Daniel I. Chasman[55,306], Jaakko Kaprio[46], Päivi Pajukanta[307], Teresa Tusié-Luna[308,309], Carlos A Aguilar-Salinas[310], Linda S Adair[311,312], Sonny Augustin Bechayda[313,314], H. Janaka de Silva[315], Ananda R Wickremasinghe[316], Ronald Krauss[317], Jer-Yuarn Wu[318], Wei Zheng[177], Anneke I den Hollander[176], Dwaipayan Bharadwaj[175,319], Adolfo Correa[320], James G Wilson[321], Lars Lind[322], Chew-Kiat Heng[323], Amanda E Nelson[169,324], Yvonne M Golightly[169,325,326,327], James F Wilson[168,146], Brenda Penninx[328,80], Hyung-Lae Kim[329], John Attia[330,163], Rodney J Scott[330,163], D C Rao[331], Donna K Arnett[332], Mark Walker[333], Heikki A Koistinen[334,335,336], Giriraj R Chandak[161,337], Chittaranjan S Yajnik[338], Josep M Mercader[339,340,341], Teresa Tusie-Luna[342], Carlos Aguilar-Salinas[343], Clicerio Gonzalez Villalpando[344], Lorena Orozco[345], Myriam

Fornage[156,346], E Shyong Tai[347,88], Rob M van Dam[88,347], Terho Lehtimäki[232,233], Nish Chaturvedi[153], Mitsuhiro Yokota[348], Jianjun Liu[151], Dermot F Reilly[349], Amy Jayne McKnight[149], Frank Kee[149], Karl-Heinz Jöckel[147], Mark I McCarthy[66,350#], Colin NA Palmer[351], Veronique Vitart[146], Caroline Hayward[146], Eleanor Simonsick[352], Cornelia M van Duijn[143], Fan Lu[353], Jia Qu[353], Haretsugu Hishigaki[140], Xu Lin[354], Winfried März[355,356,137], Esteban J Parra[136], Miguel Cruz[357], Vilmundur Gudnason[135,358], Jean-Claude Tardif[134,359], Guillaume Lettre[134,359], Leen M t Hart[133,360,132], Petra JM Elders[361], Daniel J Rader[362], Scott M Damrauer[363,364], Meena Kumari[365], Mika Kivimaki[130], Pim van der Harst[128], Tim D Spector[126], Ruth J.F. Loos[125,366], Michael A Province[120], Bruce M Psaty[367,368], Ivan Brandslund[123,369], Peter P Pramstaller[122], Kaare Christensen[370], Samuli Ripatti[46,371,372], Elisabeth Widén[46], Hakon Hakonarson[373,374], Struan F.A. Grant[374,375,376], Lambertus ALM Kiemeney[118], Jacqueline de Graaf[118], Markus Loeffler[116,117], Florian Kronenberg[115], Dongfeng Gu[113,377], Jeanette Erdmann[378], Heribert Schunkert[111,112], Paul W Franks[114], Allan Linneberg[89,379], J. Wouter Jukema[107,380], Amit V Khera[381,382,383,384], Minna Männikkö[385], Marjo-Riitta Jarvelin[106,386,387], Zoltan Kutalik[103,104], Francesco Cucca[388,389], Dennis O Mook-Kanamori[390,391], Ko Willems van Dijk[392,393,394], Hugh Watkins[96,97], David P Strachan[395], Niels Grarup[94], Peter Sever[396], Neil Poulter[397], Jerome I Rotter[90], Thomas M Dantoft[89], Fredrik Karpe[398,399], Matt J Neville[398,399], Nicholas J Timpson[85,87], Ching-Yu Cheng[172,400], Tien-Yin Wong[172,400], Chiea Chuen Khor[151], Charumathi Sabanayagam[172,400], Annette Peters[84,401,402], Christian Gieger[83,84,402], Andrew T Hattersley[403], Nancy L Pedersen[404], Patrik KE Magnusson[404], Dorret I Boomsma[405,79,406], Eco JC de Geus[78,80], L Adrienne Cupples[77,407], Joyce B.J. van Meurs[75,76], Mohsen Ghanbari[76,408], Penny Gordon-Larsen[311,312], Wei Huang[409], Young Jin Kim[13], Yasuharu Tabara[74], Nicholas J Wareham[73], Claudia Langenberg[73], Eleftheria Zeggini[220,221,410], Johanna Kuusisto[411], Markku Laakso[411], Erik Ingelsson[412,413,414,72], Goncalo Abecasis[71,415], John C Chambers[416,68,69,417], Jaspal S Kooner[69,70,418,419], Paul S de Vries[67], Alanna C Morrison[67], Kari E. North[65], Martha Daviglus[420], Peter Kraft[64,421], Nicholas G Martin[422], John B Whitfield[422], Shahid Abbas[423], Danish Saleheen[61,424,425], Robin G Walters[426,427,428], Michael V Holmes[58,59,429], Corri Black[430], Blair H Smith[431], Anne E Justice[432], Aris Baras[433], Julie E Buring[55,434], Paul M Ridker[55,434], Daniel I Chasman[55,434], Charles Kooperberg[54], Wei-Qi Wei[435], Gail P Jarvik[436], Bahram Namjou[437], M. Geoffrey Hayes[438,439,440], Marylyn D Ritchie[48], Pekka Jousilahti[47], Veikko Salomaa[47], Kristian Hveem[43,441,442], Bjørn Olav Åsvold[43,441,443], Michiaki Kubo[444], Yoichiro Kamatani[32,445], Yukinori Okada[34,32,446,447], Yoshinori Murakami[448], Unnur Thorsteinsdottir[29,449], Kari Stefansson[29,449], Yuk-Lam Ho[450], Julie A Lynch[451,452], Daniel Rader[453], Philip S Tsao[2,3,454], Kyong-Mi Chang[455,453], Kelly Cho[450,456], Christopher J O'Donnell[450,456], J. Michael Gaziano[450,456], Peter Wilson[457,458], Charles N Rotimi[16], Scott Hazelhurst[15,459], Michèle Ramsay[15,460], Richard C Trembath[461], David A van Heel[18], Gen Tamiya[14], Masayuki Yamamoto[14], Bong-Jo Kim[13], Karen L Mohlke[180], Timothy M Frayling[81], Joel N Hirschhorn[9,10,462], Sekar Kathiresan[463,382,384], VA Million Veteran Program, Global Lipids Genetics Consortium, Michael Boehnke[6], Pradeep Natarajan[464,465,466,467], Gina M Peloso[468†], Christopher D Brown[7†], Andrew P Morris[469†], Themistocles L Assimes[2,3,454†*], Panos Deloukas[5,470†], Yan V Sun[20,21†], Cristen J Willer[1,471,472†*]

Affiliations:
[1]Department of Internal Medicine, Division of Cardiology, University of Michigan, Ann Arbor, MI 48109, USA, [2]VA Palo Alto Health Care system, Palo Alto, California, USA, [3]Department of Medicine, Division of Cardiovascular Medicine, Stanford University School of Medicine, Stanford, California, USA, [4]Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI, USA, [5]William Harvey Research Institute, Barts and the London School of Medicine and Dentistry, Queen Mary University of London, Charterhouse square, EC1M 6BQ, UK, [6]Department of Biostatistics and Center for Statistics Genetics, University of Michigan, Ann Arbor, MI 48109, [7]Department of Genetics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, [8]Clinical Pharmacology, William Harvey Research Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, EC1M 6BQ UK, [9]Endocrinology, Boston Childrens Hospital, Boston 02115 MA,USA, [10]Medical and Population Genetics, Broad Institute, 75 Ames street, Cambridge, MA 02142,USA, [11]Department of Genetic Epidemiology, University of Regensburg, Regensburg, Germany, [12]McDonnell Genome Institute and Department of Medicine, Washington University, St. Louis, MO, 63108, [13]Division of Genome Research, Center for Genome Science, National Institute of Health, Chungcheongbuk-do, South Korea, [14]Tohoku Medical Megabank Organization, Tohoku University, Sendai 980-8573, Japan, [15]Sydney Brenner Institute for Molecular Bioscience, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa., [16]Center for Research on Genomics and Global Health, National Human Genome Research Institute, National Institutes of Health, 12 South Drive, Room 4047, Bethesda, MD, 20892, USA, [17]Department of Genetics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA., [18]Blizard Institute, Barts and the London School of Medicine and Dentistry, Queen Mary University of London, London, UK, [19]Wellcome Sanger Institute, Hinxton, UK, [20]Department of Epidemiology, Emory University Rollins School of Public Health, Atlanta, Georgia, USA, [21]Atlanta VA Health Care System, Decatur, Georgia, USA, [22]Malcolm Randall VA Medical Center, Gainesville, FL, [23]Division of Vascular Surgery and Endovascular Therapy, University of Florida College of Medicine, Gainesville, FL, [24]Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, Massachusetts, USA, [25]Department of Statistics, The Pennsylvania State University, University Park, PA, USA, [26]Huck Institutes of the Life Sciences, The Pennsylvania State University, University Park, PA, USA, [27]VA Palo Alto Health Care System, Palo Alto, CA, USA, [28]Department of Statistics, Stanford University, Stanford, CA, USA, [29]deCODE genetics/Amgen, Inc. Sturlugata 8, Reykjavik, 102, Iceland, [30]School of Engineering and Natural Sciences, University of Iceland, Sæmundargötu 2, Reykjavik, 102, Iceland, [31]Department of Clinical Biochemistry, Landspitali - National University Hospital of Iceland, Hringbraut, Reykjavik, 101, Iceland, [32]Laboratory for Statistical Analysis, RIKEN Center for Integrative Medical Sciences, [33]Department of Ophthalmology, Graduate School of Medical Sciences, Kyushu University, Fukuoka, Japan, [34]Department of Statistical Genetics, Osaka University Graduate School of Medicine, Osaka, Japan, [35]Department of Allergy and Rheumatology, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan, [36]Laboratory for Statistical and Translational Genetics, RIKEN Center for Integrative Medical

Sciences, Yokohama, Japan., [37]Laboratory for Statistical Analysis, RIKEN Center for Integrative Medical Sciences, Yokohama, Japan., [38]Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, MA, USA., [39]Department of Biomedical Informatics, Harvard Medical School, Boston, MA, USA., [40]Analytic and Translational Genetics Unit, Massachusetts General Hospital, Boston, Massachusetts, USA, [41]Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, Michigan, USA, [42]Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, Cambridge, Massachusetts, USA, [43]K.G. Jebsen Center for Genetic Epidemiology, Department of Public Health and Nursing, NTNU, Norwegian University of Science and Technology, Trondheim, Norway, [44]MRC Integrative Epidemiology Unit, University of Bristol, UK, [45]Department of Thoracic Medicine, St. Olavs Hospital, Trondheim University Hospital, Trondheim, Norway, [46]Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Tukholmankatu 8, 00014 Helsinki, Finland, [47]Finnish institute for Health and Welfare, Helsinki, Finland, [48]Department of Genetics, Institute for Biomedical Informatics, University of Pennsylvania, Perelman School of Medicine, Philadelphia, PA 19104, USA, [49]Division of Clinical Pharmacology, Department of Medicine, Vanderbilt University Medical Center, Nashville, TN, [50]Department of Medicine (Medical Genetics), University of Washington, [51]Division of Biomedical Informatics, Cincinnati Children's Hospital Medical Center, [52]Center for Genetic Medicine, Northwestern University, [53]Genentech, 1 DNA Way, South San Francisco, 94084, USA, [54]Fred Hutchinson Cancer Research Center, Division of Public Health Sciences, Seattle WA 9810, USA, [55]Division of Preventive Medicine, Brigham and Women's Hospital, Boston, MA 02215, [56]Centre for Genomic and Experimental Medicine , Institute of Genetics & Molecular Medicine, University of Edinburgh, Western General Hospital, Edinburgh EH4 2XU, United Kingdom, [57]Usher Institute for Population Health Sciences and Informatics, The University of Edinburgh, Nine, Edinburgh Bioquarter, 9 Little France Road, Edinburgh, EH16 4UX., [58]Clinical Trial Service Unit and Epidemiological Studies Unit, Nuffield Department of Population Health, University of Oxford, Oxford OX3 7LF, [59]Medical Research Council Population Health Research Unit, Nuffield Department of Population Health, University of Oxford, Oxford OX3 7LF, [60]Department of Population Medicine, Qatar University College of Medicine, QU Health, Doha, Qatar, [61]Center for Non-Communicable Diseases, Karachi, Sindh, Pakistan, [62]Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor, MI, 48104, [63]Department of Epidemiology, School of Public Health, University of Michigan, Ann Arbor, MI, 48109, [64]Program in Genetic Epidemiology and Statistical Genetics, Department of Epidemiology, Harvard T.H. Chan School of Public Health, 677 Huntington Avenue, Boston, MA, 02115, USA, [65]Department of Epidemiology, University of North Carolina, Chapel Hill, NC, [66]Wellcome Centre for Human Genetics, University of Oxford, UK, [67]Human Genetics Center, Department of Epidemiology, Human Genetics, and Environmental Sciences, School of Public Health, The University of Texas Health Science Center at Houston, Houston, Texas, 77030, USA, [68]Department of Epidemiology and Biostatistics, Imperial College London, London W2 1PG, UK, [69]Department of Cardiology, Ealing Hospital, London North West University Healthcare NHS Trust, Middlesex UB1 3HW, UK, [70]Imperial College Healthcare NHS Trust, London W12 0HS, UK, [71]Department of Biostatistics, Center for Statistical Genetics, University of Michigan, Ann Arbor, [72]Department of Medical Sciences,

Molecular Epidemiology and Science for Life Laboratory, Uppsala University, Uppsala, Sweden., [73]MRC Epidemiology Unit, University of Cambridge School of Clinical Medicine, Cambridge, CB2 0QQ, UK, [74]Center for Genomic Medicine, Kyoto University Graduate School of Medicine, [75]Department of Internal Medicine, Erasmus MC, University Medical Center Rotterdam, [76]Department of Epidemiology, Erasmus MC, University Medical Center Rotterdam, [77]Department of Biostatistics, Boston University School of Public Health, 801 Massachusetts Ave, Boston, MA 02118, USA, [78]Department of Biological Psychology, Behavioral and Movement Sciences, Vrije Universiteit Amsterdam, [79]Amsterdam Public Health, VU medical center Amsterdam, [80]Amsterdam Public Health research institute, VU medical center Amsterdam, [81]Genetics of Complex Traits, University of Exeter Medical School, University of Exeter, Exeter, EX2 5DW, UK, [82]Department of Clinical Acupuncture and Moxibustion, Nanjing University of Chinese Medicine, Nanjing, Jiangsu 210029, China, [83]Research Unit of Molecular Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany, [84]Institute of Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany, [85]MRC Integrative Epidemiology Unit at the University of Bristol, Oakfield Road, Bristol, BS8 2BN, United Kingdom, [86]Bristol Dental School, University of Bristol, Lower Maudlin Street, Bristol BS1 2LY, United Kingdom, [87]Population Health Sciences, Bristol Medical School, University of Bristol, Oakfield Grove, Bristol, BS8 2BN, United Kingdom, [88]Saw Swee Hock School of Public Health, National University of Singapore and National University Health System, 117549, Singapore, [89]Center for Clinical Research and Prevention, Bispebjerg and Frederiksberg Hospital, Copenhagen, Denmark, [90]The Institute for Translational Genomics and Population Sciences, Department of Pediatrics, Lundquist Institute for Biomedical Innovations (Formerly LABioMed) at Harbor-UCLA Medical Center, Torrance, CA 90502, USA, [91]Center for Public Health Genomics, University of Virginia, Charlottesville, VA 22903 USA, [92]William Harvey Research Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, John Vane Science Centre, Charterhouse Square, London, EC1M 6BQ, UK, [93]NIHR Barts Cardiovascular Biomedical Research Centre, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, EC1M 6BQ, UK, [94]Novo Nordisk Foundation Center for Basic Metabolic Research, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark, [95]Center for Clinical Research and Prevention, Bispebjerg and Frederiksberg Hospital, Copenhagen, Denmark, Nordre Fasanvej 57, DK-2000 Frederiksberg, Denmark, [96]Division of Cardiovascular Medicine, Radcliffe Department of Medicine, John Radcliffe Hospital, University of Oxford, Oxford. UK. OX3 9DU, [97]Wellcome Centre for Human Genetics, University of Oxford, Oxford. UK. OX3 7BN, [98]Group of Genomics of Complex Diseases. Research Institute of Hospital de la Santa Creu i Sant Pau (IIB Sant Pau), Barcelona, Spain, [99]Cardiovascular Medicine Unit, Department of Medicine, Karolinska Institutet, Center for Molecular Medicine, Karolinska University Hospital, Stockholm, Sweden, [100]Department of Internal Medicine, Section Gerontology and Geriatrics, Leiden University Medical Center, Leiden, the Netherlands, [101]Institute for Genetic and Biomedical Research, Italian National Council of Research (IRGB CNR), Cagliari Italy, [102]Institute for Genetic and Biomedical Research, Italian National Council of Research (IRGB CNR), Lanusei, Italy, [103]University Center for Primary Care and Public Health, University of

Lausanne, Rte de la Corniche 10, Lausanne, 1010, Switzerland, [104]Swiss Institute of Bioinformatics, Lausanne, 1015, Switzerland, [105]Department of Medicine, Internal Medicine, Lausanne University Hospital and University of Lausanne, Rue du Bugnon 46, Lausanne, 1011, Switzerland, [106]Department of Epidemiology and Biostatistics, MRC-PHE Centre for Environment and Health, School of Public Health, Imperial College London, London, UK, [107]Dept of Cardiology, Leiden University Medical Center, Leiden, the Netherlands, [108]Dept of Internal Medicine, Section of Gerontology and Geriatrics, Leiden university Medical Center, Leiden, the Netherlands, [109]BHF Glasgow Cardiovascular Research Centre, Faculty of Medicine, Glasgow, United Kingdom, [110]Institute for Cardiogenetics, University of Lübeck, DZHK (German Research Centre for Cardiovascular Research), partner site Hamburg/Lübeck/Kiel, University Heart Center Lübeck, Lübeck and Charité – University Medicine Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, Institute for Dental and Craniofacial Sciences, Department of Periodontology and Synoptic Dentistry, Berlin, Germany, [111]Deutsches Herzzentrum München, Klinik für Herz- und Kreislauferkrankungen, Technische Universität München,Munich, Germany., [112]Deutsches Zentrum für Herz-Kreislauf-Forschung (DZHK) e.V., partner site Munich Heart Alliance,Munich, Germany., [113]Key Laboratory of Cardiovascular Epidemiology & 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 100037, China, [114]Lund University Diabetes Centre, Malmö, Sweden, [115]Institute of Genetic Epidemiology, Department of Genetics and Pharmacology, Medical University of Innsbruck, Innsbruck, Austria and German Chronic Kidney Disease study, [116]Institute for Medical Informatics, Statistics and Epidemiology, University of Leipzig, Haertelstrasse 16-18, 04107 Leipzig, Germany, [117]LIFE Research Centre for Civilization Diseases, University of Leipzig, Philipp-Rosenthal-Straße 27, 04103 Leipzig, Germany, [118]Radboud university medical center, Radboud Institute for Health Sciences, Nijmegen, The Netherlands, [119]Quantinuum Research LLC, Wayne, PA, 19087 USA, [120]Division of Statistical Genomics, Department of Genetics; Washington University School of Medicine; St. Louis, MO, USA, [121]Department of Epidemiology; University of Pittsburgh; Pittsburgh, PA, USA, [122]Institute for Biomedicine, Eurac Research, Affiliated Institute of the University of Lübeck, Via Galvani 31, 39100, Bolzano, Italy, [123]Department of Clinical Biochemistry, Lillebaelt Hospital, Vejle, Denmark, [124]Cardiovascular Health Research Unit, Department of Medicine, University of Washington, Seattle, 98101, USA, [125]The Charles Bronfman Institute for Personalized Medicine, Icahn School of Medicine at Mount Sinai, New York, NY, 10029, USA, [126]Department of Twin Research and Genetic Epidemiology, King's College London, London SE1 7EH, UK, [127]NIHR Biomedical Research Centre at Guy's and St Thomas' Foundation Trust, London SE1 9RT, UK, [128]University of Groningen, University Medical Center Groningen, Department of Cardiology, 9700RB Groningen, The Netherlands, [129]Institute of Cardiovascular Sciences, University College London, Gower Street, WC1E 6BT London, UK, [130]Department of Epidemiology and Public Health, University College London, 1-19 Torrington Place, WC1E 6BT London, United Kingdom, [131]Department of Psychiatry, University of Pennsylvania, Philadelphia, PA, 19104, USA, [132]Amsterdam UMC, Department of Epidemiology and Biostatistics, Amsterdam Public Health Research Institute, Amsterdam, 1081HV, the

Netherlands., [133]Leiden University Medical Center, Department of Cell and Chemical Biology, Leiden, 2333ZA, The Netherlands, [134]Montreal Heart Institute, 5000 Belanger Street, Montreal, Quebec, H1T 1C8, Canada, [135]Icelandic Heart Association, 201 Kopavogur, Iceland, [136]Department of Anthropology, University of Toronto at Mississauga, Mississauga, ON L5L 1C6, Canada, [137]Vth Department of Medicine, Medical Faculty Mannheim, Heidelberg University, 68167 Mannheim, Germany, [138]SYNLAB MVZ Humangenetik Mannheim GmbH, 68163 Mannheim, Germany, [139]CAS Key Laboratory of Nutrition, Metabolism and Food Safety, Shanghai Institute of Nutrition and Health, University of Chinese Academy of Sciences, Chinese Academy of Sciences, [140]Biomedical Technology Research Center, Tokushima Research Institute, Otsuka Pharmaceutical Co., Ltd., Tokushima, Japan, [141]Institute for Molecular Bioscience, The University of Queensland, Brisbane, Queensland 4072, Australia, [142]Institute for Advanced Research, Wenzhou Medical University, Wenzhou, Zhejiang 325027, China, [143]Nuffield Department of Population Health, University of Oxford, Oxford, United Kingdom, [144]Laboratory of Neurogenetics, National Institute on Aging, NIH, Bethesda MD, USA, [145]Data Tecnica International, Glen Echo MD, USA, [146]MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh, Western General Hospital, Crewe Road, Edinburgh, EH4 2XU, Scotland, [147]Institute for Medical Informatics, Biometrie and Epidemiology, University of Duisburg-Essen, Essen, Germany, [148]Department of Computer Science, University of Applied Sciences and Arts Dortmund, Emil-Figge-Str. 42, 44227 Dortmund, Germany, [149]Centre for Public Health, Queen's University of Belfast, Northern Ireland, [150]Department of Epidemiology and Biostatistics, School of Public Health, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China, [151]Genome Institute of Singapore, Agency for Science, Technology and Research, 138672, Singapore, [152]Public Health Informatics Unit, Department of Integrated Health Sciences, Nagoya University Graduate School of Medicine, Nagoya, 461-8673, Japan, [153]MRC Unit for Lifelong Health and Ageing at UCL, 1-19 Torrington Place, London, WC1E 7HB, United Kingdom, [154]Department of Pediatrics, Tampere University Hospital, Tampere 33521, Finland, [155]Department of Pediatrics, Faculty of Medicine and Health Technology, Tampere University, Tampere 33014, Finland, [156]Brown Foundation Institute of Molecular Medicine, McGovern Medical School, University of Texas Health Science Center at Houston, Houston TX 77030, USA, [157]CONACYT, Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán, Ciudad de Mexico, Mexico, [158]Departamento de Genómica Computacional, Instituto Nacional de Medicina Genómica, Ciudad de Mexico, Mexico, [159]Center for diabetes research, University of Bergen, Bergen, Norway, [160]Lund University Diabetes Center, Lunds University, Malmö, Sweden, [161]Genomic Research on Complex diseases (GRC Group), CSIR-Centre for Cellular and Molecular Biology, Hyderabad, Telangana, India, [162]University of Alabama at Birmingham, Epidemiology, School of Public Health, [163]Hunter Medical Research Institute, Newcastle, Australia, [164]Medical Research Institute, Kangbuk Samsung Hospital, Sungkyunkwan University School of Medicine, Seoul, 03181, Korea, [165]Department of Clinical Research Design & Evaluation, SAIHST, Sungkyunkwan University, Seoul, 06355, Korea, [166]Center for Cohort Studies, Total Healthcare Center, Kangbuk Samsung Hospital, Sungkyunkwan University School of Medicine, Seoul, 04514, Korea, [167]Department of Occupational and Environmental Medicine, Kangbuk Samsung Hospital, Sungkyunkwan University School of Medicine, Seoul, 03181,

Korea, [168]Centre for Global Health Research, Usher Institute, University of Edinburgh, Teviot Place, Edinburgh, EH8 9AG, Scotland, [169]Thurston Arthritis Research Center, University of North Carolina, Chapel Hill, North Carolina, USA, [170]Genome Institute of Singapore, Agency for Science, Technology and Research, Singapore, [171]Division of Biomedical Informatics and Personalized Medicine, Department of Medicine, Anschutz Medical Campus, University of Colorado, Denver, Aurora, CO 80045, USA, [172]Ocular Epidemiology, Singapore Eye Research Institute, Singapore National Eye Centre, 168751, Singapore, [173]Department of Ophthalmology, National University of Singapore and National University Health System, 119228, Singapore, [174]Genomics and Molecular Medicine Unit, CSIR-Institute of Genomics and Integrative Biology, New Delhi - 110020, India., [175]Academy of Scientific and Innovative Research, CSIR-Institute of Genomics and Integrative Biology Campus, New Delhi 110020, India., [176]Departments of Ophthalmology and Human Genetics, Radboud University Nijmegen Medical Center, Philips van Leydenlaan 15, Nijmegen, 6525 EX, the Netherlands, [177]Vanderbilt Epidemiology Center, Division of Epidemiology, Vanderbilt University Medical Center, [178]Department of Pediatrics, University of California San Francisco, Oakland, CA 94609 USA, [179]National Center for Global Health and Medicine, Tokyo, 1628655, Japan, [180]Department of Genetics, University of North Carolina, Chapel Hill, NC 27599 USA, [181]Department of Biostatistics and Epidemiology, University of Massachusetts-Amherst, Amherst, MA 01003 USA, [182]Department of Cardiovascular Sciences, University of Leicester, Leicester, UK, [183]NIHR Leicester Biomedical Research Centre, Glenfield Hospital, Leicester, UK, [184]Beijing Institute of Ophthalmology, Beijing Key Laboratory of Ophthalmology and Visual Sciences, Beijing Tongren Eye Center, Beijing Tongren Hospital, Capital Medical University, 17 Hougou Lane, Chong Wen Men, Beijing, 100005 , China, [185]Beijing Tongren Eye Center, Beijing Tongren Hospital, Capital Medical University, 1 Dong Jiao Min Xiang, Dong Cheng District, Beijing, 100730, China, [186]Institute of Genetics and Biophysics “Adriano Buzzati-Traverso" - CNR, Naples, Italy, [187]IRCCS Neuromed, Pozzilli, Isernia, Italy, [188]Division of Biostatistics, Washington University, St. Louis, MO 63110, [189]Division of Endocrinology and Metabolism, Tri-Service General Hospital Songshan Branch, Taipei, Taiwan, [190]Rush Alzheimer’s Disease Center, Rush University Medical Center, [191]Department of Neurological Sciences, Rush University Medical Center, [192]Department of Nephrology, University Hospital Regensburg, Regensburg, Germany, [193]Department of Medicine, Surgery and Health Sciences, University of Trieste, Strada di Fiume 447, 34149, Trieste, Italy, [194]Department of Molecular Epidemiology, German Institute of Human Nutrition Potsdam-Rehbruecke, Nuthetal, Germany, [195]German Center for Diabetes Research (DZD), München-Neuherberg, Germany, [196]Department of Epidemiology, School of Public Health, University of Michigan, Ann Arbor, MI 48109, [197]Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor, MI, 48104, [198]Department of Genetics and Bioinformatics, Dasman Diabetes Institute, Kuwait, [199]Department of Nutrition and Dietetics, School of Health Science and Education, Harokopio University of Athens, Athens, Greece, [200]Department of Population Science and Experimental Medicine, University College London, London, UK, [201]Clinical Division of Neurogeriatrics, Department of Neurology, Medical University of Graz, Graz, Austria, [202]Institute for Medical Informatics, Statistics and Documentation, Medical University of Graz, Graz, Austria, [203]Department of Bioinformatics and Genomics, University of North Carolina at Charlotte, NC 28223 USA, [204]Department of Internal

Medicine, University of Michigan Medical Center, [205]Institute for Maternal and Child Health, IRCCS Burlo Garofolo, Trieste, Italy, [206]Institute of Genetic and Biomedical Research, National Research Council of Italy, UOS of Sassari, Sassari, Italy, [207]University of Groningen, University Medical Center Groningen, Department of Epidemiology, Groningen, 9700 RB, the Netherlands, [208]Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, Turku, Finland, [209]Centre for Population Health Research, University of Turku and Turku University Hospital, Turku, Finland, [210]Sleep Medicine and Circadian Disorders, Brigham and Women's Hospital, Boston, Massachusetts 02115, USA, [211]Division of Sleep Medicine, Harvard Medical School, Boston, Massachusetts 02115, USA, [212]Central Diagnostics Laboratory, Division Laboratories, Pharmacy, and Biomedical genetics, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands, [213]Laboratory of Epidemiology and Population Science National Institute on Aging Intramural Research Program, NIH 251 Bayview Blvd, NIH Biomedical Research Center, NIA, Baltimore, MD 21224, [214]Interfaculty Institute for Genetics and Functional Genomics, Department of Functional Genomics, University of Greifswald and University Medicine Greifswald, Greifswald, Germany, [215]Oneomics. co. ltd. 2F, Soonchunhyang Mirai Medical Center 173, Buheuyng-ro, Bucheon-si Gyeonggi-do, 14585, Korea, [216]Department of Biomedical Science, Hallym University, Chuncheon, Gangwon-do 24252, Korea, [217]Centre for Bone and Arthritis Research, Department of Internal Medicine and Clinical Nutrition, Institute of Medicine, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden., [218]Bioinformatics Core Facility, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden, [219]Institute of Medical Informatics and Statistics, Kiel University, Kiel, Germany., [220]Institute of Translational Genomics, Helmholtz Zentrum München – German Research Center for Environmental Health, Neuherberg, Germany, [221]Wellcome Trust Sanger Institute, Hinxton, CB10 1SA, UK, [222]Wellcome Centre for Human Genetics, University of Oxford, Oxford, UK, [223]Oxford Centre for Diabetes Endocrinology and Metabolism, Oxford, UK, [224]Wellcome Sanger Institute, Hinxton, Cambridge, HH CB10 1 UK, [225]School of Medicine and Public Health, Faculty of Medicine and Health, University of Newcastle, Newcastle, New South Wales, 2308, Australia, [226]Center for Geriatrics and Gerontology, Division of Endocrinology and Metabolism, Department of Internal Medicine, Taichung Veterans General Hospital, Taichung, Taiwan, [227]School of Medicine, National Yang-Ming University, Taipei, Taiwan, [228]School of Medicine, National Defense Medical Center, Taipei, Taiwan, [229]Division of Endocrinology and Metabolism, Department of Internal Medicine, Taichung Veterans General Hospital, Taichung, Taiwan, [230]Department of Medicine, School of Medicine, National Yang-Ming University, Taipei, Taiwan, [231]Dept of Kinesiology, Université Laval, Québec, Canada, [232]Department of Clinical Chemistry, Fimlab Laboratories, Tampere 33520, Finland, [233]Department of Clinical Chemistry, Finnish Cardiovascular Research Center - Tampere, Faculty of Medicine and Health Technology, Tampere University, Tampere 33014, Finland, [234]Department of Cardiology, Heart Center, Tampere University Hospital, Tampere 33521, Finland, [235]Department of Cardiology, Finnish Cardiovascular Research Center - Tampere, Faculty of Medicine and Health Technology, Tampere University, Tampere 33014, Finland, [236]University of Queensland Diamantina Institute, Translational Research Institute, Kent St, Woolloongabba, Brisbane, QLD, 4102, Australia., [237]Novo Nordisk Foundation Center for Basic Metabolic Research, Faculty of Health and Medical Sciences, University of Copenhagen,

Copenhagen, Denmark, Department of Medicine, Bornholms Hospital, Rønne, Denmark, [238]School of Public Health, University of Alabama at Birmingham, [239]Cardiology, Division Heart & Lungs, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands, [240]Department of Population and Quantitative Health Sciences, Case Western Reserve University, Cleveland, OH, 44106, USA, [241]Division of Statistical Genomics, Department of Genetics, Washington University School of Medicine, St. Louis, MO, USA, [242]Department of Epidemiology - Erasmus MC - University Medical Center Rotterdam, Rotterdam, the Netherlands., [243]Ohio State University, Division of Endocrinology, Columbus OH 43210, USA, [244]University of Washington, Department of Epidemiology, Seattle WA 98195, USA, [245]George Washington University, School of Medicine and Health Sciences, Washington DC 20037, USA, [246]Department of Epidemiology, School of Public Health, Peking University Health Science Center, Beijing, China, [247]Institute for Laboratory Medicine, University Hospital Leipzig, Paul-List-Strasse 13/15, 04103 Leipzig, Germany, [248]Laboratory of Epidemiology and Population Sciences, National Institute on Aging, NIH, Baltimore, MD, 20892-9205, USA, [249]Centre for Population Health Research, University of Turku and Turku University Hospital, Finland, [250]Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, Finland, [251]Department of Clinical Physiology and Nuclear Medicine, Turku University Hospital, Turku, Finland, [252]Department of Environmental and Preventive Medicine, Jichi Medical University School of Medicine, Shimotsuke, 329-0498, Japan, [253]Centre for Population Health Sciences, Usher Institute, University of Edinburgh, Teviot Place, Edinburgh, EH8 9AG, Scotland, [254]Department of Functional Pathology, Shimane University School of Medicine, Izumo, 6938501, Japan, [255]Department of Pediatrics and Adolescent Medicine, Turku University Hospital and University of Turku, Turku, Finland, [256]Department of Physiology, University of Turku, Turku, Finland, [257]Faculty of Medicine, University of Split, Šoltanska 2, HR-21000, Split, Croatia, [258]Medical Department III – Endocrinology, Nephrology, Rheumatology, University of Leipzig Medical Center, Liebigstr. 21, 04103 Leipzig, Germany, [259]Department of Nutrition-Dietetics, Harokopio University, Eleftheriou Venizelou, Athens, 17676, Greece, [260]Department of Clinical Gene Therapy, Osaka University Graduate School of Medicine, Suita, 5650871, Japan, [261]Department of Geriatric and General Medicine, Osaka University Graduate School of Medicine, Suita, 5650871, Japan, [262]Department of Vascular Surgery, Division of Surgical Specialties, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands, [263]Corneal Dystrophy Research Institute, Department of Ophthalmology, Yonsei University College of Medicine, Seoul 03722, Korea, [264]Dept of Radiology and Nuclear Medicine, Erasmus MC - University Medical Center Rotterdam, Rotterdam, the Netherlands., [265]Julius Centre for Health Sciences and Primary Care, University Medical Centre Utrecht, 3584CG, the Netherlands, [266]Second Department of Cardiology, Medical School, National and Kapodistrian University of Athens, Attikon University Hospital, Athens, Greece, [267]Center for Vision Research, Department of Ophthalmology and The Westmead Institute, University of Sydney, Hawkesbury Rd, Sydney, New South Wales, 2145, Australia., [268]Menzies Institute for Medical Research, School of Medicine, University of Tasmania, Liverpool St, Hobart, Tasmania, 7000, Australia., [269]Centre for Eye Research Australia, University of Melbourne, Melbourne, Victoria, 3002, Australia, [270]Department of Clinical Physiology, Tampere University Hospital, Tampere 33521, Finland, [271]Department of Clinical Physiology, Finnish Cardiovascular Research Center -

Tampere, Faculty of Medicine and Health Technology, Tampere University, Tampere 33014, Finland, [272]Institute of Nutrition and Functional Foods (INAF), Université Laval, Québec, Canada, [273]Pennington Biomedical Research Center, Baton Rouge, LA 70808, USA, [274]Medical Department III – Endocrinology, Nephrology, Rheumatology, University of Leipzig Medical Center, Liebigstr. 18, 04103 Leipzig, Germany, [275]Medical School, Faculty of Health and Medical Sciences, University of Western Australia, Perth, Western Australia, 6000, Australia, [276]Institute of Epidemiology, Kiel University, Kiel, Germany, [277]Institute of Clinical Molecular Biology, Kiel University, Kiel, Germany, [278]Sahlgrenska University Hospital, Department of Drug Treatment, Gothenburg, Sweden, [279]Geriatric Medicine, Institute of Medicine, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden, [280]Department of Internal Medicine, EwhaWomans University School of Medicine, Seoul, Korea, [281]Division of Cancer Control and Population Sciences, UPMC Hillman Cancer Center, University of Pittsburgh, Pittsburgh, PA 15232, USA, [282]Department of Epidemiology, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA 15232, USA, [283]Health Services and Systems Research, Duke-NUS Medical School, 169857, Singapore, [284]Saw Swee Hock School of Public Health, National University of Singapore and National University Health System, [285]Department of Endocrinology and Metabolism, Kyung Hee University School of Medicine, Seoul 02447, Korea, [286]Institute for Community Medicine, University Medicine Greifswald, Germany, [287]Laboratory of Epidemiology and Population Science National Institute on Aging Intramural Research Program, NIH 251 Bayview Blvd, NIH Biomedical Research Center, Baltimore, MD 21224, [288]Paavo Nurmi Centre, Sports and Exercise Medicine Unit, Department of Physical Activity and Health, University of Turku, Turku, Finland, [289]University of Groningen, University Medical Center Groningen, Interdisciplinary Center Psychopathology and Emotion Regulation (ICPE), Groningen, 9700 RB, the Netherlands, [290]Institute of Molecular Genetics, National Research Council of Italy, Pavia, Italy, [291]Gottfried Schatz Research Center for Cell Signaling, Metabolism and Aging, Medical University of Graz, Graz, Austria, [292]Local Health Unit Toscana Centro, Firenze, Italy., [293]Department of Genetics and Bioinformatics, Dasman Diabetes Institute, Kuwait, [294]Institute of Nutritional Science, University of Potsdam, Nuthetal, Germany, [295]Institute for Maternal and Child Health IRCCS "Burlo Garofolo", Via dell'Istria 65/1, 34137, Trieste, Italy, [296]Dept of Nephrology, University Hospital Regensburg, Regensburg, Germany, [297]Dept of Nephrology, Diabetology, Rheumatology; Traunstein Hospital, Traunstein, Germany, [298]KfH Kidney Center Traunstein, Traunstein, Germany, [299]Center for Translational and Systems Neuroimmunology, Department of Neurology, Columbia University Medical Center, New York, NY, USA, [300]Program in Medical and Population Genetics, Broad Institute, Cambridge, MA, USA, [301]Medical School, National and Kapodistrian University Athens, 75 M. Assias Street, 115 27 Athens, Greece, [302]Dromokaiteio Psychiatric Hospital, 124 61 Athens, Greece, [303]Clinical Pharmacology, William Harvey Research Institute, Queen Mary University of London, London, EC1M 6BQ,UK, [304]Department of Ophthalmology, Medical Faculty Mannheim, Heidelberg University, Kutzerufer 1, Mannheim, 68167, Germany, [305]Institute of Clinical and Scientific Ophthalmology and Acupuncture Jonas & Panda, Heidelberg, Germany, [306]Harvard Medical School, Boston MA 02115, [307]Department of Human Genetics, David Geffen School of Medicine at UCLA, University of California, Los Angeles, CA, USA, [308]Unidad de Biología Molecular y Medicina Genómica, Instituto Nacional de Ciencias Médicas y Nutrición Salvador

Zubirán, Mexico 14080, Mexico., [309]Instituto de Investigaciones Biomédicas, UNAM, [310]Departamento de Endocrinología y Metabolismo, Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán, Mexico 14080, Mexico., [311]Department of Nutrition, Gillings School of Global Public Health, University of North Carolina, Chapel Hill, North Carolina, 27599 USA, [312]Carolina Population Center, University of North Carolina, Chapel Hill, North Carolina, 27516 USA, [313]USC–Office of Population Studies Foundation, University of San Carlos, Cebu City, 6000, Philippines, [314]Department of Anthropology, Sociology, and History, University of San Carlos, Cebu City, 6000 Philippines, [315]Department of Medicine, Faculty of Medicine, University of Kelaniya, Ragama, 11010, Sri Lanka, [316]Department of Public Health, Faculty of Medicine, University of Kelaniya, Ragama, 11010, Sri Lanka, [317]Children's Hospital Oakland Research Institute, Oakland, CA 94609 USA, [318]Institute of Biomedical Sciences, Academia Sinica, Taiwan, [319]Systems Genomics Laboratory, School of Biotechnology, Jawaharlal Nehru University, New Delhi - 110067, India, [320]Department of Medicine, University of Mississippi Medical Center, Jackson, MS, 39216, USA, [321]Department of Physiology and Biophysics, University of Mississippi Medical Center, Jackson, MS, 39216, USA, [322]Department of Medical Sciences, Uppsala University, Sweden, [323]Department of Paediatrics, Yong Loo Lin School of Medicine, National University of Singapore; and Khoo Teck Puat - National University Children's Medical Institute, National University Health System, [324]Department of Medicine, University of North Carolina, Chapel Hill, NC, USA, [325]Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina, Chapel Hill, North Carolina, USA, [326]Injury Prevention Research Center, University of North Carolina, Chapel Hill, North Carolina, USA, [327]Division of Physical Therapy, University of North Carolina, Chapel Hill, North Carolina, USA, [328]Department of Psychiatry, Amsterdam UMC, Vrije Universiteit Amsterdam, [329]Department of Biochemistry, College of Medicine, Ewha Womans University, Seoul 07804, Korea, [330]Faculty of Health and Medicine, University of Newcastle, Australia, [331]Washington University School of Medicine, Division of Biostatistics, [332]University of Kentucky, College of Public Health, [333]Institute of Cellular Medicine (Diabetes), The Medical School, Newcastle University, Framlington Place, Newcastle upon Tyne, NE2 4HH, [334]Department of Public Health Solutions, Finnish Institute for Health and Welfare, P.O. Box 30, FI-00271 Helsinki, Finland., [335]University of Helsinki and Department of Medicine, Helsinki University Central Hospital, P.O.Box 340, Haartmaninkatu 4, Helsinki, FI-00029, Finland., [336]Minerva Foundation Institute for Medical Research, Biomedicum 2U, Tukholmankatu 8, Helsinki, FI-00290, Finland., [337]Academy of Scientific and Innovative Research (AcSIR), New Delhi, India, [338]Diabetology Research Centre, KEM Hospital and Research Centre, Pune, Maharashtra, India, [339]Programs in Metabolism and Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, MA, USA, [340]Diabetes Unit and Center for Genomic Medicine, Massachusetts General Hospital, Boston, MA, USA10, [341]Harvard Medical School, Boston, Massachusetts, USA, [342]Unidad de Biología Molecular y Medicina Genómica, Instituto de Investigaciones Bimédicas UNAM/ Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán, Mexico City, [343]Dirección de Nutrición and Unidad de Estudios de Enfermedades Metabólicas, Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán, Mexico City, [344]InstitutoNacional de Salud Publica y Centro de Estudios en Diabetes, Mexico, [345]Instituto Nacional de Medicina Genómica, Mexico, [346]Human Genetics Center,

School of Public Health, University of Texas Health Science Center at Houston, Houston TX 77030, USA, [347]Yong Loo Lin School of Medicine, National University of Singapore and National University Health System, 119228, Singapore, [348]Kurume University School of Medicine, Kurume, 830-0011, Japan, [349]Genetics, Merck Sharp & Dohme Corp., Kenilworth, NJ, 07033, US, [350]Oxford Centre for Diabetes, Endocrinology & Metabolism, University of Oxford, UK, [351]Population Health and Genomics, University of Dundee, Ninwells Hospital and Medical School, Dundee, DD1 9SY, UK, [352]Intramural Research Program, National Institute on Aging, 3001 S. Hanover St., Baltimore, MD 21225, [353]The Eye Hospital, School of Ophthalmology & Optometry, Wenzhou Medical University, Wenzhou, Zhejiang 325027, China, [354]Shanghai Institute of Nutrition and Health University of Chinese Academy of Sciences, Chinese Academy of Sciences, [355]Synlab Academy, SYNLAB Holding Deutschland GmbH, Mannheim and Augsburg, Germany, [356]Clinical Institute of Medical and Chemical Laboratory Diagnostics, Medical University of Graz, Austria, [357]Unidad de Investigacion Medica en Bioquimica, Hospital de Especialidades, Centro Medico Nacional Siglo XXI, Instituto Mexicano del Seguro Social, Mexico City, Mexico, [358]Faculty of Medicine, University of Iceland, 101 Reykjavik, Iceland, [359]Department of Medicine, Faculty of Medicine, Université de Montréal, 2900 Edouard Montpetit Blvd, Montreal, Quebec, H3T 1J4, Canada, [360]Leiden University Medical Center, Department of Biomedical Data Sciences, Section Molecular Epidemiology, Leiden, 2333ZA, The Netherlands, [361]Amsterdam UMC, Department of General Practice and Elderly Care, Amsterdam Public Health Research Institute, Amsterdam, 1081HV, The Netherlands, [362]Department of Genetics, University of Pennsylvania, Philadelphia, PA, 19104, USA, [363]Department of Surgery, University of Pennsylvania, Philadelphia, PA, 19104, USA, [364]Corporal Michael Crescenz VA Medical Center, Philadelphia, Pennsylvania, PA, 19104, USA, [365]Institute of Social and Economic Research, University of Essex, Wivenhoe Park, CO4 3SQ, United Kingdom, [366]The Mindich Child Health and Development Institute, Icahn School of Medicine at Mount Sinai, New York, NY, 10029, USA, [367]Cardiovascular Health Research Unit, Departments of Medicine, Epidemiology and Health Services, University of Washington, Seattle, 98101, WA, USA, [368]Kaiser Permanent Washington Health Research Institute, Seattle, 98101, WA, USA, [369]Institute of Regional Health Research, University of Southern Denmark, Odense, Denmark, [370]Danish Aging Research Center, University of Southern Denmark; Odense C, Denmark, [371]Public Health, Faculty of Medicine, University of Helsinki, Finland, [372]Broad Institute of MIT and Harvard, Cambridge, MA, [373]Center for Applied Genomics, Children's Hospital of Philadelphia, Philadelphia, PA, 19104 USA, [374]Department of Pediatrics, The University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, 19104 USA, [375]Division of Human Genetics, Children’s Hospital of Philadelphia, Philadelphia, PA, 19104 USA, [376]Department of Genetics, University of Pennsylvania, Philadelphia, PA, 19104 USA, [377]School of Medicine, Southern University of Science and Technology, Shenzhen, China, [378]Institute for Cardiogenetics, University of Lübeck, DZHK (German Research Centre for Cardiovascular Research), partner site Hamburg/Lübeck/Kiel, and University Heart Center Lübeck, Lübeck, Germany, [379]Department of Clinical Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark, [380]Netherlands Heart Institute, Utrecht, the Netherlands, [381]Division of Cardiology, Department of Medicine, Massachusetts General Hospital, Boston, Massachusetts, USA, [382]Program of Medical and Population Genetics,

Broad Institute, Cambridge, Massachusetts, USA, [383]Center for Genomic Medicine, Massachusetts General Hospital, Boston, Massachusetts, USA, [384]Department of Medicine, Harvard Medical School, Boston, Massachusetts, USA, [385]Northern Finland Birth Cohorts, Infrastructure for population studies, Faculty of Medicine, University of Oulu, Oulu, Finland, [386]Center for Life Course Health Research, Faculty of Medicine, University of Oulu, Oulu, Fin, [387]Biocenter of Oulu, University of Oulu, Oulu, Finl, [388]Institute for Genetic and Biomedical Research, Italian National Council of Research (IRGB CNR), Cagliari, Italy, [389]University of Sassari, Sassari, Italy, [390]Department of Clinical Epidemiology, Leiden University Medical Center, Leiden, the Netherlands, [391]Department of Public Health and Primary Care, Leiden University Medical Center, Leiden, the Netherlands, [392]Department of Internal Medicine, Division of Endocrinology, Leiden University Medical Center, Leiden, the Netherlands, [393]Einthoven Laboratory for Experimental Vascular Medicine, Leiden University Medical Center, Leiden, the Netherlands, [394]Department of Human Genetics, Leiden University Medical Center, Leiden, the Netherlands, [395]Population Health Research Institute, St George's, University of London, London SW17 0RE, UK, [396]National Heart and Lung Institute, Imperial College London, London, W2 1PG, UK, [397]School of Public Health, Imperial College London, London, W2 1PG, UK, [398]OCDEM, University of Oxford, Churchill Hospital, Oxford OX3 7LE, UK, [399]NIHR Oxford Biomedical Research Centre, Churchill Hospital, Oxford, UK, [400]Ophthalmology & Visual Sciences Academic Clinical Program (Eye ACP), Duke-NUS Medical School, 169857, Singapore, [401]DZHK (German Centre for Cardiovascular Research), Munich Heart Alliance partner site, Munich, Germany., [402]German Center for Diabetes Research (DZD), Neuherberg, Germany., [403]University of Exeter Medical School, University of Exeter, Exeter, EX2 5DW, UK, [404]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden, [405]Netherlands Twin Register, Department of Biological Psychology, Behavioral and Movement Sciences, Vrije Universiteit Amsterdam, [406]Amsterdam Reproduction & Development research institute, VU medical center Amsterdam, [407]Framingham Heart Study, National Heart, Lung, and Blood Institute, US National Institutes of Health, Bethesda, MD, USA., [408]Department of Genetics, School of Medicine, Mashhad University of Medical Sciences, Mashhad, Iran., [409]Department of Genetics, Shanghai-MOST Key Laboratory of Health and Disease Genomics, Chinese National Human Genome Center at Shanghai, Shanghai, 201203 China, [410]TUM School of Medicine, Technical University of Munich and Klinikum Rechts der Isar, Munich, Germany, [411]Institute of Clinical Medicine, Internal Medicine, University of Eastern Finland and Kuopio University Hospital, [412]Department of Medicine, Division of Cardiovascular Medicine, Stanford University School of Medicine, Stanford, CA 94305, [413]Stanford Cardiovascular Institute, Stanford University, Stanford, CA 94305, [414]Stanford Diabetes Research Center, Stanford University, Stanford, CA 94305, [415]Regeneron Pharmaceuticals, Tarrytown, NY, USA, [416]Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore 308232, Singapore, [417]Imperial College Healthcare NHS Trust, Imperial College London, London W12 0HS, UK, [418]MRC-PHE Centre for Environment and Health, Imperial College London, London W2 1PG, UK, [419]National Heart and Lung Institute, Imperial College London, London W12 0NN, UK, [420]Institute for Minority Health Research, University of Illinois College of Medicine, Chicago, Illinois, [421]Department of Biostatistics, Harvard T.H. Chan School of Public Health, 677 Huntington Avenue, Boston, MA,

02115, USA, [422]QIMR Berghofer Medical Research Institute, 300 Herston Road, Brisbane, Queensland 4006, Australia, [423]Center for Non-Communicable Diseases, Karachi, Sindh, Pakistan & Faisalabad Institute of Cardiology, Faislabad, Pakistan, [424]Department of Medicine, Columbia University Irving Medical Center, New York, NY, USA, [425]Department of Cardiology, Columbia University Irving Medical Center, New York, NY, USA, [426]Clinical Trial Service Unit and Epidemiological Studies Unit, Nuffield Department of Population Health, University of Oxford, Oxford OX3 7LF, UK, [427]Medical Research Council Population Health Research Unit, Nuffield Department of Population Health, University of Oxford, Oxford OX3 7LF, UK, [428]Big Data Institute, University of Oxford, Oxford OX3 7LF, UK, [429]National Institute for Health Research Oxford Biomedical Research Centre, Oxford University Hospitals, Oxford, UK, [430]Aberdeen Centre for Health Data Science, 1:042 Polwarth Building School of Medicine, Medical Science and Nutrition University of Aberdeen Foresterhill Aberdeen AB25 2ZD, [431]Division of Population Health and Genomics, Ninewells Hospital and Medical School, University of Dundee, Dundee DD1 9SY, United Kingdom, [432]Biomedical and Translational Informatics, Geisinger Health, Danville, PA 17822, [433]Regeneron Pharmaceuticals, Tarrytown, NY, USA., [434]Harvard Medical School, Boston, MA 02115, [435]Department of Biomedical Informatics, Vanderbilt University Medical Center, Nashville, TN, [436]Departments of Medicine (Medical Genetics) and Genome Sciences, University of Washington, [437]Center for Autoimmune Genomics and Etiology, Cincinnati Children's Hospital Medical Center (CCHMC), Cincinnati, OH, USA., [438]Division of Endocrinology, Metabolism, and Molecular Medicine, Department of Medicine, Northwestern University, Feinberg School of Medicine, Chicago, IL 60618, USA, [439]Department of Anthropology, Northwestern University, Evanston, IL 60208, USA, [440]Center for Genetic Medicine, Northwestern University, Feinberg School of Medicine, Chicago, IL 60618, USA, [441]HUNT Research Centre, Department of Public Health and Nursing, NTNU, Norwegian University of Science and Technology, Levanger, 7600 Norway, [442]Department of Medicine, Levanger Hospital, Nord-Trøndelag Hospital Trust, Levanger, 7600 Norway, [443]Department of Endocrinology, St. Olavs Hospital, Trondheim University Hospital, Trondheim, Norway, [444]RIKEN Center for Integrative Medical Sciences, [445]Laboratory of Complex Trait Genomics, Department of Computational Biology and Medical Sciences, Graduate School of Frontier Sciences, The University of Tokyo, Tokyo, Japan, [446]Laboratory of Statistical Immunology, WPI Immunology Frontier Research Center, Osaka University, Osaka, Japan, [447]Integrated Frontier Research for Medical Science Division, Institute for Open and Transdisciplinary Research Initiatives, Osaka University, Osaka, Japan, [448]Division of Molecular Pathology, Institute of Medical Science, The University of Tokyo, Tokyo, Japan, [449]Faculty of Medicine, University of Iceland, Sæmundargötu 2, Reykjavik, 102, Iceland, [450]VA Boston Healthcare System, Boston, MA, USA, [451]VA Informatics and Computing Infrastructure, VA Salt Lake City Health Care System, Salt Lake City, UT, USA, [452]University of Massachusetts, Boston, MA, USA, [453]Department of Medicine, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, USA, [454]Cardiovascular Institute, Stanford University School of Medicine, Stanford, California, USA, [455]Corporal Michael J. Crescenz VA Medical Center, Philadelphia, PA, USA, [456]Department of Medicine, Brigham Women's Hospital, Boston, MA, USA, [457]Atlanta VA Medical Center, Atlanta, GA, USA, [458]Division of Cardiology, Emory University School of Medicine, Atlanta, GA, USA, [459]School of Electrical and Information

Engineering, University of the Witwatersrand, Johannesburg, South Africa, [460]Division of Human Genetics, National Health Laboratory Service and School of Pathology, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa., [461]School of Basic and Medical Biosciences, Faculty of Life Sciences and Medicine, King's College London, London, UK, [462]Departments of Pediatrics and Genetics, Harvard Medical School, Boston, MA, USA, [463]Center for Genomic Medicine, Department of Medicine, Massachusetts General Hospital, Boston, Massachusetts, USA, [464]Cardiology Division, Massachusetts General Hospital, Harvard Medical School, Boston, MA, [465]Department of Medicine, Massachusetts General Hospital, Harvard Medical School, Boston, MA, [466]Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, Cambridge, MA, [467]Cardiovascular Research Center and Center for Genomic Medicine, Massachusetts General Hospital, Boston, MA, [468]Department of Biostatistics, Boston University School of Public Health, Boston, MA, [469]Centre for Genetics and Genomics Versus Arthritis, Centre for Musculoskeletal Research, Division of Musculoskeletal and Dermatological Sciences, The University of Manchester, Manchester, UK, [470]Princess Al-Jawhara Al-Brahim Centre of Excellence in Research of Hereditary Disorders (PACER-HD), King Abdulaziz University, Jeddah, Saudi Arabia, [471]Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI 48109, USA, [472]Department of Human Genetics, University of Michigan, Ann Arbor, MI 48019, USA

[‡]These authors contributed equally to this work (as co-second authors)

[†]These authors jointly supervised this work

[#]Present Address: Genentech, 1 DNA Way, South San Francisco, CA 94080

A full list of consortia members can be found in the supplementary information

Elevated blood lipid levels are heritable risk factors of cardiovascular disease with varying prevalence worldwide due to differing dietary patterns and medication use[1]. Despite advances in prevention and treatment, particularly through the lowering of low-density lipoprotein cholesterol levels[2], heart disease remains the leading cause of death worldwide[3]. Genome-wide association studies (GWAS) of blood lipid levels have led to important biological and clinical insights, as well as new drug targets, for cardiovascular disease. However, most previous GWAS[4-23] have been conducted in European ancestry populations and may have missed genetic variants contributing to lipid level variation in other ancestry groups due to differences in allele frequencies, effect sizes, and linkage-disequilibrium (LD) patterns[24]. Here we conduct a multi-ancestry genome-wide genetic discovery meta-analysis of lipid levels in ~1.65 million individuals, including 350,000 of non-European ancestries. We quantify the gain in studying non-European ancestries and provide evidence to support expanding recruitment into new ancestries even with relatively smaller sample sizes. We find that increasing diversity rather than studying additional European ancestry individuals results in substantial improvements in fine-mapping functional variants and portability of polygenic prediction, with modest gains in the number of discovered loci and ancestry-specific variants. As GWAS expands its emphasis beyond identifying genes and fundamental biology towards using genetic variants for preventive and precision medicine[25], we anticipate that increased participant diversity will lead to more accurate and equitable[26] application of polygenic scores in clinical practice.

Main Text

The Global Lipids Genetics Consortium aggregated GWAS results from 1,654,960 individuals from 201 primary studies representing five genetic ancestry groups: Admixed African or African (AdmAFR, N=99.4k, 6.0% of sample), East Asian (EAS, N=146.5k, 8.9%), European (EUR, N=1.32m, 79.8%), Hispanic (HIS, N=48.1k, 2.9%), and South Asian (SAS, N=41.0k, 2.5%) (**Table 1**, **Supplementary Table 1**, **Supplementary Figure 1**). We performed GWAS for five blood lipid traits: low-density lipoprotein cholesterol (LDL-C), high-density lipoprotein cholesterol (HDL-C), triglycerides (TG), total cholesterol (TC), and non-high-density lipoprotein cholesterol (nonHDL-C). Of the 91 million variants imputed from the Haplotype Reference Consortium or 1000 Genomes Phase 3 that successfully passed variant-level QC, 52 million

variants were present in at least two cohorts and had sufficient minor allele counts (> 30 in the meta-analysis) to be evaluated as a potential index variant.

Ancestry-specific genetic discovery

We first quantified the number of genome-wide significant loci identified in at least one of the five ancestry-specific meta-analyses. We found 773 lipid-associated genomic regions containing 1,765 distinct index variants that reached genome-wide significance (p-value < $5x10^{-8}$, ±500 kb, **Supplementary Tables 2-3**, **Supplementary Figures 2-3**) for at least one ancestry group and lipid trait. Of these regions, 237 were novel based on the most-significant index variant in each region being >500 kb from variants previously reported as associated with any of the five lipid traits[4-23,27]. Of these loci, 76% were identified only in the European ancestry-specific analyses (N~1.3m, 80% of sample). Of the non-European ancestries, the African ancestry GWAS (N~99k, primarily African American) identified more ancestry-specific loci (15 unique to AdmAFR) than any other non-European ancestry group (six loci unique to EAS, six to HIS, one to SAS). The difference is likely attributable to allele frequencies being most different between African and European ancestry populations (**Figure 1a-d**) and to African populations having greater genetic diversity[28].

Trans-ancestry genetic discovery

We next performed trans-ancestry meta-analyses using the meta-regression approach implemented in MR-MEGA[30] to account for heterogeneity in variant effect sizes on lipids between ancestry groups. A total of 1,750 index variants at 923 loci (±500 kb regions) reached genome-wide significance for at least one lipid trait. These included 168 regions not identified by ancestry-specific analysis, 120 (71%) of which were novel (**Supplementary Tables 4-5**, **Supplementary Figure 4, Extended Data Figure 1**). Almost all (98%) index variants from the ancestry-specific analysis remained significant (p-value<$5x10^{-8}$) after meta-analysis across all ancestry groups, although fifteen AdmAFR, nine EAS, three HIS, and one SAS index variants from ancestry-specific analysis did not (trans-ancestry p-value $7.7x10^{-6}$ to $5.9x10^{-8}$, **Supplementary Figure 5**, **Supplementary Note**). In total, we identified 941 lipid-associated loci including 355 novel loci from either single- or trans-ancestry analyses.

Next, we compared the number of loci identified per 100,000 participants in each ancestry group and the combined dataset (**Figure 1e**). African and Hispanic ancestry-specific analyses identified the most loci per genotyped individual, perhaps due to African ancestry and/or increased genetic diversity. European and trans-ancestry analyses identified slightly fewer loci per 100,000 individuals, likely reflecting a slight reduction in the benefit from new samples added to very large sample sizes (>1m). For the genome-wide significant variants discovered in each ancestry, we estimated the proportion of ancestry-enriched variants by enumerating the number of other ancestries with sufficient power to detect association (range 0 to 4). We estimated the power for discovery of each variant by assuming an equivalent discovery sample size in the other ancestries, fixed effect size, and observed allele frequencies from the other ancestries (**Figure 1f**). To allow for comparison at similar sample sizes across ancestry groups, we selected European ancestry index variants identified from a meta-analysis of ~100,000 individuals subsampled from the present study. African ancestry index-variants were most ancestry-enriched, with only 61% of index variants demonstrating sufficient power in at least one other ancestry group (equal N, power>80% to reach alpha=$5x10^{-8}$), likely due to population-enriched allele frequencies. In comparison, 88% of South Asian index variants had estimated power >80% in at least one other ancestry.

Finally, we found that both the number of identified variants and the mean observed chi-squared values from genome-wide lipid association tests were approximately linearly related to meta-analysis sample size across ancestries (**Supplementary Table 6**, **Extended Data Figure 2**). However, in the European ancestry group the incremental increase in either the number of loci or chi-squared value was slightly attenuated at the largest sample sizes. Taken together, these results suggest that once sufficiently well-powered GWAS sample sizes are reached within a given ancestry group, assembling large sample sizes of other under-represented groups will modestly enhance variant discovery relative to increasing the sample size of the dominant ancestry.

Comparison of effects across ancestries

Differences in association signals across ancestries despite similar sample sizes could be due to variation in allele frequencies and/or effect sizes. This could reflect differing patterns of LD with

the underlying causal variant or an interaction with an environmental risk factor whose prevalence varies by ancestry and/or geography. We found that effect size estimates of individual variants were largely similar based on pairwise comparison between ancestries (r^2=0.93 for variants with p-value<$5x10^{-8}$) (**Extended Data Figure 3**, **Supplementary Table 7**, **Supplementary Figure 6**). We additionally tested for genome-level differences in effect size correlation between East Asian, European, and South Asian ancestry groups using Popcorn[29], which were not significantly different from 1 (p-value>0.05, **Supplementary Figures 7** and **8**). We tested for differences in genetic correlation between Admixed African and European ancestries in the UK Biobank and Million Veteran Program (MVP) using bivariate GREML[30,31] as the Popcorn method does not account for long-range LD in admixed populations. Genetic correlation between Admixed African and European ancestries for HDL-C (r=0.84) was not significantly different from 1 in the UK Biobank (possibly due to relatively small numbers of African ancestry individuals), while correlations for the other traits ranged from 0.52-0.60 in UK Biobank and 0.47-0.69 in MVP (**Supplementary Table 8**). These results indicate moderately high correlation in lipid effect sizes across ancestry groups when considering all genome-wide variants.

Of the 2,286 variants that reached genome-wide significance in the trans-ancestry meta-analysis across all five lipid traits, 159 (7%) showed significant heterogeneity of effect size due to ancestry (p-value<$2.2x10^{-5}$; Bonferroni correction for 2,286 variants, **Supplementary Table 5**). Of these 159, 31 showed the largest effect in African ancestry analyses, 24 in East Asian, 67 in European, 20 in Hispanic, and 17 in South Asian. Only 49 (2%) of these variants from trans-ancestry meta-analysis showed significant residual heterogeneity not due to ancestry, which may be attributable to differences in ascertainment or analysis strategy between cohorts (**Supplementary Table 5**), suggesting cohort-related factors are a less important driver of heterogeneity than genetic ancestry.

Trans-ancestry analyses aid fine-mapping

We next assessed whether trans-ancestry fine-mapping narrowed the set of likely causal variants at each of the independent trans-ancestry association signals (LD r^2<0.7), assuming one shared causal variant per ±500 kb region (**Supplementary Table 9**). 19% of the association signals had

only one variant in the 99% credible set and 55% (816/1,486) had ≤10. In contrast, 5% (73/1486) had >100. Of the 407 variants with >90% posterior probability of being the causal variant at a locus in the trans-ancestry meta-analysis, 56 (14%) were missense variants, 7 (2%) were splice-region variants, and 4 (1%) were stop-gain variants (*CD36, HBB, ANGPTL8, PDE3B*). (**Supplementary Tables 10-12**).

The median number of variants in 99% credible sets from European ancestry analysis was 13; this was reduced to 8 in the trans-ancestry analysis. Of 1,486 association signals, 825 (56%) had reduced credible set size in the trans-ancestry analysis. At these 825 loci, the number of variants in the trans-ancestry credible sets were reduced by 40% relative to the minimum credible set size in either Admixed African (the most genetically diverse group) or European ancestry analyses (**Extended Data Figure 4**). We estimate that increasing the sample size of European ancestry samples to that of the trans-ancestry analysis would yield a 20% reduction in credible set size, approximately half of the 40% reduction observed in trans-ancestry analysis. This suggests that sample size differences alone do not explain the reduction, rather differences in LD patterns and effect sizes across ancestries likely contribute to the improved fine-mapping (**Supplementary Note**). For example, rs900776, an intronic variant in the *DMTN* region with many high LD variants in the European ancestry group, has a posterior probability of being causal of 0.86 in the African ancestry derived credible sets, >0.99 in the trans-ancestry analysis, but only 0.51 in the European ancestry-specific analysis (**Figure 2**).

Trans-ancestry PRS are most predictive

We evaluated the potential of polygenic scores to predict elevated LDL-C, a major causal risk factor of CAD, in diverse ancestry groups. We created three non-overlapping datasets to separately: i) perform ancestry-specific or trans-ancestry GWAS to estimate variant effect sizes, ii) optimize risk score parameters, and iii) evaluate the utility of the resulting scores. For each ancestry-specific or trans-ancestry GWAS we created multiple polygenic score weights -- either genome-wide with PRS-CS[32] or using pruning and thresholding to select independent variants. We tested each score in the optimizing dataset, which was matched for ancestry to the GWAS (AdmAFR, EAS, EUR, SAS, ALL from UK biobank or HIS from Michigan Genomics Initiative (MGI), **Extended Data Figures 5** and **6**, **Supplementary Tables 13-15**). The top-performing

score from each GWAS was selected: PRS-CS for East Asian ancestry, European ancestry, and European ancestry 2010 scores from a previous GLGC GWAS[4], and an optimized pruning and threshold-based score for all others. We then evaluated the polygenic scores in 8 cohorts of individuals (N=295,577, **Supplementary Table 16**), not included in the discovery GWAS, from 6 ancestral groups: East Asian (146,477), European American (85,571), African American (21,730), African (2,452 East Africa, 4,972 South Africa, 7,309 West Africa), South Asian (15,242), Hispanic American (7,669), and Asian American (4,155).

The polygenic score developed from trans-ancestry meta-analysis consistently showed the best or near-best performance in each group tested, with improved or comparable prediction relative to ancestry-matched scores (adjusted R^2 ~ 0.10-0.16, **Figure 3, Supplementary Table 17, Extended Data Figure 7**). This observation was especially evident for ancestries with smaller GWAS sample sizes, as was the case for HIS and SAS. For African Americans in MGI and MVP, polygenic prediction was similar for individuals with different levels of recent African ancestry admixture (**Extended Data Figure 8**) and reached the level of prediction observed for European ancestry individuals from the same dataset. The increase in LDL-C per each standard deviation increase in the polygenic score was also similar between ancestry groups in MVP: 13.2±0.22 mg/dL for African American, 8.9±0.47 mg/dL for Asian (EAS/SAS), 10.5±0.10 mg/dL for European, and 10.6±0.32 mg/dL for Hispanic ancestry individuals. We repeated the evaluation of trans-ancestry vs single-ancestry polygenic scores with a set GWAS with sample size of ~100k individuals and with fixed methodology; results were consistent with those from the full dataset (**Figure 3b**, **Supplementary Figure 9**). Thus, polygenic prediction for LDL-C in all ancestries appears to benefit the most from adding samples of diverse ancestries once relatively large numbers of European ancestry individuals have already been included. Additional studies are needed to determine if this applies to other phenotypes with different genetic architectures and heritabilities.

Discussion

Genome-wide discovery for blood lipid traits based on ~1.65 million individuals from five ancestry groups confirmed that the contributions of common genetic variation to blood lipids are largely similar across diverse populations. First, we found that the number of significant loci

relative to sample size was similar within each ancestry group, and approximately linearly related to sample size, with a small increase in ancestry-specific variants observed in African ancestry cohorts relative to the others. Second, we demonstrated that inclusion of additional ancestries through trans-ancestry fine-mapping reduces the set of candidate causal variants in credible sets and does so more rapidly than in single-ancestry analysis. Trans-ancestry GWAS should therefore facilitate identification of effector genes at GWAS loci and allow for accelerated biological insight and identification of potential drug targets. Third, we found that a polygenic score derived from ~88k African ancestry and ~830k European ancestry individuals was correlated with observed lipid levels among individuals with admixed African ancestry as well as among individuals with European ancestry. We hypothesize that the inclusion of African ancestry individuals in the GWAS yields improvement in polygenic prediction performance through the general fine-mapping of loci and the improved prioritization of trans-ancestry causal variants. Fourth, and perhaps most important, the trans-ancestry score was generally most informative across all major population groups examined. This provides useful information for other genetic discovery efforts and investigations of the utility of the polygenic scores in diverse populations.

Generalizability of these findings regarding portability of polygenic scores from the trans-ancestry meta-analysis to other traits may depend on the heritability, degree of polygenicity, level of genetic correlation, allele frequencies of causal variants across ancestry groups, gene-environment interactions, and representation of diverse populations in the GWAS[33,34]. While many traits show a high degree of shared genetic correlation across ancestries[31,35,36] others have distinct genetic variants with large effects that are more common in specific ancestry groups[33] which may limit the utility of trans-ancestry polygenic scores for particular phenotypes in some ancestries.

The benefits from genetic discovery efforts as GWAS sample sizes increase will likely not be measured just by the number of loci discovered. Rather, the focus will increasingly turn to improving our understanding of the biology at established loci, identifying potential therapeutic targets, and efficiently identifying individuals at high-risk of adverse health outcomes across population groups without exacerbating existing health disparities. Considering the results

presented here, and those of related studies[37-39], we believe future genetic studies will benefit substantially from meta-analysis across participants of diverse ancestries. Further gains in the depth and number of sequenced individuals of diverse ancestries[40,41] may additionally improve discovery of novel variants and loci in diverse cohorts, particularly variants absent from arrays and imputation reference panels. Our results suggest that diversifying the populations under study, rather than simply increasing the sample size, is now the single most efficient approach to achieving these goals, at least for blood lipids and likely for tightly related downstream adverse health outcomes such as cardiovascular disease. However, if costs for recruitment of diverse populations are higher than recruitment of individuals from previously studied ancestry groups, and total number of genome-wide significant index variants is the goal, then continued low-cost recruitment of majority ancestry groups is expected to still provide some benefit. Taken together, our results also strongly support ongoing and future large-scale recruitment efforts targeted at the enrollment and DNA collection of non-European ancestry participants. Geneticists and those responsible for cohort development must continue diversifying genetic discovery datasets, while increasing sample size in a cost-effective manner, to ensure genetic studies reduce rather than exacerbate existing health inequities across race, ancestry, geographic region, and nationality.

References

1 Taddei, C. *et al.* Repositioning of the global epicentre of non-optimal cholesterol. *Nature* **582**, 73-77, doi:10.1038/s41586-020-2338-1 (2020).

2 Ference, B. A. *et al.* Low-density lipoproteins cause atherosclerotic cardiovascular disease. 1. Evidence from genetic, epidemiologic, and clinical studies. A consensus statement from the European Atherosclerosis Society Consensus Panel. *Eur Heart J* **38**, 2459-2472, doi:10.1093/eurheartj/ehx144 (2017).

3 Roth, G. A. *et al.* Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980-2017: a systematic analysis for the Global Burden of Disease Study 2017. *The Lancet* **392**, 1736-1788, doi:10.1016/S0140-6736(18)32203-7 (2018).

4 Teslovich, T. M. *et al.* Biological, clinical and population relevance of 95 loci for blood lipids. *Nature* **466**, 707-713, doi:10.1038/nature09270 (2010).

5 Willer, C. J. *et al.* Discovery and refinement of loci associated with lipid levels. *Nature genetics* **45**, 1274-1283, doi:10.1038/ng.2797 (2013).

6 Liu, D. J. *et al.* Exome-wide association study of plasma lipids in >300,000 individuals. *Nature genetics* **49**, 1758-1766, doi:10.1038/ng.3977 (2017).

7 Lu, X. *et al.* Exome chip meta-analysis identifies novel loci and East Asian-specific coding variants that contribute to lipid levels and coronary artery disease. *Nature genetics* **49**, 1722-1730, doi:10.1038/ng.3978 (2017).

8 Kathiresan, S. *et al.* A genome-wide association study for blood lipid phenotypes in the Framingham Heart Study. *BMC Med Genet* **8 Suppl 1**, S17-S17, doi:10.1186/1471-2350-8-S1-S17 (2007).

9 Kathiresan, S. *et al.* Polymorphisms Associated with Cholesterol and Risk of Cardiovascular Events. *New England Journal of Medicine* **358**, 1240-1249, doi:10.1056/NEJMoa0706728 (2008).

10 Peloso, G. M. *et al.* Association of low-frequency and rare coding-sequence variants with blood lipids and coronary heart disease in 56,000 whites and blacks. *American journal of human genetics* **94**, 223-232, doi:10.1016/j.ajhg.2014.01.009 (2014).

11 Hoffmann, T. J. *et al.* A large electronic-health-record-based genome-wide study of serum lipids. *Nature genetics* **50**, 401-413, doi:10.1038/s41588-018-0064-5 (2018).

12 Surakka, I. *et al.* The impact of low-frequency and rare variants on lipid levels. *Nature genetics* **47**, 589-597, doi:10.1038/ng.3300 (2015).

13 Klarin, D. *et al.* Genetics of blood lipids among ~300,000 multi-ethnic participants of the Million Veteran Program. *Nature genetics* **50**, 1514-1523, doi:10.1038/s41588-018-0222-9 (2018).

14 Holmen, O. L. *et al.* Systematic evaluation of coding variation identifies a candidate causal variant in TM6SF2 influencing total cholesterol and myocardial infarction risk. *Nature genetics* **46**, 345-351, doi:10.1038/ng.2926 (2014).

15 Asselbergs, F. W. *et al.* Large-scale gene-centric meta-analysis across 32 studies identifies multiple lipid loci. *American journal of human genetics* **91**, 823-838, doi:10.1016/j.ajhg.2012.08.032 (2012).

16 Albrechtsen, A. *et al.* Exome sequencing-driven discovery of coding polymorphisms associated with common metabolic phenotypes. *Diabetologia* **56**, 298-310, doi:10.1007/s00125-012-2756-1 (2013).

17 Saxena, R. *et al.* Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science (New York, N.Y.)* **316**, 1331-1336, doi:10.1126/science.1142358 (2007).
18 Iotchkova, V. *et al.* Discovery and refinement of genetic loci associated with cardiometabolic risk using dense imputation maps. *Nature genetics* **48**, 1303-1312, doi:10.1038/ng.3668 (2016).
19 Tachmazidou, I. *et al.* A rare functional cardioprotective APOC3 variant has risen in frequency in distinct population isolates. *Nature Communications* **4**, 2872, doi:10.1038/ncomms3872 (2013).
20 Tang, C. S. *et al.* Exome-wide association analysis reveals novel coding sequence variants associated with lipid traits in Chinese. *Nature Communications* **6**, 10206, doi:10.1038/ncomms10206 (2015).
21 van Leeuwen, E. M. *et al.* Genome of the Netherlands population-specific imputations identify an ABCA6 variant associated with cholesterol levels. *Nature Communications* **6**, 6065, doi:10.1038/ncomms7065 (2015).
22 Spracklen, C. N. *et al.* Association analyses of East Asian individuals and trans-ancestry analyses with European individuals reveal new loci associated with cholesterol and triglyceride levels. *Hum Mol Genet* **26**, 1770-1784, doi:10.1093/hmg/ddx062 (2017).
23 Kanai, M. *et al.* Genetic analysis of quantitative traits in the Japanese population links cell types to complex human diseases. *Nature genetics* **50**, 390-400, doi:10.1038/s41588-018-0047-6 (2018).
24 Sirugo, G., Williams, S. M. & Tishkoff, S. A. The Missing Diversity in Human Genetic Studies. *Cell* **177**, 26-31, doi:10.1016/j.cell.2019.02.048 (2019).
25 Khera, A. V. *et al.* Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nature genetics* **50**, 1219-1224, doi:10.1038/s41588-018-0183-z (2018).
26 Duncan, L. *et al.* Analysis of polygenic risk score usage and performance in diverse human populations. *Nature Communications* **10**, 3328, doi:10.1038/s41467-019-11112-0 (2019).
27 Buniello, A. *et al.* The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res* **47**, D1005-d1012, doi:10.1093/nar/gky1120 (2019).
28 Tishkoff, S. A. *et al.* The genetic structure and history of Africans and African Americans. *Science (New York, N.Y.)* **324**, 1035-1044, doi:10.1126/science.1172257 (2009).
29 Brown, B. C., Ye, C. J., Price, A. L. & Zaitlen, N. Transethnic Genetic-Correlation Estimates from Summary Statistics. *American journal of human genetics* **99**, 76-88, doi:10.1016/j.ajhg.2016.05.001 (2016).
30 Lee, S. H., Yang, J., Goddard, M. E., Visscher, P. M. & Wray, N. R. Estimation of pleiotropy between complex diseases using single-nucleotide polymorphism-derived genomic relationships and restricted maximum likelihood. *Bioinformatics* **28**, 2540-2542, doi:10.1093/bioinformatics/bts474 (2012).
31 Guo, J. *et al.* Quantifying genetic heterogeneity between continental populations for human height and body mass index. *Scientific reports* **11**, 5240, doi:10.1038/s41598-021-84739-z (2021).

32 Ge, T., Chen, C.-Y., Ni, Y., Feng, Y.-C. A. & Smoller, J. W. Polygenic prediction via Bayesian regression and continuous shrinkage priors. *Nature Communications* **10**, 1776, doi:10.1038/s41467-019-09718-5 (2019).
33 Majara, L. *et al.* Low generalizability of polygenic scores in African populations due to genetic and environmental diversity. *bioRxiv*, 2021.2001.2012.426453, doi:10.1101/2021.01.12.426453 (2021).
34 Lehmann, B. C. L., Mackintosh, M., McVean, G. & Holmes, C. C. High trait variability in optimal polygenic prediction strategy within multiple-ancestry cohorts. *bioRxiv*, 2021.2001.2015.426781, doi:10.1101/2021.01.15.426781 (2021).
35 Shi, H. *et al.* Population-specific causal disease effect sizes in functionally important regions impacted by selection. *Nature Communications* **12**, 1098, doi:10.1038/s41467-021-21286-1 (2021).
36 Martin, A. R. *et al.* Clinical use of current polygenic risk scores may exacerbate health disparities. *Nature genetics* **51**, 584-591, doi:10.1038/s41588-019-0379-x (2019).
37 Cavazos, T. B. & Witte, J. S. Inclusion of variants discovered from diverse populations improves polygenic risk score transferability. *Human Genetics and Genomics Advances* **2**, 100017, doi:https://doi.org/10.1016/j.xhgg.2020.100017 (2021).
38 Wojcik, G. L. *et al.* Genetic analyses of diverse populations improves discovery for complex traits. *Nature* **570**, 514-518, doi:10.1038/s41586-019-1310-4 (2019).
39 Bentley, A. R. *et al.* Multi-ancestry genome-wide gene–smoking interaction study of 387,272 individuals identifies new loci associated with serum lipids. *Nature genetics* **51**, 636-648, doi:10.1038/s41588-019-0378-y (2019).
40 Taliun, D. *et al.* Sequencing of 53,831 diverse genomes from the NHLBI TOPMed Program. *Nature* **590**, 290-299, doi:10.1038/s41586-021-03205-y (2021).
41 Kowalski, M. H. *et al.* Use of >100,000 NHLBI Trans-Omics for Precision Medicine (TOPMed) Consortium whole genome sequences improves imputation quality and detection of rare variant associations in admixed African and Hispanic/Latino populations. *PLoS genetics* **15**, e1008500, doi:10.1371/journal.pgen.1008500 (2019).

Table 1: Meta-analysis sample size by ancestry group

Ancestry Group	Sample Size	Number of Cohorts	Mean Sample Size per Cohort (range)	Number of Variants
European	1,320,016	146	10,928 (173-389,344)	47 M
East Asian	146,492	40	7,448 (150-131,050)	17 M
Admixed African/ African	99,432	19	5,330 (473-62,022)	33 M
Hispanic	48,057	10	6,032 (1,496-22,302)	27 M
South Asian	40,963	7	6,413 (1,796-16,110)	17 M
Total	1,654,960	201		52 M

The present meta-analysis represents a 6-fold overall increase in sample size relative to the most recent 2018 Million Veteran Program blood lipid meta-analysis[13], with a 2-fold increase in sample size of Admixed African and Hispanic individuals.

Figures

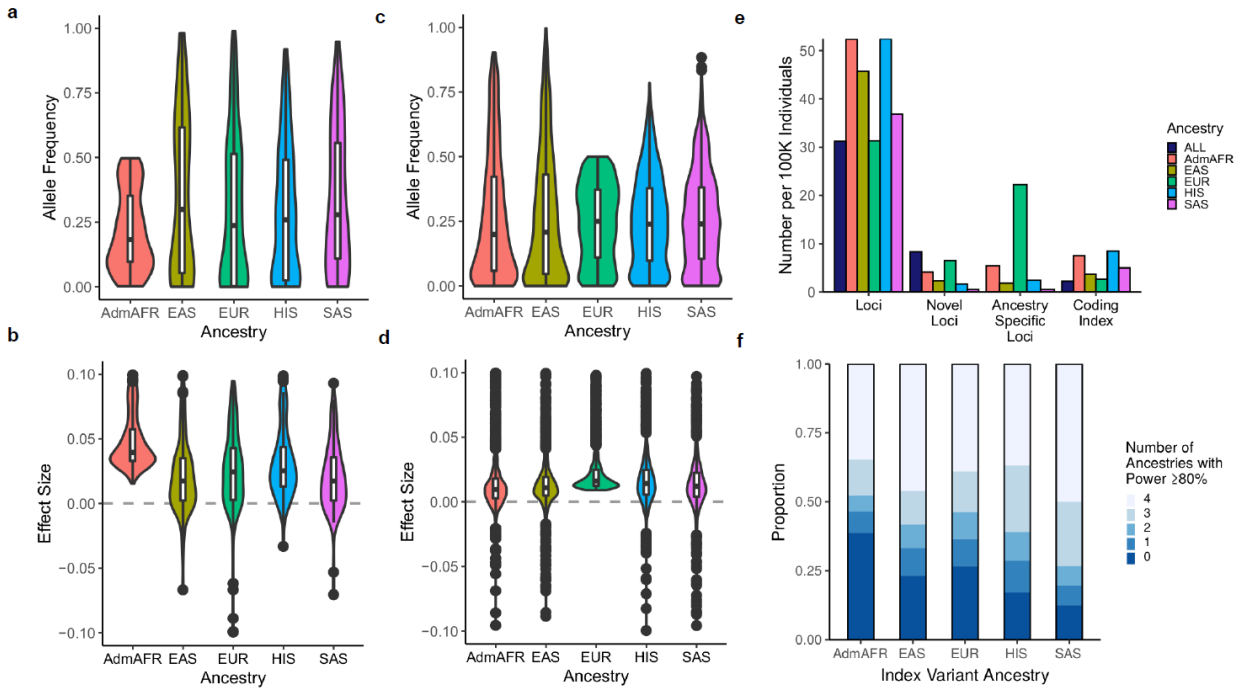


Figure 1: Comparison of identified loci across ancestry groups
a) Allele frequency distribution and b) effect sizes of Admixed African ancestry index variants in non-African ancestry populations. c) Allele frequency distribution and d) effect sizes of European ancestry index variants in non-European ancestry populations. Boxplots depict the median value as the center, first and third quartiles as box boundaries and whiskers extending 1.5 times the inter-quartile range, with points beyond this region shown individually. Sample sizes for each ancestry are provided in Table 1. The mean effect size of Admixed African ancestry identified index variants is larger than from European ancestry analysis, reflecting the difference in power to detect an association within each group as a result of the >10-fold difference in sample size. e) Number of loci identified within each ancestry group, normalized to a constant sample size of 100,000 individuals and averaged across lipid traits. At currently available sample sizes, trans-ancestry and European ancestry analyses identify a lower proportion of loci relative to the number of individuals than analyses of other ancestry groups. However, the larger sample size of European or trans-ancestry analyses leads to a greater relative proportion of novel loci and a higher proportion of loci significant only in European ancestry analyses. f) Proportion of index variants identified from each ancestry-specific meta-analysis that would be well-powered to detect an association of the same effect size but with ancestry-specific frequencies in the other ancestry groups. Dark blue regions indicate variants likely to be detected at an equivalent sample size only in the original ancestry group (i.e. ancestry-specific). Additional comparisons of allele frequencies and effect sizes across ancestries are provided in Supplementary Figure 3.

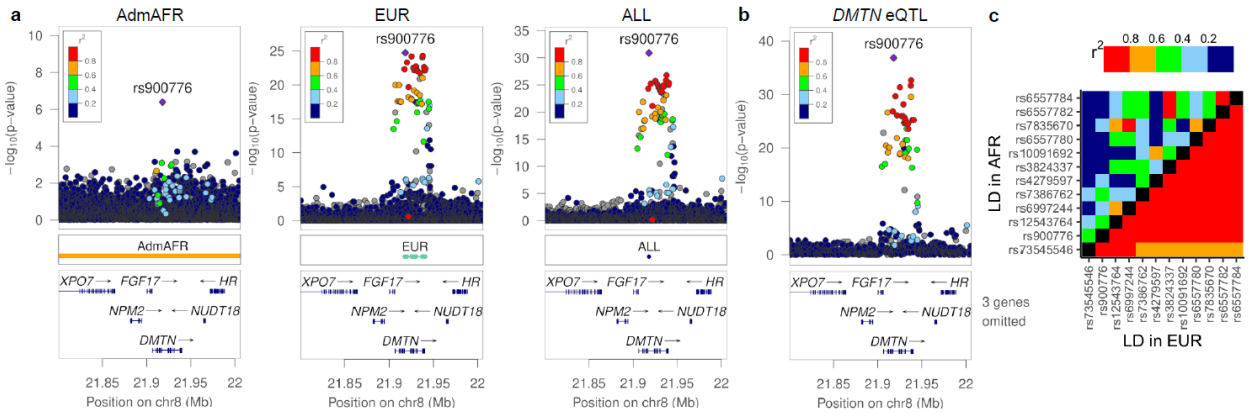


Figure 2: Inclusion of multiple ancestries drives improved fine-mapping
a) Association of the *DMTN* intron variant rs900776 with LDL-C or b) *DMTN* expression. The region spanned by the 99% credible sets are shown in the center box. The LDL-C association signal significantly colocalizes with the GTEx eQTL signal of *DMTN* in liver. c) The LD patterns for variants in the European ancestry 99% credible set differ greatly between African and European ancestry individuals in 1000 Genomes. The lead variant has a posterior probability of 0.86 in Admixed African, 0.51 in European, and >0.99 in the trans-ancestry analysis.

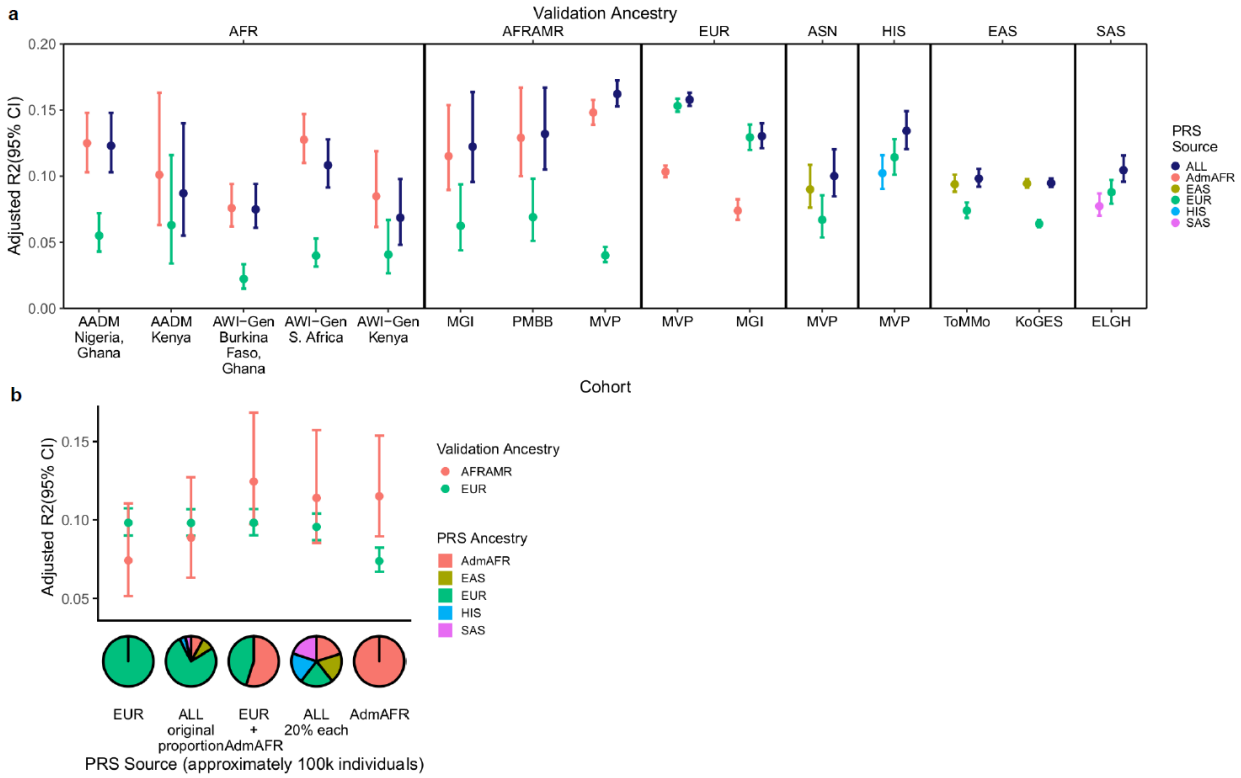


Figure 3: Trans-ancestry LDL-C PRS show similar performance across ancestry groups a) Polygenic scores generated from trans-ancestry meta-analysis show equivalent or better performance across most ancestry groups relative to ancestry-specific PRS within each cohort, whereas European ancestry-specific scores show less transferability. Adjusted R^2 is calculated with the risk score as a predictor of LDL-C in a linear model with covariates. AFR: African, AFRAMR: African American, ASN: Asian American b) Trans-ancestry scores derived from equal proportions of each ancestry group predict LDL-C better for African Americans in MGI than predominantly European ancestry scores at constant sample size. Error bars depict 95% confidence intervals. Sample sizes for each cohort are provided in Supplementary Table 16.

Methods:

Cohort level analysis

Each cohort contributed GWAS summary statistics for HDL-C, LDL-C, nonHDL-C, TC and TG, imputation quality statistics, and analysis metrics for quality control (QC), following a detailed analysis plan (**Supplementary File 1**). Briefly, we requested that each cohort perform imputation to 1000 Genomes Phase 3 (1KGP3), with European ancestry cohorts additionally imputing with the Haplotype Reference Consortium (HRC) panel using the Michigan Imputation Server (https://imputationserver.sph.umich.edu/index.html#!) which uses Minimac software[42]. Detailed pre-imputation QC guidelines were provided; these included removing samples with

call rate < 95%, samples with heterozygosity > median + 3(interquartile range), ancestry outliers from principal component analysis within each ancestry group, and variants deviating from Hardy-Weinberg equilibrium (p-value < 10^{-6}) or with variant call rate < 98%. Analyses were carried out separately by ancestry group and were additionally stratified by cases and controls where appropriate (i.e. for a disease-focused cohort such as CAD). Residuals were generated separately in males and females adjusting for age, age^2, principal components of ancestry, and any necessary study-specific covariates. Triglyceride levels were natural-log transformed before generating residuals. Inverse normalization was then done on the residual values. Individuals on cholesterol lowering medication had their pre-medication levels[43] approximated by dividing the LDL-C value by 0.7 and the TC value by 0.8. Association analysis of the residuals for the majority of cohorts was carried out using a linear mixed-model approach in rvtests or with other similar software including BOLT-LMM[44], SAIGE[45], or deCode association software.

Quality Control

Each input file was assessed for quality control using the EasyQC software[46] (www.genepi-regensburg.de/easyqc). We generated QQ plots by minor allele frequency (MAF) bins, assessed trends in standard errors relative to sample size for each cohort, and checked MAF of submitted variants relative to their expected value based on the imputation reference panel. In addition, we checked that each cohort reproduced the expected direction of effect at most known loci relative to the cohort sample size. Cohorts identified to have issues with the submitted files were contacted and corrected files were submitted or the cohort was excluded from meta-analysis. Results from either sex-stratified analysis or sex-combined analysis with sex as a covariate were used. During the QC process, within each cohort we removed poorly imputed variants (info score or r^2 < 0.3), variants deviating from Hardy-Weinberg Equilibrium (HWE p-value < 10^{-8}, except for MVP which used HWE p-value < 10^{-20}), and variants with minor allele count < 3. An imputation info score threshold of 0.3 was selected to balance the inclusion of variants across diverse studies while removing poorly imputed variants. Summary statistics were then genomic-control (GC) corrected using the λ_{GC} value calculated from the median p-value of variants with MAF > 0.5%. To capture as many variants as possible, summary statistics from cohorts that had submitted both HRC and 1KGP3 imputed files were joined, selecting variants imputed from HRC where both imputed versions of a variant existed. For variants imputed by both panels, we

observed that variants imputed from the HRC panel resulted in a higher imputation info score for 94% of variants when compared to the imputation info score from 1KGP3.

Meta-analysis

Ancestry-specific meta-analysis was performed using RAREMETAL[47] (https://github.com/SailajaVeda/raremetal). Trans-ancestry meta-analysis was performed using MR-MEGA[48] with 5 principal components of ancestry. The choice of 5 principal components was made after comparing the λ_{GC} values across minor allele frequency bins from meta-analysis of HDL-C with MR-MEGA using from 2 up to 10 principal components. In addition, fixed-effects meta-analysis was carried out with METAL[49] to calculate effect sizes for use in the creation of polygenic scores. Study-level principal components were plotted for each cohort by ancestry group to verify that the reported ancestry for each cohort was as expected. Following meta-analysis, we identified loci based on a genome-wide significance threshold of $5x10^{-8}$ after GC correction using the λ_{GC} value calculated from the median p-value of variants with MAF > 0.5%. The choice of double-GC correction was made to be most conservative and to minimize potential false-positive findings. Observed λ_{GC} values were within the expected range for similarly sized studies and are included in **Supplementary Tables 2** and **4**. Variants with a cumulative minor allele count ≤ 30 and those found in a single study were excluded from index variant selection. Index variants were identified following an iterative procedure starting with the most significant variant and grouping the surrounding region into a locus based on the larger of either ± 500 kb or ± 0.25 cM. cM positions were interpolated using the genetic map distributed with Eagle v2.3.2 (genetic_map_hg19_withX.txt)[50]. Variants were annotated using WGSA[51] including the summary of each variant from SnpEff[52] and the closest genes for intergenic variants from ANNOVAR[53]. Annotation of variants as known or novel was done based on manual review of previously published variants and with variants reported in the GWAS catalog[27] for any of the studied lipid traits (accessed May 2020, provided as **Supplementary Table 18**). For comparison between ancestries and lipid traits, index variants were grouped into genomic regions starting with the most significantly associated variant and grouping all surrounding index variants within ± 500 kb into a single region.

Power to detect association within each ancestry was determined using the effect size and sample size of the variant within the original discovery ancestry group and the observed allele frequency from the other ancestry groups with alpha set to $5x10^{-8}$. We excluded variants that were only successfully imputed in a single ancestry group to account for imputation panel differences between groups (ie. Haplotype Reference Consortium for European ancestry individuals and 1000 Genomes for other ancestries). Variants that were successfully imputed in 2 or more ancestries were assumed to have zero power in any other ancestry where the variant was not successfully imputed. The proportion of variance explained by each variant was estimated as $2\beta^2(1-f)f$ where β is the effect size from METAL and f is the effect allele frequency (**Supplementary Table 19**). The proportion of variance explained within each ancestry was estimated using the trans-ancestry effect size from METAL with the ancestry-specific allele frequency. Coverage of the genome by associated genetic regions was calculated using BEDTools[54] for the regions defined by the minimum and maximum position within each locus having p-value < $5x10^{-8}$.

Conditional analysis

Approximate conditional analysis was performed using rareGWAMA[55] to identify index variants that were shadows of nearby, more significant associations. LD reference populations were taken from UK Biobank specific to Admixed African, European (subset of 40,000), or South Asian ancestry individuals or from the 1000 Genomes project (1KGP3) for East Asian or Hispanic ancestry individuals. Conditional analysis was carried out using the individual cohort level summary statistics as was done for meta-analysis with RAREMETAL. rareGWAMA requires imputation quality scores which were set to 1 for all variants that had previously passed quality control (pre-filtered at imputation info/r^2 > 0.3). The European ancestry subset of UK Biobank was used as the reference population for the conditional analysis of the trans-ancestry meta-analysis (~80% European ancestry). Stepwise conditional analysis was performed sequentially for the index variants within each chromosome ranked by most to least significant. Index variants were then flagged as not independent from other more significant variants if the absolute value of the ratio of the original effect size to the effect size after conditional analysis was greater than the 95^{th} percentile of all values (**Supplementary Figure 10**). This threshold was selected to remove variants whose effects were driven by nearby, more strongly associated

variants in LD. This corresponded to a ratio of original to conditional effect size of 1.6 for ancestry-specific conditional analysis and a ratio of 1.7 for the trans-ancestry conditional analysis. The effect sizes from meta-analysis with METAL were used for comparison with the trans-ancestry conditional analysis results. Variants flagged as non-independent were excluded from the summary results in the manuscript and are flagged as non-independent in **Supplementary Tables 3** and **5**.

Genetic correlation

Popcorn[29] was used to assess the degree of correlation in effect sizes between ancestry groups for each of the lipid traits with 1000 Genomes phase 3 as the reference LD panel. Only variants with MAF > 0.01 in each ancestry individually were included in the comparison. Both the genetic effect and genetic impact models were tested. Bivariate GREML from GCTA was used to calculate the genetic correlation between unrelated Admixed Africans and a subset of white British individuals in the UK Biobank following the method of Guo et al[30,31]. HapMap3 variants with MAF > 0.01 in each ancestry were used to construct the genetic relationship matrix (GRM) with the allele frequencies standardized in each population. Individuals with genetic relatedness > 0.05 were removed. A total of up to 5,575 AdmAfr and 38,668 white British individuals from UK Biobank were included in the analysis of each trait after removal of related individuals. The measured lipid traits were corrected for medication use and were inverse-normalized after correction for age, sex, and batch. Principal components 1-20 constructed from the GRM were included as covariates in the calculation of genetic correlation. Analysis within the Million Veteran Program included 24,502 European ancestry and 21,950 African American unrelated individuals. Maximum measured values were used for LDL-C, TC, and triglycerides and minimum values for HDL-C. Lipid traits were inverse-normalized after correction for age and sex with principal components 1-20 included as covariates in the calculation of genetic correlation.

Credible sets

Credible sets of potentially causal variants were generated for each of the loci identified in the trans-ancestry meta-analysis. We determined 99% credible sets of variants that encompassed the causal variant with 99% posterior probability. Regions for construction of credible sets were

defined as the ± 500 kb region around each index variant. Bayes factors[56,57] (BF) for each variant in the ancestry-specific meta-analysis were approximated by:

$$BF \approx \exp\left[0.5\left(\frac{\beta^2}{SE^2} - \log(N_{AS})\right)\right]$$

where β and SE are the effect sizes and standard errors from the RAREMETAL meta-analysis, and N_{AS} is the ancestry-specific sample size. A full derivation is included in the **Supplementary Methods**. To account for the difference in sample sizes between ancestry groups, we additionally approximated the Bayes factors after adjustment for the total trans-ancestry sample size for each trait (N_{TE}) relative to the ancestry-specific sample size for that trait using the following equation:

$$BF \approx \exp\left[0.5\left(\frac{\beta^2 N_{TE}}{SE^2 N_{AS}} - \log(N_{TE})\right)\right]$$

Credible sets for the trans-ancestry meta-analysis were generated using the Bayes factors as output by MR-MEGA. The credible sets within each region were generated by ranking all variants by Bayes factor and calculating the number of variants required to reach a cumulative probability of 99%. In addition, we calculated credible sets in the same manner using the European ancestry and trans-ancestry meta-analysis results but including only the set of variants present in the AdmAFR meta-analysis. To summarize the size of the credible sets across the 5 lipid traits examined, we identified the set of independent index variants from the trans-ancestry meta-analysis after grouping variants based on LD. For each ± 500kb region centered around the most-significantly associated index variant for any trait, we determined the pairwise LD between all index variants in this region using LDpair[58] with all reference populations (1000 Genomes AFR, AMR, EAS, EUR, and SAS) included. We considered variants to be independent if they were outside of this region, had LD $r^2 < 0.7$, or were not available in the LDpair reference populations. Variants within the credible sets were annotated with SnpEff[52] using WGSA[51] and with VEP[59]. The number of variants in LD with an index variant was determined using LDproxy[58] (**Supplementary Table 20**). Protein numbering was taken from dbSNP[60]. eQTL

colocalization was performed using coloc[61] version 3.2.1 with R version 3.4.3 using the default parameters. Results from GTEx V8[62] were compared with the GWAS signals in the region defined by the larger of ±0.25cM or ±500kb surrounding each index variant. The eQTL and GWAS signals (based on p-values from MR-MEGA) were considered to be colocalized if PP3 + PP4 ≥ 0.8 and if PP4/(PP3+PP4) > 0.9, where PP3 is the probability of two independent causal variants while PP4 is the probability of a single, shared causal variant.

LDL-C polygenic scores

Weights for the LDL-C polygenic scores were derived from beta estimates generated from each of the ancestry-specific meta-analyses and from the trans-ancestry results using METAL. Additional meta-analyses were carried out using the 2010 Global Lipids Genetics Consortium LDL-C meta-analysis results[4] in combination with the i) Admixed African or ii) Admixed African, East Asian, Hispanic, and South Asian ancestry results from the present meta-analysis for comparison. Furthermore, we performed a meta-analysis of European ancestry cohorts randomly selected to reach a total sample size near 100K, 200K, or 400K to understand the role of increasing European ancestry sample size and the influence of imputation panel. In addition, we tested possible methods for improving performance of European ancestry derived scores in African ancestry individuals by separately fitting the European ancestry polygenic scores in the UK Biobank Admixed African ancestry subset to determine the best set of risk score parameters (various pruning and thresholding parameters or PRS-CS, **Supplementary Note**).

We generated polygenic score weights using both: i) significant variants only (at a variety of p-value thresholds) and ii) using genome-wide methods. Meta-analysis results were first filtered to variants present in UK Biobank, MGI, and MVP with imputation info score > 0.3. Pruning and thresholding was performed in PLINK[63] with ancestry-matched subsets of UK Biobank individuals (AdmAFR N=7,324, EUR N=40,000, SAS N=7,193, trans-ancestry: N=10,000 (80% EUR, 15% AdmAFR, 5% SAS)) or 1KGP3 (HIS N=347 , EAS N=504) used for LD reference. We additionally tested 1000 Genomes phase 3 with all populations included as the LD reference panel for the trans-ancestry score (results not shown), which gave very similar results to those of the UK Biobank trans-ancestry reference set originally selected for its larger sample size. P-value thresholds (after GC correction) of $5x10^{-10}$, $5x10^{-9}$, $5x10^{-8}$, $5x10^{-7}$, $5x10^{-6}$, $5x10^{-5}$, $5x10^{-4}$, $5x10^{-3}$,

and $5x10^{-2}$ were tested with distance thresholds of 250 and 500 kb and LD r^2 thresholds of 0.1 and 0.2. Polygenic score weights were also generated using PRS-CS[32] with the LD reference panels for African, East Asian, and European ancestry populations from 1000 Genomes provided by the developers. PRS-CS LD reference panels for the other ancestries were generated using 1000 Genomes following the same protocol as provided by the PRS-CS authors[32]. This included removing variants with MAF ≤ 0.01, ambiguous A/T or G/C variants, and restricting to variants included in HapMap3. Pairwise LD matrices within pre-defined LD blocks[64] (using European LDetect blocks for Hispanic and trans-ancestry LD calculations and Asian blocks for South Asian) were then calculated using PLINK and converted to HDF5 format.

For each individual in the testing cohorts, polygenic scores were calculated as the sum of the dosages multiplied by the given weight at each variant. UK Biobank individuals not present in datasets used to generate the summary statistics (either Admixed African, white British, both Admixed African and white British, East Asian, South Asian, or all individuals excluding South Asian) were used to select the best performing Admixed African, European, Admixed African+European, East Asian, South Asian, and trans-ancestry polygenic scores, respectively. UK Biobank South Asian ancestry individuals were included in the trans-ancestry risk score weights but excluded from the UK Biobank trans-ancestry testing set due to an initial focus on comparing predictions among European and African ancestry individuals. Sample sizes of the ancestry groups in UK Biobank used to test PRS performance included: AdmAFR N=6,863; EAS N=1,441; EUR N=389,158; SAS N=6,814; ALL=461,918. The best performing Hispanic ancestry polygenic score weights were selected based on performance in Hispanic ancestry individuals in the Michigan Genomics Initiative dataset. Model fit was assessed by the adjusted R^2 of a linear model for LDL-C value at initial assessment adjusted for cholesterol medication (divided by 0.7 to estimate pre-medication levels) with sex, batch, age at initial assessment, and PCs1-4 as covariates (**Supplementary Tables 21-23**). Python and R were used for analysis of PRS models.

The best performing polygenic score in each ancestry group was then tested in the validation cohorts: the Michigan Genomics Initiative (EUR N=17,190; AFRAMR N=1,341), East London Genes and Health[65] (ELGH; SAS N=15,242), Tohoku Medical Megabank Community Cohort Study (ToMMo; EAS N=28,217), Korean Genome and Epidemiology Study[66] (KoGES; EAS

N=118,260), Penn Medicine BioBank (PMBB; AFRAMR=2,138), Africa America Diabetes Mellitus (AADM; 3,566 West AFR; 707 East AFR), Africa Wits-INDEPTH partnership for Genomic Studies (AWI-Gen; 1,744 East AFR; 4,972 South AFR; 3,744 West AFR) and Million Veteran Program participants not included in the discovery meta-analysis (MVP; EUR N=68,381; AFRAMR N=18,251; EAS/SAS N=4,155; HIS N=7,669). Adjusted R^2 values were reported for each cohort and ancestry group, with 95% confidence intervals for the adjusted R^2 values calculated using bootstrapping. Within each cohort, covariates used were: MGI- sex, batch, PC1-4, and birth year; PMBB- birth year, sex, and PC1-4; ELGH- age, sex, and PC1-10; MVP- sex, PC1-4, birth year, and mean age; ToMMo-sex, age, recruitment method, and PC1-20 (only participants from Miyagi Prefecture were included); KoGES-age, sex, and recruitment area, AADM-age, sex, PC1-3, AWI-Gen East Africa- age, sex, PC1-6, AWI-Gen South Africa- age, sex, PC1-6, and AWI-Gen West Africa- age, sex, and PC1-4. The type of LDL-C value used in the model varied depending on the measurements selected by each cohort. Mean LDL-C values were used for MGI, MVP and PMBB, maximum LDL-C values for ELGH, and baseline measurements for AADM, AWI-Gen, ToMMo and KoGES. A descriptive summary of each validation cohort is included in **Supplementary Table 16**. African admixture for MGI was calculated using all African ancestry individuals in 1000 Genomes with ADMIXTURE v1.3[67]. African admixture for MVP was calculated using the YRI and LWK African ancestry individuals in 1000 Genomes.

Methods References

42 Das, S. *et al.* Next-generation genotype imputation service and methods. *Nature genetics* **48**, 1284-1287, doi:10.1038/ng.3656 (2016).

43 Efficacy and safety of cholesterol-lowering treatment: prospective meta-analysis of data from 90 056 participants in 14 randomised trials of statins. *The Lancet* **366**, 1267-1278, doi:https://doi.org/10.1016/S0140-6736(05)67394-1 (2005).

44 Loh, P.-R. *et al.* Efficient Bayesian mixed-model analysis increases association power in large cohorts. *Nature genetics* **47**, 284-290, doi:10.1038/ng.3190 (2015).

45 Zhou, W. *et al.* Efficiently controlling for case-control imbalance and sample relatedness in large-scale genetic association studies. *Nature genetics* **50**, 1335-1341, doi:10.1038/s41588-018-0184-y (2018).

46 Winkler, T. W. *et al.* Quality control and conduct of genome-wide association meta-analyses. *Nat Protoc* **9**, 1192-1212, doi:10.1038/nprot.2014.071 (2014).

47 Feng, S., Liu, D., Zhan, X., Wing, M. K. & Abecasis, G. R. RAREMETAL: fast and powerful meta-analysis for rare variants. *Bioinformatics* **30**, 2828-2829, doi:10.1093/bioinformatics/btu367 (2014).

48 Mägi, R. *et al.* Trans-ethnic meta-regression of genome-wide association studies accounting for ancestry increases power for discovery and improves fine-mapping resolution. *Human Molecular Genetics* **26**, 3639-3650, doi:10.1093/hmg/ddx280 (2017).
49 Willer, C. J., Li, Y. & Abecasis, G. R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190-2191, doi:10.1093/bioinformatics/btq340 (2010).
50 Loh, P.-R., Palamara, P. F. & Price, A. L. Fast and accurate long-range phasing in a UK Biobank cohort. *Nature genetics* **48**, 811-816, doi:10.1038/ng.3571 (2016).
51 Liu, X. *et al.* WGSA: an annotation pipeline for human genome sequencing studies. *Journal of Medical Genetics* **53**, 111-112, doi:10.1136/jmedgenet-2015-103423 (2016).
52 Cingolani, P. *et al.* A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. *Fly* **6**, 80-92, doi:10.4161/fly.19695 (2012).
53 Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Research* **38**, e164-e164, doi:10.1093/nar/gkq603 (2010).
54 Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841-842, doi:10.1093/bioinformatics/btq033 (2010).
55 Liu, D. J. *et al.* Meta-analysis of gene-level tests for rare variant association. *Nature genetics* **46**, 200-204, doi:10.1038/ng.2852 (2014).
56 Maller, J. B. *et al.* Bayesian refinement of association signals for 14 loci in 3 common diseases. *Nature genetics* **44**, 1294-1301, doi:10.1038/ng.2435 (2012).
57 Kass, R. E. & Raftery, A. E. Bayes Factors. *Journal of the American Statistical Association* **90**, 773-795, doi:10.1080/01621459.1995.10476572 (1995).
58 Machiela, M. J. & Chanock, S. J. LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. *Bioinformatics* **31**, 3555-3557, doi:10.1093/bioinformatics/btv402 (2015).
59 McLaren, W. *et al.* The Ensembl Variant Effect Predictor. *Genome Biology* **17**, 122, doi:10.1186/s13059-016-0974-4 (2016).
60 Sherry, S. T. *et al.* dbSNP: the NCBI database of genetic variation. *Nucleic acids research* **29**, 308-311, doi:10.1093/nar/29.1.308 (2001).
61 Giambartolomei, C. *et al.* Bayesian Test for Colocalisation between Pairs of Genetic Association Studies Using Summary Statistics. *PLoS genetics* **10**, e1004383, doi:10.1371/journal.pgen.1004383 (2014).
62 The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science (New York, N.Y.)* **369**, 1318-1330, doi:10.1126/science.aaz1776 (2020).
63 Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *American journal of human genetics* **81**, 559-575, doi:10.1086/519795 (2007).
64 Berisa, T. & Pickrell, J. K. Approximately independent linkage disequilibrium blocks in human populations. *Bioinformatics* **32**, 283-285, doi:10.1093/bioinformatics/btv546 (2016).

65 Finer, S. *et al.* Cohort Profile: East London Genes & Health (ELGH), a community-based population genomics and health study in British Bangladeshi and British Pakistani people. *International Journal of Epidemiology* **49**, 20-21i, doi:10.1093/ije/dyz174 (2019).
66 Moon, S. *et al.* The Korea Biobank Array: Design and Identification of Coding Variants Associated with Blood Biochemical Traits. *Scientific reports* **9**, 1382, doi:10.1038/s41598-018-37832-9 (2019).
67 Alexander, D. H., Novembre, J. & Lange, K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res* **19**, 1655-1664, doi:10.1101/gr.094052.109 (2009).

Data Availability

The GWAS meta-analysis results (including both ancestry-specific and trans-ancestry analyses) and risk score weights are available at: http://csg.sph.umich.edu/willer/public/glgc-lipids2021. The optimized trans-ancestry and single-ancestry polygenic score weights will be deposited within the PGS Catalog (https://www.pgscatalog.org/).

Code Availability

EasyQC: www.genepi-regensburg.de/easyqc; Raremetal: https://github.com/SailajaVeda/raremetal

Acknowledgments

Funding for the Global Lipids Genetics Consortium was provided by the NIH (R01-HL127564). This research has been conducted using the UK Biobank Resource under application number 24460. Computing support and file management for central meta-analysis by Sean Caron is gratefully acknowledged. This research is based on data from the Million Veteran Program, Office of Research and Development, Veterans Health Administration, and was supported by awards #2I01BX003362-03A1 and 1I01BX004821-01A1#. This publication does not represent the views of the Department of Veteran Affairs or the United States Government. Study-specific acknowledgements are provided in the supplemental material.

Author Contributions

All authors reviewed the manuscript. Consortium management: G.M.P., P.N., T.L.A., M.B., S.K., C.J.W. Study design, interpretation of results, and drafting of manuscript: S.E.G., S.L.C., K-H.H.W., S.K., G.J.M.Z., S.R., I.S., I.N., E.M., K.L.M., T.M.F., J.N.H., S.K., M.B., P.N., G.M.P., C.D.B., A.P.M., Y.V.S., P.D., T.L.A., C.J.W. Primary meta-analysis and quality control: S.E.G., S.V., T.W.W., A.E.L. Polygenic score analysis and development: S.E.G., S.L.C., K-H.H.W., S.K., M.Y.H., S.H., A.N., A.C., A.R.B., K.E., A.V., B.T., H.C.M., K.A.H., C.N.R., S.H., M.R., R.C.T., D.A.vH., G.T., M.Y., B-J.K. Individual study genetic analysis: S.E.G., S.K., S.V., A.E.L., K.L.M., G.M.P., P.D., C.J.W., Q.H., D.K., X.Z., G.T., A.H., D.F.G., H.H., I.O., M.A., S.S., C.T., M.K., W.Z., B.M.B., H.R., S. E.R., A.S.H., Y.V., Q.F., E.A.R., T.L., J.A.P., S.A.P., J.H., F.G., Y.B., J.E.M., A.C., K.L., I.Y.M ., G.H., A.R., J.D.F., W.Z., D.R.W., C.T., H.H., M.G., A.M., M.R.B., W.Z., K.Y., E.M.S., A.P., S.G., X.Y., J.L., J.Z., F.M., H.J., K.Y., C.M., A.P., J.H., G.W., A.R.W., Y.J., Z.G., S.H., R.E.M., J.C., M.A., J.Y., A.M., H.R.W., J.R., J.B., L.L.K., A.G., M.S., R.N., C.S., E.F., A.F.M., P.M., M .W., S.T., N.S., L.T.M., B.H.T., M.M., L.Z., J.H., B.Y., A.P., A.K., C.L., L.F., M.S., T.E.G., J.P. B., E.W.D., J.M.Z., J.S.M., C.F., H.C., J.A.B., M.F.F., M.K.W., M.P., M.M., P.C., N.V., J.W.B., J.E., R.L.K., R.C.S., K.L., N.R.Z., P.L., M.E.K., G.E.D., S.H., D.D.I., H.I., J.Y., J.L., H.L.L., J. M., B.S., M.A., L.J.S., M.C., C.W., M.N., A.W., N.H., X.S., R.X., A.H., J.C.F., V.L., M.A., A.U .J., M.R.I., C.O., H.K., S.R., P.R.T., L.A., R.D., L.A.L., X.C., G.P., L.L., M.P., J.L., X.L., E.T., F.T., C.N.S., A.L., S.B., S.C.W., Y.W., W.B.W., T.N., D.R., Y.S., Y.H., S.C., F.L., J.Y., K.A.K., M.G., M.B., K.M., L.F.B., J.A.S., P.H., A.F., E.H., M.L., C.X., J.Z., M.C., S.V., P.J.v., N.P., B. E.C., J.L., S.v., K.C., S.W., M.E.Z., J.L., H.C., M.N., S.F., L.S., N.W.R., C.A.W., S.L., J.W., C. C., L.L., K.N., G.C., H.V., B.H., O.G., Q.C., M.O.O., J.v., X.L., K.S., N.T., J.S., R.D.J., A.P.R., L.W.M., Z.C., L.L., H.M.H., K.L.Y., T.K., J.T., J.C.B., G.N.N., L.J.L., H.L., M.A.N., O.T.R., S.I ., S.H.W., C.P.N., H.C., S.J., T.N., F.A., H.N., P.S.B., I.K., P.K., T.G., T.K., K.B., D.d., G.d., E. K., H.H.A., M.I., X.Z., F.W.A., A.O.K., J.W.B., X.S., L.S.R., O.P., T.H., P.M., A.W.H., M.K., L .P., C.B., A.T., Y.C., C.E.P., T.A.M., W.L., A.F., C.O., D.M., Y.C., H.L., J.Y., W.K., S.R., J.W., I.M.H., K.J.S., H.V., G.H., M.K.E., A.B.Z., O.P., G.P., I.E.H., S.R., K.P., A.J.O., H.S., G.B., R. S., H.S., Y.E.C., S.B., G.D., T.T., S.L.K., N.K., M.B.S., G.G., B.J., C.A.B., P.K.J., D.A.B., P.L. D., X.L., V.M., M.B., M.J.C., P.B.M., X.G., M.C., J.B.J., N.J.S., D.I.C., J.K., P.P., T.T., C.A.A.,

L.S.A., S.A.B., H.d., A.R.W., R.K., J.W., W.Z., A.I.d., D.B., A.C., J.G.W., L.L., C.H., A.E.N., Y .M.G., J.F.W., B.P., H.K., J.A., R.J.S., D.C.R., D.K.A., M.W., H.A.K., G.R.C., C.S.Y., J.M.M., T.T., C.A., C.G.V., L.O., M.F., E.T., R.M.v., T.L., N.C., M.Y., J.L., D.F.R., A.M., F.K., K.J., M. I.M., C.N.P., V.V., C.H., E.S., C.M.v., F.L., J.Q., H.H., X.L., W.M., E.J.P., M.C., V.G., J.T., G. L., L.M.t., P.J.E., D.J.R., S.M.D., M.K., M.K., P.v., T.D.S., R.J.L., M.A.P., B.M.P., I.B., P.P.P., K.C., S.R., E.W., H.H., S.F.G., L.A.K., J.d., M.L., F.K., D.G., J.E., H.S., P.W.F., A.L., J.W.J., A. V.K., M.M., M.J., Z.K., F.C., D.O.M., K.W., H.W., D.P.S., N.G., P.S., N.P., J.I.R., T.M.D., F.K. , M.J.N., N.J.T., C.C., T.W., C.K., C.S., A.P., C.G., A.T.H., N.L.P., P.K.M., D.I.B., E.J.d., L.A.C ., J.B.v., M.G., P.G., W.H., Y.K., Y.T., N.J.W., C.L., E.Z., J.K., M.L., E.I., G.A., J.C.C., J.S.K., P.S.d., A.C.M., K.E.N., M.D., P.K., N.G.M., J.B.W., S.A., D.S., R.G.W., M.V.H., C.B., B.H.S., A.E.J., A.B., J.E.B., P.M.R., D.I.C., C.K., W.W., G.P.J., B.N., M.H., M.D.R., P.J., V.S., K.H., B. Å., M.K., Y.K., Y.O., Y.M., U.T., K.S., Y.H., J.A.L., D.R., P.S.T., K.C., K.C., C.J.O., J.M.G., P. W.

Competing interests

G.C-P. is currently an employee of 23andMe Inc. M.J.C. is the Chief Scientist for Genomics England, a UK Government company. B.M.P. serves on the steering committee of the Yale Open Data Access Project funded by Johnson & Johnson. G.T., A.H., D.F.G., H.H., U.T., and K.S. are employees of deCODE/Amgen Inc. V.S. has received honoraria for consultations from Novo Nordisk and Sanofi and has an ongoing research collaboration with Bayer Ltd. M.M. has served on advisory panels for Pfizer, NovoNordisk and Zoe Global, has received honoraria from Merck, Pfizer, Novo Nordisk and Eli Lilly, and research funding from Abbvie, Astra Zeneca, Boehringer Ingelheim, Eli Lilly, Janssen, Merck, NovoNordisk, Pfizer, Roche, Sanofi Aventis, Servier, and Takeda. M.M. and A.M. are employees of Genentech and a holders of Roche stock. M.S. receives funding from Pfizer Inc. for a project unrelated to this work. M.E.K. is employed by SYNLAB MVZ Mannheim GmbH. W.M. has received grants from Siemens Healthineers, grants and personal fees from Aegerion Pharmaceuticals, grants and personal fees from AMGEN, grants from Astrazeneca, grants and personal fees from Sanofi, grants and personal fees from Alexion Pharmaceuticals, grants and personal fees from BASF, grants and personal fees from Abbott Diagnostics, grants and personal fees from Numares AG, grants and personal fees from Berlin-Chemie, grants and personal fees from Akzea Therapeutics, grants from Bayer Vital

GmbH , grants from bestbion dx GmbH, grants from Boehringer Ingelheim Pharma GmbH Co KG, grants from Immundiagnostik GmbH, grants from Merck Chemicals GmbH, grants from MSD Sharp and Dohme GmbH, grants from Novartis Pharma GmbH, grants from Olink Proteomics, other from Synlab Holding Deutschland GmbH, all outside the submitted work. A.V.K. has served as a consultant to Sanofi, Medicines Company, Maze Pharmaceuticals, Navitor Pharmaceuticals, Verve Therapeutics, Amgen, and Color Genomics; received speaking fees from Illumina, the Novartis Institute for Biomedical Research; received sponsored research agreements from the Novartis Institute for Biomedical Research and IBM Research, and reports a patent related to a genetic risk predictor (20190017119). S.K. is an employee of Verve Therapeutics, and holds equity in Verve Therapeutics, Maze Therapeutics, Catabasis, and San Therapeutics. He is a member of the scientific advisory boards for Regeneron Genetics Center and Corvidia Therapeutics; he has served as a consultant for Acceleron, Eli Lilly, Novartis, Merck, Novo Nordisk, Novo Ventures, Ionis, Alnylam, Aegerion, Haug Partners, Noble Insights, Leerink Partners, Bayer Healthcare, Illumina, Color Genomics, MedGenome, Quest, and Medscape; he reports patents related to a method of identifying and treating a person having a predisposition to or afflicted with cardiometabolic disease (20180010185) and a genetics risk predictor (20190017119). D.K. accepts consulting fees from Regeneron Pharmaceuticals. D.O.M-K. is a part-time clinical research consultant for Metabolon, Inc. D.S. has received support from the British Heart Foundation, Pfizer, Regeneron, Genentech, and Eli Lilly pharmaceuticals. The spouse of C.J.W. is employed by Regeneron.

Supplementary Information:

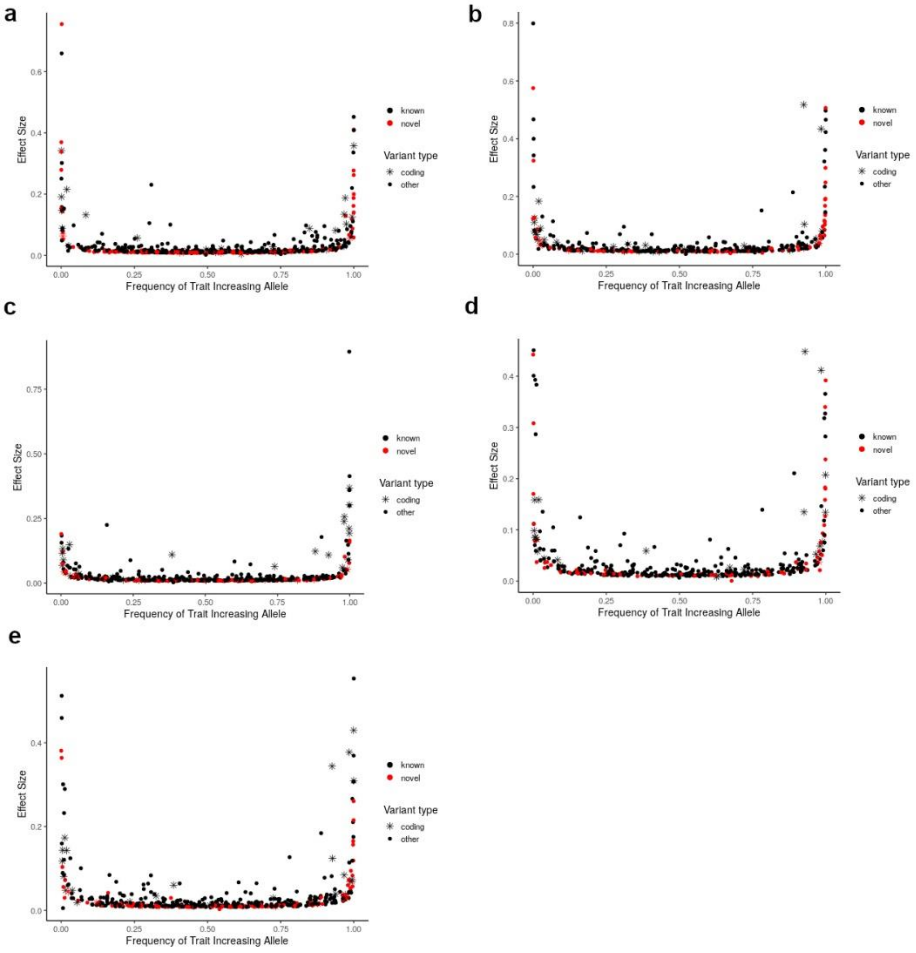
Supplementary Information: This file contains acknowledgements for each cohort, VA Million Veteran Program and Global Lipids Genetics Consortium authors, Supplementary Tables 2,4,8,13,21-23, Supplementary Figures 1-10, the Supplementary Notes and Supplementary Methods.

Supplementary Tables: This file contains Supplementary Tables 1,3,5-7,9-12,14-20

Correspondence:

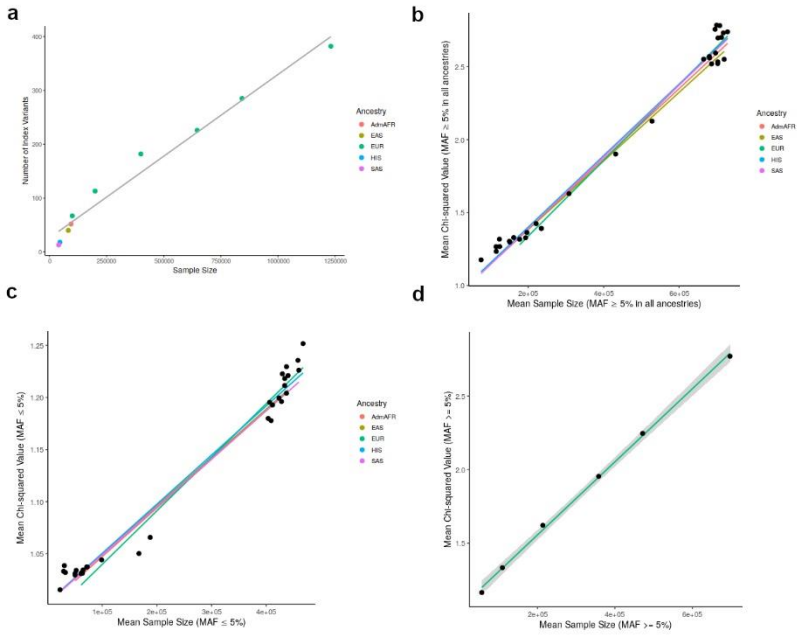
Correspondence and requests for materials should be addressed to Cristen Willer, cristen@umich.edu, or Themistocles L. Assimes, tassimes@stanford.edu

Extended Data Figures



Extended Data Figure 1: Effect sizes of identified index variants from trans-ancestry meta-analysis

Index variants associated with a) HDL cholesterol, b) LDL cholesterol, c) triglycerides, d) nonHDL cholesterol and e) total cholesterol include both common variants of small to moderate effect and low frequency variants of moderate to large effect.



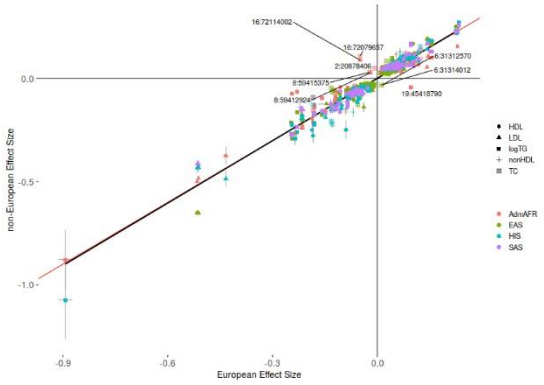
Extended Data Figure 2: Comparison of the number of index variants by sample size

a) Comparison of the number of index variants reaching genome-wide significance ($p < 5x10^{-8}$) from meta-analysis of LDL-C in each ancestry group. A meta-analysis of five random subsets of European cohorts selected to reach sample sizes of approximately 100,000, 200,000, 400,000, 600,000, or 800,000 individuals is also shown.

b) Comparison of chi-squared values from meta-analysis of LDL-C for each possible combination of ancestry groups (without genomic-control correction) for variants with minor allele frequency (MAF) ≥ 5%. The colored lines indicate a linear regression model of all meta-analyses for a specific ancestry (eg. all analyses including European individuals).

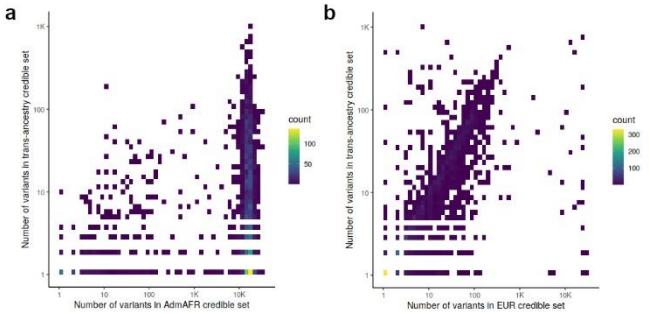
c) Comparison of chi-squared values from meta-analysis of LDL-C for variants with MAF ≤ 5%.

d) Comparison of chi-squared valued for variants with MAF ≥ 5% for LDL-C without genomic-control correction in a meta-analysis of all European cohorts as well as five subsets selected to reach sample sizes of approximately 100,000, 200,000, 400,000, 600,000, or 800,000 individuals.



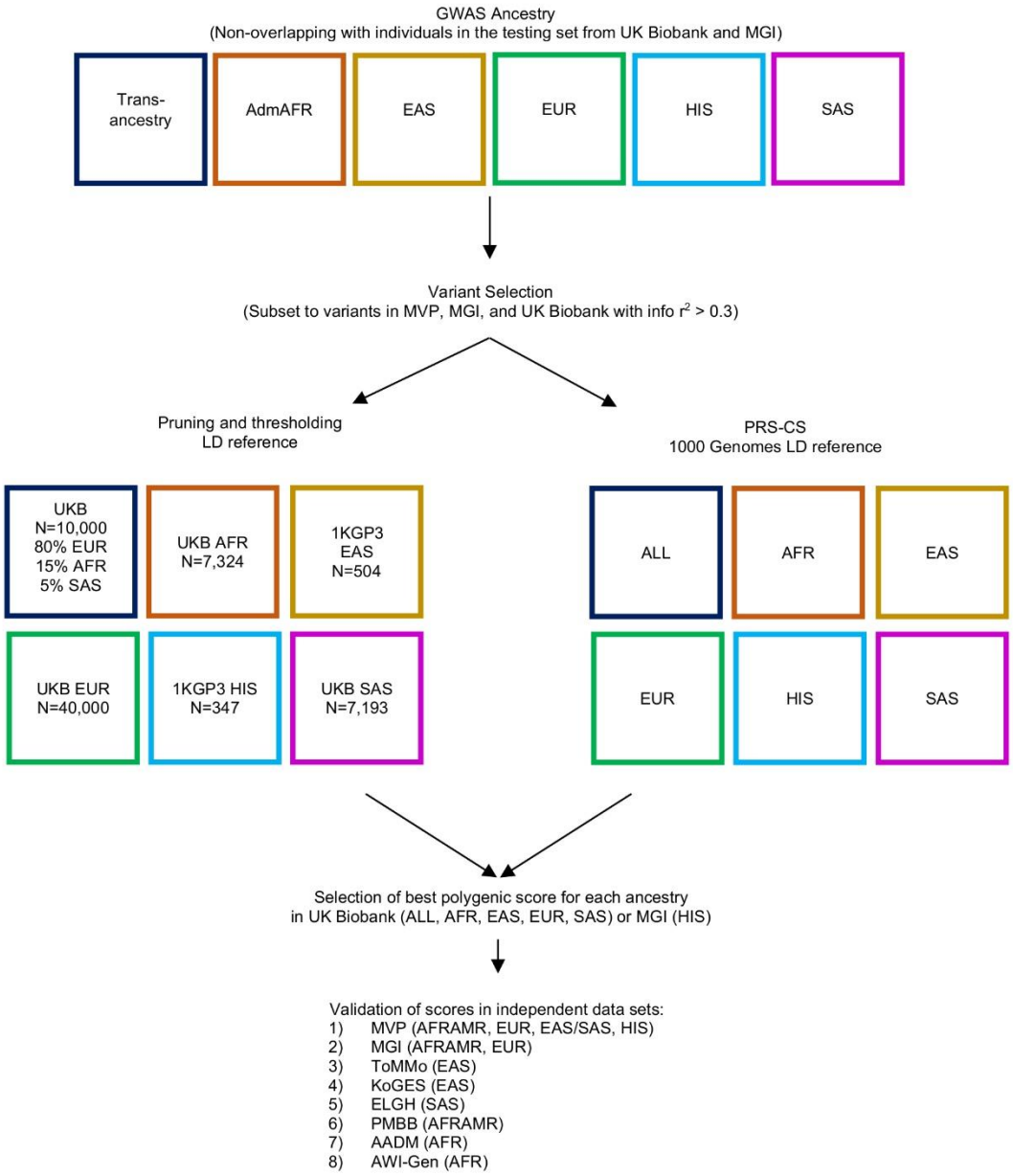
Extended Data Figure 3: Effect sizes by ancestry for unique index variants from ancestry-specific meta-analysis

Comparison of effect sizes and standard errors for variants reaching genome-wide significance (p-value < $5x10^{-8}$ as given by RAREMETAL) in both ancestry groups. Variants with discordant directions of effect between ancestries are labeled by chromosome and position (build 37). Association results for all index variants are given in Supplementary Table 3. The red line depicts an equivalent European ancestry and non-European ancestry effect size while the black line depicts a linear regression model. R^2=0.93



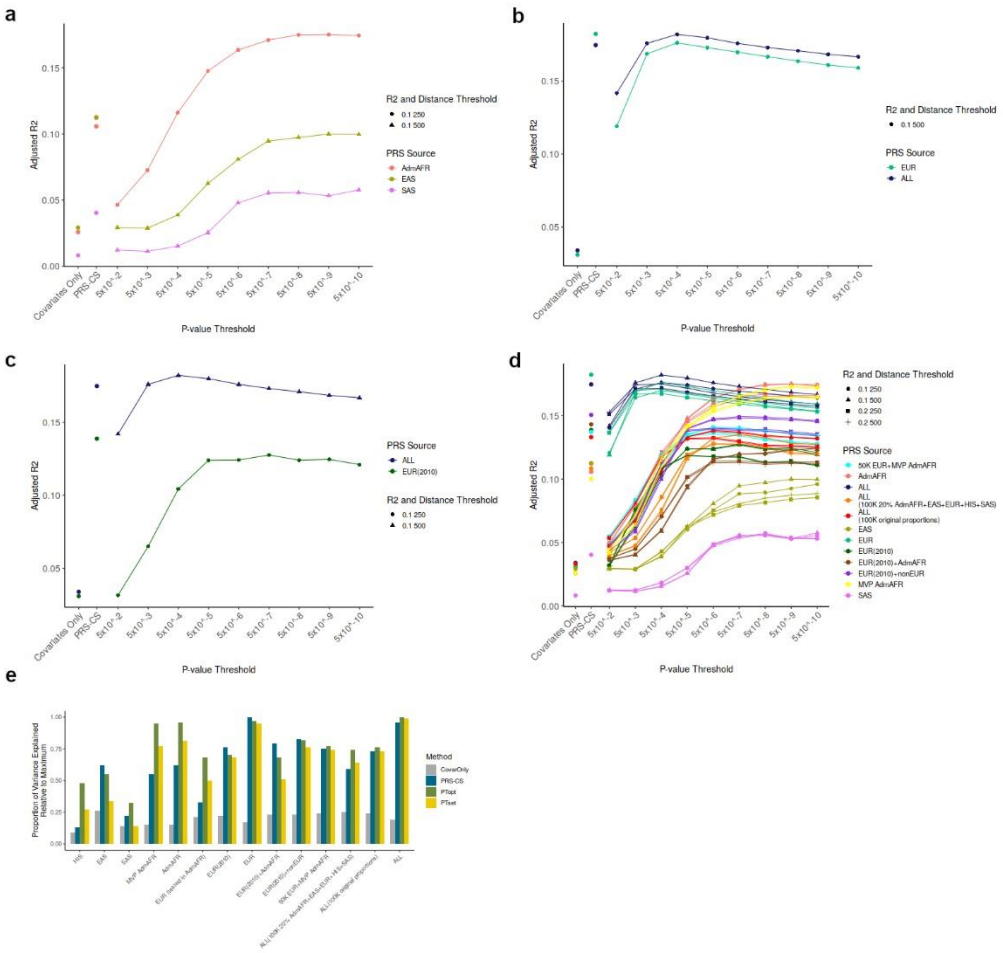
Extended Data Figure 4: Comparison of credible set size

The number of variants in the 99% credible sets for each association signal are compared between a) Admixed African ancestry and trans-ancestry analysis and b) European ancestry and trans-ancestry analysis



Extended Data Figure 5: Overview of LDL-C polygenic score generation and validation

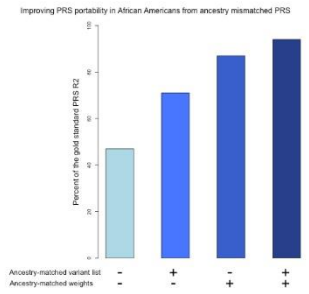
Polygenic scores were calculated separately in each ancestry group or in all ancestries combined using either pruning and thresholding or PRS-CS. The polygenic scores were then taken forward for testing in ancestry-matched participants followed by validation in independent data sets.



Extended Data Figure 6: Optimal polygenic score threshold by ancestry group for either PRS-CS or pruning and thresholding based LDL-C polygenic scores

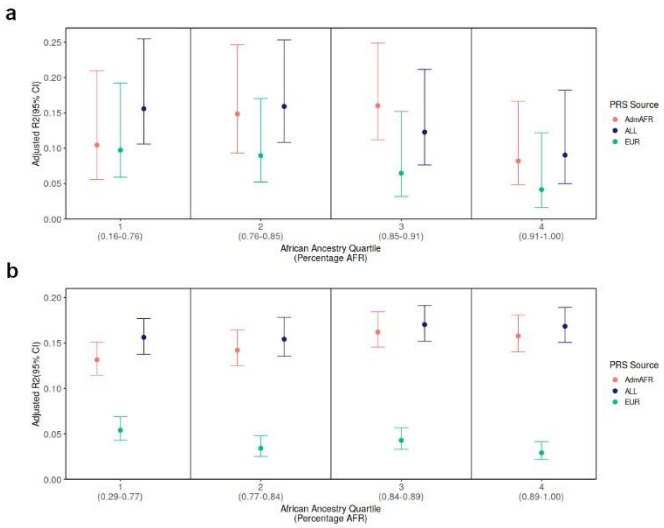
Adjusted R^2 estimated upon testing in UK Biobank ancestry-matched participants (not included in GWAS summary statistics).

a) Admixed African, East Asian and South Asian ancestry polygenic scores
b) European and trans-ancestry polygenic scores
c) European ancestry (GLGC 2010) and trans-ancestry polygenic scores
d) All polygenic scores across all thresholds used for score construction
e) Comparison of adjusted R^2 across ancestry groups relative to the maximum for covariates alone, polygenic scores from PRS-CS or polygenic scores from pruning and thresholding



Extended Data Figure 7: Comparison of PRS performance by admixture quartile

We divided the testing cohorts into quartiles by proportion of African ancestry and estimated the performance of the PRS separately within each quartile in a) the Michigan Genomics Initiative (N = 1,341) and b) in the Million Veteran Program (N = 18,251). Error bars represent 95% confidence intervals.



Extended Data Figure 8: Improvement in PRS performance in African Americans when starting with ancestry-mismatched European ancestry scores by updating weights, updating variant lists, or updating both variants and weights to be ancestry-matched.

By comparison to the gold-standard performance of the trans-ancestry-derived PRS in African Americans (adjusted $R^2 = 0.12$), a European ancestry derived score capture only 47% of the variance explained by the trans-ancestry PRS. When LD and association information from the target population is used to optimize the list of variants for inclusion in the PRS, but with ancestry-mismatched weights from European ancestry GWAS, the variance explained reaches 71% of the gold standard. If the PRS variant list selected in European ancestry individuals were genotyped in the target population, and PRS weights were updated using a GWAS from the target population, the variance explained reached 87% of the gold standard. Finally, deriving both the marker list and weights from the target population (single-ancestry GWAS) explained 94% of the variance relative to the gold-standard trans-ancestry PRS.