A multi-layer functional genomic analysis to understand noncoding genetic variation in lipids

Shweta Ramdas[1*], Jonathan Judd[2*], Sarah E Graham[3*], Stavroula Kanoni[4*], Yuxuan Wang[5*], Ida Surakka[3], Brandon Wenz[1], Shoa L Clarke[6,7], Alessandra Chesi[8], Andrew Wells[1], Konain Fatima Bhatti[4], Sailaja Vedantam[9,10], Thomas W Winkler[11], Adam E Locke[12], Eirini Marouli[4], Greg JM Zajac[13], Kuan-Han H Wu[14], Ioanna Ntalla[15], Qin Hui[16,17], Derek Klarin[18,19,20], Austin T Hilliard[6], Zeyuan Wang[16,17], Chao Xue[3], Gudmar Thorleifsson[21], Anna Helgadottir[21], Daniel F Gudbjartsson[21,22], Hilma Holm[21], Isleifur Olafsson[23], Mi Yeong Hwang[24], Sohee Han[24], Masato Akiyama[25,26], Saori Sakaue[27,28,29], Chikashi Terao[30], Masahiro Kanai[25,31,32], Wei Zhou[33,14,20], Ben M Brumpton[34,35,36], Humaira Rasheed[34,35,37], Aki S Havulinna[38,39], Yogasudha Veturi[40], Jennifer Allen Pacheco[41], Elisabeth A Rosenthal[42], Todd Lingren[43], QiPing Feng[44], Iftikhar J. Kullo[45], Akira Narita[46], Jun Takayama[46], Hilary C Martin[47], Karen A Hunt[48], Bhavi Trivedi[48], Jeffrey Haessler[49], Franco Giulianini[50], Yuki Bradford[40], Jason E Miller[40], Archie Campbell[51,52], Kuang Lin[53], Iona Y Millwood[53,54], Asif Rasheed[55], George Hindy[56], Jessica D Faul[57], Wei Zhao[58], David R Weir[57], Constance Turman[59], Hongyan Huang[59], Mariaelisa Graff[60], Ananyo Choudhury[61], Dhriti Sengupta[61], Anubha Mahajan[62], Michael R Brown[63], Weihua Zhang[64,65,66], Ketian Yu[13], Ellen M Schmidt[13], Anita Pandit[13], Stefan Gustafsson[67], Xianyong Yin[13], Jian'an Luan[68], Jing-Hua Zhao[69], Fumihiko Matsuda[70], Hye-Mi Jang[24], Kyungheon Yoon[24], Carolina Medina-Gomez[71], Achilleas Pitsillides[5], Jouke Jan Hottenga[72,73], Andrew R Wood[74], Yingji Ji[74], Zishan Gao[75,76,77], Simon Haworth[35,78], Ruth E Mitchell[35,79], Jin Fang Chai[80], Mette Aadahl[81,82], Anne A Bjerregaard[81], Jie Yao[83], Ani Manichaikul[84], PhD Wen-Jane[85], PhD Chao A[86], Helen R Warren[87,88], Julia Ramirez[87], Jette Bork-Jensen[89], Line L Kårhus[81], Anuj Goel[90,91], Maria Sabater-Lleal[92,93], Raymond Noordam[94], Pala Mauro[95], Floris Matteo[96,95], Aaron F McDaid[97,98], Pedro Marques-Vidal[99], Matthias Wielscher[100], Stella Trompet[101,102], Naveed Sattar[103], Line T Møllehave[81], Matthias Munz[104], Lingyao Zeng[105,106], Jianfeng Huang[107], Bin Yang[107], Alaitz Poveda[108], Azra Kurbasic[108],

Sebastian Schönherr[109], Lukas Forer[109], Markus Scholz[110,111], Tessel E. Galesloot[112], Jonathan P. Bradfield[113], Sanni E Ruotsalainen[38], E Warwick Daw[114], Joseph M Zmuda[115], Jonathan S Mitchell[116], Christian Fuchsberger[116], Henry Christensen[117], Jennifer A Brody[118], Phuong Le[119,120], Mary F Feitosa[114], Mary K Wojczynski[114], Daiane Hemerich[121], Michael Preuss[121], Massimo Mangino[122,123], Paraskevi Christofidou[122], Niek Verweij[124], Jan W Benjamins[124], Jorgen Engmann[125,126], Tsao L. Noah[127], Anurag Verma[1], Roderick C Slieker[128,129], Ken Sin Lo[130], Nuno R Zilhao[131], Marcus E Kleber[132,133], Graciela E Delgado[132], Shaofeng Huo[134], Daisuke D Ikeda[135], Hiroyuki Iha[135], Jian Yang[136,137,138], Jun Liu[139,140], Ayşe Demirkan[141,142], Hampton L Leonard[143,144], Jonathan Marten[145], Carina Emmel[146], Börge Schmidt[146], Laura J Smyth[147], Marisa Cañadas-Garre[147,148,149,150], Chaolong Wang[151,152], Masahiro Nakatochi[153], Andrew Wong[154], Nina Hutri-Kähönen[155], Xueling Sim[80], Rui Xia[156], Alicia Huerta-Chagoya[157], Juan Carlos Fernandez-Lopez[158], Valeriya Lyssenko[159,160], Suraj S Nongmaithem[161], Alagu Sankareswaran[161,162], Marguerite R Irvin[154], Christopher Oldmeadow[163], Han-Na Kim[164,165], Seungho Ryu[166,167], Paul RHJ Timmers[168,145], Liubov Arbeeva[169], Rajkumar Dorajoo[152,170], Leslie A Lange[171], Gauri Prasad[172,173], Laura Lorés-Motta[174], Marc Pauper[174], Jirong Long[175], Xiaohui Li[83], Elizabeth Theusch[176], Fumihiko Takeuchi[177], Cassandra N Spracklen[178,179], Anu Loukola[38], Sailalitha Bollepalli[38], Sophie C Warner[180,181], Ya Xing Wang[182], Wen B. Wei[183], Teresa Nutile[184], Daniela Ruggiero[184,185], Yun Ju Sung[186], Shufeng Chen[107], Fangchao Liu[107], Jingyun Yang[187,188], Katherine A Kentistou[168], Bernhard Banas[189], Anna Morgan[190], Karina Meidtner[191,192], Lawrence F Bielak[58], Jennifer A Smith[58,57], Prashantha Hebbar[193], Aliki-Eleni Farmaki[194,195], Edith Hofer[196,197], Maoxuan Lin[198], Maria Pina Concas[199], Simona Vaccargiu[200], Peter J van der Most[201], Niina Pitkänen[202,203], Brian E Cade[204,205], Sander W. van der Laan[206], Kumaraswamy Naidu Chitrala[207,208], Stefan Weiss[209], Amy R Bentley[210], Ayo P Doumatey[210], Adebowale A Adeyemo[210], Jong Young Lee[211], Eva RB Petersen[212], Aneta A Nielsen[213], Hyeok Sun Choi[214], Maria Nethander[215,216], Sandra Freitag-Wolf[217], Lorraine Southam[218,47], Nigel W Rayner[219,220,47,218], Carol A Wang[221], Shih-Yi Lin[222,223,224], Jun-Sing

Wang[225,226], Christian Couture[227], Leo-Pekka Lyytikäinen[228,229], Kjell Nikus[230,231], Gabriel Cuellar-Partida[232], Henrik Vestergaard[89,233], Bertha Hidalgo[234], Olga Giannakopoulou[4], Qiuyin Cai[175], Morgan O Obura[128], Jessica van Setten[235], Karen Y. He[236], Hua Tang[237], Natalie Terzikhan[238], Jae Hun Shin[214], Rebecca D Jackson[239], Alexander P Reiner[240], Lisa Warsinger Martin[241], Zhengming Chen[53,54], Liming Li[242], Takahisa Kawaguchi[70], Joachim Thiery[243,111], Joshua C Bis[118], Lenore J Launer[244], Huaixing Li[245], Mike A Nalls[143,144], Olli T Raitakari[246,247,248], Sahoko Ichihara[249], Sarah H Wild[250], Christopher P Nelson[180,181], Harry Campbell[168], Susanne Jäger[191,192], Toru Nabika[251], Fahd Al-Mulla[193], Harri Niinikoski[252,253], Peter S Braund[180,181], Ivana Kolcic[254], Peter Kovacs[255], Tota Giardoglou[256], Tomohiro Katsuya[257,258], Dominique de Kleijn[259], Gert J. de Borst[259], Eung Kweon Kim[260], Hieab H.H. Adams[238,261], M. Arfan Ikram[238], Xiaofeng Zhu[236], Folkert W Asselbergs[235], Adriaan O Kraaijeveld[235], Joline WJ Beulens[128,262], Xiao-Ou Shu[175], Loukianos S Rallidis[263], Oluf Pedersen[89], Torben Hansen[89], Paul Mitchell[264], Alex W Hewitt[265,266], Mika Kähönen[267,268], Louis Pérusse[227,269], Claude Bouchard[270], Anke Tönjes[271], Yii-Der Ida Chen[83], Craig E Pennell[221], Trevor A Mori[272], Wolfgang Lieb[273], Andre Franke[274], Claes Ohlsson[215,275], Dan Mellström[215,276], Yoon Shin Cho[214], Hyejin Lee[277], Jian-Min Yuan[278,279], Woon-Puay Koh[280,281], Sang Youl Rhee[282], Jeong-Taek Woo[282], Iris M Heid[11], Klaus J Stark[11], Martina E Zimmermann[11], Henry Völzke[283], Georg Homuth[209], Michele K Evans[284], Alan B Zonderman[284], Ozren Polasek[254,285], Gerard Pasterkamp[206], Imo E Hoefer[206], Susan Redline[204,205], Katja Pahkala[202,203,286], Albertine J Oldehinkel[287], Harold Snieder[201], Ginevra Biino[288], Reinhold Schmidt[196], Helena Schmidt[289], Stefania Bandinelli[290], George Dedoussis[194], Thangavel Alphonse Thanaraj[193], Patricia A Peyser[58], Norihiro Kato[177], Matthias B Schulze[191,192,291], Giorgia Girotto[292,199], Carsten A Böger[189,293,294], Bettina Jung[189,293,294], Peter K Joshi[168], David A Bennett[187,188], Philip L De Jager[295,296], Xiangfeng Lu[107], Vasiliki Mamakou[297,298], Morris Brown[299,88], Mark J Caulfield[87,88], Patricia B Munroe[87,88], Xiuqing Guo[83], Marina Ciullo[184,185], Jost B. Jonas[300,182,301,302], Nilesh J Samani[180,181], Jaakko Kaprio[38], Päivi Pajukanta[303], Teresa Tusié-Luna[304,305], Carlos A Aguilar-Salinas[306], Linda S Adair[307,308], Sonny Augustin Bechayda[309,310], H.

Janaka de Silva[311], Ananda R Wickremasinghe[312], Ronald M Krauss[313], Jer-Yuarn Wu[314], Wei Zheng[175], Anneke I den Hollander[174], Dwaipayan Bharadwaj[173,315], Adolfo Correa[316], James G Wilson[317], Lars Lind[318], Chew-Kiat Heng[319], Amanda E Nelson[169,320], Yvonne M Golightly[169,321,322,323], James F Wilson[168,145], Brenda Penninx[324,325], Hyung-Lae Kim[326], John Attia[327,163], Rodney J Scott[327,163], D C Rao[328], Donna K Arnett[329], Mark Walker[330], Laura J Scott[13], Heikki A Koistinen[39,331,332], Giriraj R Chandak[161,162,333], Josep M Mercader[334,335,336], Teresa Tusie-Luna[337], Carlos Aguilar-Salinas[338], Clicerio Gonzalez Villalpando[339], Lorena Orozco[340], Myriam Fornage[156,341], E Shyong Tai[342,80], Rob M van Dam[80,342], Terho Lehtimäki[228,229], Nish Chaturvedi[343], Mitsuhiro Yokota[344], Jianjun Liu[152], Dermot F Reilly[345], Amy Jayne McKnight[147], Frank Kee[147], Karl-Heinz Jöckel[146], Mark I McCarthy[62,346], Colin NA Palmer[347], Veronique Vitart[145], Caroline Hayward[145], Eleanor Simonsick[348], Cornelia M van Duijn[139,140], Zi-Bing Jin[349,350], Fan Lu[350], Haretsugu Hishigaki[135], Xu Lin[134], Winfried März[351,352,132], Vilmundur Gudnason[131,353], Jean-Claude Tardif[130], Guillaume Lettre[130], Leen M t Hart[354,355,356], Petra JM Elders[357], Daniel J Rader[358], Scott M Damrauer[359,360], Meena Kumari[361], Mika Kivimaki[126], Pim van der Harst[362], Tim D Spector[122], Ruth J.F. Loos[121,363,89], Michael A Province[114], Esteban J Parra[119], Miguel Cruz[364], Bruce M Psaty[118,365,366], Ivan Brandslund[117,367], Peter P Pramstaller[116], Charles N Rotimi[368], Kaare Christensen[369], Samuli Ripatti[38,370,371], Elisabeth Widén[38], Hakon Hakonarson[372,373], Struan F.A. Grant[373,374,375], Lambertus ALM Kiemeney[376], Jacqueline de Graaf[376], Markus Loeffler[110,111], Florian Kronenberg[109], Dongfeng Gu[107,377], Jeanette Erdmann[378], Heribert Schunkert[105,106], Paul W Franks[108], Allan Linneberg[81,82], J. Wouter Jukema[101,379], Amit V Khera[380,381,382,383], Minna Männikkö[384], Marjo-Riitta Jarvelin[100,385,386], Zoltan Kutalik[97,98], Cucca Francesco[387,388], Dennis O Mook-Kanamori[389,390], Ko Willems van Dijk[391,392,393], Hugh Watkins[90,91], David P Strachan[394], Niels Grarup[89], Peter Sever[395], Neil Poulter[396], Wayne Huey-Herng Sheu[397,398], Jerome I Rotter[83], Thomas M Dantoft[81], Fredrik Karpe[399,400], Matt J Neville[399,400], Nicholas J Timpson[35,79], Ching-Yu Cheng[401,402], Tien-Yin Wong[401,402], Chiea Chuen Khor[152], Hengtong Li[403], Charumathi Sabanayagam[401,402], Annette Peters[77,404,405],

Christian Gieger[76,77,405], Andrew T Hattersley[406], Nancy L Pedersen[407], Patrik KE Magnusson[407], Dorret I Boomsma[72,408], Eco JC de Geus[72,408], L Adrienne Cupples[5,409], Joyce B.J. van Meurs[71,140], Arfan Ikram[140], Mohsen Ghanbari[140,410], Penny Gordon-Larsen[307,308], Wei Huang[411], Young Jin Kim[24], Yasuharu Tabara[70], Nicholas J Wareham[68], Claudia Langenberg[68], Eleftheria Zeggini[218,47,412], Jaakko Tuomilehto[39,413,414], Johanna Kuusisto[415], Markku Laakso[415], Erik Ingelsson[7,416,417,418], Goncalo Abecasis[13,419], John C Chambers[420,64,65,421], Jaspal S Kooner[65,66,422,423], Paul S de Vries[63], Alanna C Morrison[63], Scott Hazelhurst[61,424], Michèle Ramsay[61], Kari E. North[425], Martha Daviglus[426], Peter Kraft[59,427], Nicholas G Martin[428], John B Whitfield[428], Shahid Abbas[429], Danish Saleheen[55,430,431], Robin G Walters[53,54,432], Michael V Holmes[53,54,433], Corri Black[434], Blair H Smith[435], Aris Baras[419], Anne E Justice[436], Julie E Buring[50,437], Paul M Ridker[50,437], Daniel I Chasman[50,437], Charles Kooperberg[49], Gen Tamiya[46], Masayuki Yamamoto[46], David A van Heel[48], Richard C Trembath[438], Wei-Qi Wei[439], Gail P Jarvik[440], Bahram Namjou[441], M. Geoffrey Hayes[442,443,444], Marylyn D Ritchie[40], Pekka Jousilahti[39], Veikko Salomaa[39], Kristian Hveem[34,445,446], Bjørn Olav Åsvold[34,445,447], Michiaki Kubo[448], Yoichiro Kamatani[25,449], Yukinori Okada[27,25,450,451], Yoshinori Murakami[452], Bong-Jo Kim[453], Unnur Thorsteinsdottir[21,454], Kari Stefansson[21,454], Jifeng Zhang[3], Y Eugene Chen[3], Yuk-Lam Ho[455], Julie A Lynch[456,457], Daniel Rader[458], Philip S Tsao[6,7,459], Kyong-Mi Chang[460,458], Kelly Cho[455,461], Christopher J O'Donnell[455,461], John M Gaziano[455,461], Peter Wilson[17,462], Karen L Mohlke[178], Timothy M Frayling[74], Joel N Hirschhorn[463,10,464], Sekar Kathiresan[465,381,383], Michael Boehnke[13], Million Veterans Program, Global Lipids Genetics Consortium, Struan Grant[1,374], Pradeep Natarajan[466,467,20,468], Yan V Sun[16,17], Andrew P Morris[469], Panos Deloukas[4,470], Gina Peloso[471], Themistocles L Assimes[6,7,459], Cristen J Willer[3,472,14$], Xiang Zhu[473,474,475,476], Christopher D Brown[1]

[1]Department of Genetics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA, [2]Department of Genetics, Stanford University School of Medicine, Stanford, CA, USA, [3]Department of Internal Medicine, Division of Cardiology, University of Michigan, Ann Arbor, MI, 48109, USA, [4]William Harvey Research Institute, Barts and the London School of Medicine and Dentistry, Queen Mary University of London, Charterhouse square, EC1M 6BQ, UK, [5]Department of Biostatistics, Boston University School of Public Health, 801 Massachusetts Ave, Boston, MA 02118, USA, [6]VA Palo Alto Health Care Systems, Palo Alto, California, USA, [7]Department of Medicine, Division of Cardiovascular Medicine, Stanford University School of Medicine, Stanford, CA 94305, USA, [8]Department of Pathology and Laboratory Medicine, University of Pennsylvania, Philadelphia, PA, USA, [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, USA, [13]Department of Biostatistics, Center for Statistical Genetics, University of Michigan, Ann Arbor, MI, USA, [14]Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI, USA, [15]Clinical Pharmacology, William Harvey Research Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, EC1M 6BQ UK, [16]Department of Epidemiology, Emory University Rollins School of Public Health, Atlanta, Georgia, USA, [17]Atlanta VA Health Care System, Decatur, GA, USA, [18]Malcolm Randall VA Medical Center, Gainesville, FL, USA, [19]Division of Vascular Surgery and Endovascular Therapy, University of Florida College of Medicine, Gainesville, FL, USA, [20]Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, Cambridge, MA, USA, [21]deCODE genetics/Amgen, Inc. Sturlugata 8, Reykjavik, 102, Iceland, [22]School of

149 Engineering and Natural Sciences, University of Iceland, Sæmundargötu 2, Reykjavik, 102,
150 Iceland, [23]Department of Clinical Biochemistry, Landspitali - National University Hospital of
151 Iceland, Hringbraut, Reykjavik, 101, Iceland, [24]Division of Genome Science, Department of
152 Precision Medicine, National Institute of Health, Chungcheongbuk-do, South Korea,
153 [25]Laboratory for Statistical Analysis, RIKEN Center for Integrative Medical Sciences,
154 Yokohama, Japan, [26]Department of Ophthalmology, Graduate School of Medical Sciences,
155 Kyushu University, Fukuoka, Japan, [27]Department of Statistical Genetics, Osaka University
156 Graduate School of Medicine, Osaka, Japan, [28]Laboratory for Statistical Analysis, RIKEN
157 Center for Integrative Medical Sciences, [29]Department of Allergy and Rheumatology,
158 Graduate School of Medicine, The University of Tokyo, Tokyo, Japan, [30]Laboratory for
159 Statistical and Translational Genetics, RIKEN Center for Integrative Medical Sciences,
160 Yokohama, Japan, [31]Program in Medical and Population Genetics, Broad Institute of MIT and
161 Harvard, Cambridge, MA, USA., [32]Department of Biomedical Informatics, Harvard Medical
162 School, Boston, MA, USA, [33]Analytic and Translational Genetics Unit, Massachusetts
163 General Hospital, Boston, Massachusetts, USA, [34]K.G. Jebsen Center for Genetic
164 Epidemiology, Department of Public Health and Nursing, NTNU, Norwegian University of
165 Science and Technology, Trondheim, Norway, [35]MRC Integrative Epidemiology Unit (IEU),
166 Bristol Medical School, University of Bristol, Oakfield House, Oakfield Grove, Bristol, BS8
167 2BN, UK, [36]Clinic of Medicine, St. Olavs Hospital, Trondheim University Hospital,
168 Trondheim, Norway, [37]Division of Medicine and Laboratory Sciences, University of
169 Oslo,Norway, [38]Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of
170 Helsinki, Tukholmankatu 8, 00014 Helsinki, Finland, [39]Department of Public Health and
171 Welfare, Finnish Institute for Health and Welfare, Helsinki, Finland, [40]Department of
172 Genetics, Institute for Biomedical Informatics, University of Pennsylvania, Perelman School
173 of Medicine, Philadelphia, PA 19104, USA, [41]Center for Genetic Medicine, Northwestern

University, Feinberg School of Medicine, Chicago, IL 60611, USA, [42]Department of Medicine (Medical Genetics), University of Washington, WA, USA, [43]Division of Biomedical Informatics, Cincinnati Children's Hospital Medical Center, OH, USA, [44]Division of Clinical Pharmacology, Department of Medicine, Vanderbilt University Medical Center, Nashville, TN, USA, [45]Department of Cardiovascular Medicine and the Gonda Vascular Center, Mayo Clinic, Rochester, MN, USA, [46]Tohoku Medical Megabank Organization, Tohoku University, Sendai 980-8573, Japan, [47]Wellcome Trust Sanger Institute, Hinxton, CB10 1SA, UK, [48]Blizard Institute, Barts and the London School of Medicine and Dentistry, Queen Mary University of London, London, UK, [49]Fred Hutchinson Cancer Research Center, Division of Public Health Sciences, Seattle WA 9810, USA, [50]Division of Preventive Medicine, Brigham and Women's Hospital, Boston, MA 02215, USA, [51]Centre for Genomic and Experimental Medicine , Institute of Genetics and Cancer, University of Edinburgh, Western General Hospital, Edinburgh EH4 2XU, United Kingdom, [52]Usher Institute, The University of Edinburgh, Nine, Edinburgh Bioquarter, 9 Little France Road, Edinburgh, EH16 4UX, UK, [53]Clinical Trial Service Unit and Epidemiological Studies Unit, Nuffield Department of Population Health, University of Oxford, Oxford OX3 7LF, UK, [54]Medical Research Council Population Health Research Unit, Nuffield Department of Population Health, University of Oxford, Oxford OX3 7LF, UK, [55]Center for Non-Communicable Diseases, Karachi, Sindh, Pakistan, [56]Department of Population Medicine, Qatar University College of Medicine, QU Health, Doha, Qatar, [57]Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor, MI, 48104, USA, [58]Department of Epidemiology, School of Public Health, University of Michigan, Ann Arbor, MI, 48109, USA, [59]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, [60]Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina at

Chapel Hill, NC USA, [61]Sydney Brenner Institute for Molecular Bioscience, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa, [62]Wellcome Centre for Human Genetics, University of Oxford, UK, [63]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, [64]Department of Epidemiology and Biostatistics, Imperial College London, London W2 1PG, UK, [65]Department of Cardiology, Ealing Hospital, London North West University Healthcare NHS Trust, Middlesex UB1 3HW, UK, [66]Imperial College Healthcare NHS Trust, London W12 0HS, UK, [67]Department of Medical Sciences, Molecular Epidemiology and Science for Life Laboratory, Uppsala University, Uppsala, Sweden, [68]MRC Epidemiology Unit, University of Cambridge School of Clinical Medicine, Cambridge, CB2 0QQ, UK, [69]Cardiovascular Epidemiology Unit, Department of Public Health and Primary Care, University of Cambridge, Strangeways Research Laboratory, Wort's Causeway, Cambridge CB1 8RN, UK, [70]Center for Genomic Medicine, Kyoto University Graduate School of Medicine, Kyoto, Japan, [71]Department of Internal Medicine, Erasmus MC, University Medical Center Rotterdam, the Netherlands, [72]Department of Biological Psychology, Vrije Universiteit Amsterdam, The Netherlands, [73]Amsterdam Public Health Research Institute, Amsterdam UMC, the Netherlands, [74]Genetics of Complex Traits, University of Exeter Medical School, University of Exeter, Exeter, EX2 5DW, UK, [75]Department of Clinical Acupuncture and Moxibustion, Nanjing University of Chinese Medicine, Nanjing, Jiangsu 210029, China, [76]Research Unit of Molecular Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany, [77]Institute of Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany, [78]Bristol Dental School, University of Bristol, Lower Maudlin Street, Bristol BS1 2LY, United Kingdom, [79]Population Health Sciences, Bristol Medical School,

University of Bristol, Oakfield Grove, Bristol, BS8 2BN, United Kingdom, [80]Saw Swee Hock School of Public Health, National University of Singapore and National University Health System, 117549, Singapore, [81]Center for Clinical Research and Prevention, Bispebjerg and Frederiksberg Hospital, Copenhagen, Denmark, [82]Department of Clinical Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark, [83]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, [84]Center for Public Health Genomics, University of Virginia, Charlottesville, VA 22903 USA, [85]Department of Medical Research, Taichung Veterans General Hospital, Taichung, Taiwan; No. 1650, Sec. 4, Taiwan Boulevard, Taichung City 40705, Taiwan, [86]Institute of Population Health Sciences, National Health Research Institutes, 35 Keyan Road, Zhunan Town, Miaoli County 350, Taiwan, ROC, [87]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, [88]NIHR Barts Cardiovascular Biomedical Research Centre, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, EC1M 6BQ, UK, [89]Novo Nordisk Foundation Center for Basic Metabolic Research, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark, [90]Division of Cardiovascular Medicine, Radcliffe Department of Medicine, John Radcliffe Hospital, University of Oxford, Oxford. UK. OX3 9DU, [91]Wellcome Centre for Human Genetics, University of Oxford, Oxford. UK. OX3 7BN, [92]Unit of Genomics of Complex Diseases. Sant Pau Biomedical Research Institute (IIB Sant Pau), Barcelona, Spain, [93]Cardiovascular Medicine Unit, Department of Medicine, Karolinska Institutet, Center for Molecular Medicine, Karolinska University Hospital, Stockholm, Sweden, [94]Department of Internal Medicine, Section of Gerontology and Geriatrics, Leiden University Medical Center,

Leiden, the Netherlands, [95]Istituto di Ricerca Genetica e Biomedica, Consiglio Nazionale delle Ricerche, Italy, [96]Dipartimento di Scienze Biomediche, Università degli Studi di Sassari, Sardinia, Italy, [97]University Center for Primary Care and Public Health, University of Lausanne, Rte de Berne 113, Lausanne, 1010, Switzerland, [98]Swiss Institute of Bioinformatics, Lausanne, 1015, Switzerland, [99]Department of Medicine, Internal Medicine, Lausanne University Hospital and University of Lausanne, Rue du Bugnon 46, Lausanne, 1011, Switzerland, [100]Department of Epidemiology and Biostatistics, MRC-PHE Centre for Environment and Health, School of Public Health, Imperial College London, London, UK, [101]Dept of Cardiology, Leiden University Medical Center, Leiden, the Netherlands, [102]Dept of Internal Medicine, Section of Gerontology and Geriatrics, Leiden university Medical Center, Leiden, the Netherlands, [103]BHF Glasgow Cardiovascular Research Centre, Faculty of Medicine, Glasgow, United Kingdom, [104]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, [105]Deutsches Herzzentrum München, Klinik für Herz- und Kreislauferkrankungen, Technische Universität München,Munich, Germany., [106]Deutsches Zentrum für Herz-Kreislauf-Forschung (DZHK) e.V., partner site Munich Heart Alliance,Munich, Germany., [107]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, [108]Lund University Diabetes Centre, Malmö, Sweden, [109]Institute of Genetic Epidemiology, Department of Genetics and Pharmacology, Medical University of Innsbruck,

Innsbruck, Austria and German Chronic Kidney Disease Study, Austria, [110]Institute for Medical Informatics, Statistics and Epidemiology, University of Leipzig, Haertelstrasse 16-18, 04107 Leipzig, Germany, [111]LIFE Research Centre for Civilization Diseases, University of Leipzig, Philipp-Rosenthal-Straße 27, 04103 Leipzig, Germany, [112]Radboud university medical center, Radboud Institute for Health Sciences, Nijmegen, the Netherlands, [113]Quantinuum Research LLC, Wayne, PA, 19087 USA, [114]Division of Statistical Genomics, Department of Genetics; Washington University School of Medicine; St. Louis, MO, USA, [115]Department of Epidemiology; University of Pittsburgh; Pittsburgh, PA, USA, [116]Institute for Biomedicine, Eurac Research, Affiliated Institute of the University of Lübeck, Via Galvani 31, 39100, Bolzano, Italy, [117]Department of Clinical Biochemistry, Lillebaelt Hospital, Vejle, Denmark, [118]Cardiovascular Health Research Unit, Department of Medicine, University of Washington, Seattle, 98101, USA, [119]Department of Anthropology, University of Toronto at Mississauga, Mississauga, ON L5L 1C6, Canada, [120]Department of Computer Science, University of Toronto, Toronto, ON M5S 2E4, Canada, [121]The Charles Bronfman Institute for Personalized Medicine, Icahn School of Medicine at Mount Sinai, New York, NY, USA, [122]Department of Twin Research and Genetic Epidemiology, King's College London, London SE1 7EH, UK, [123]NIHR Biomedical Research Centre at Guy's and St Thomas' Foundation Trust, London SE1 9RT, UK, [124]Department of Cardiology, University of Groningen, University Medical Center Groningen, 9700RB Groningen, the Netherlands, [125]Institute of Cardiovascular Sciences, University College London, Gower Street, WC1E 6BT London, UK, [126]Department of Epidemiology and Public Health, University College London, 1-19 Torrington Place, WC1E 6BT London, United Kingdom, [127]Department of Surgery, University of Pennsylvania, Philadelphia, PA, [128]Amsterdam UMC, Department of Epidemiology and Data Science, Amsterdam Public Health Research Institute, Amsterdam, 1081HV, the Netherlands, [129]Leiden University Medical Center, Department of Cell and Chemical Biology,

299 Leiden, 2333ZA, the Netherlands, [130]Montreal Heart Institute, Université de Montréal, 5000
300 Belanger street, Montreal, PQ, Canada H1T1C8, [131]Icelandic Heart Association, 201
301 Kopavogur, Iceland, [132]Vth Department of Medicine, Medical Faculty Mannheim, Heidelberg
302 University, 68167 Mannheim, Germany, [133]SYNLAB MVZ Humangenetik Mannheim GmbH,
303 68163 Mannheim, Germany, [134]Shanghai Institute of Nutrition and Health, University of
304 Chinese Academy of Sciences, Chinese Academy of Sciences, Shanghai, China, [135]Biomedical
305 Technology Research Center, Tokushima Research Institute, Otsuka Pharmaceutical Co.,
306 Ltd., Tokushima, Japan, [136]School of Life Sciences, Westlake University, Hangzhou, Zhejiang
307 310024, China, [137]Westlake Laboratory of Life Sciences and Biomedicine, Hangzhou,
308 Zhejiang 310024, China, [138]Institute for Molecular Bioscience, The University of Queensland,
309 Brisbane, Queensland 4072, Australia, [139]Nuffield Department of Population Health,
310 University of Oxford, Oxford, United Kingdom, [140]Department of Epidemiology, Erasmus
311 MC, University Medical Center Rotterdam, the Netherlands, [141]Department of Epidemiology,
312 Erasmus MC, University Medical Center, Rotterdam, the Netherlands, [142]Section of Statistical
313 Multi-omics, Department of Clinical and Experimental research, University of Surrey,
314 Guildford, Surrey, UK, [143]Laboratory of Neurogenetics, National Institute on Aging, NIH,
315 Bethesda MD, USA, [144]Data Tecnica International, Glen Echo MD, USA, [145]MRC Human
316 Genetics Unit, Institute of Genetics and Cancer, University of Edinburgh, Western General
317 Hospital, Crewe Road, Edinburgh, EH4 2XU, Scotland, [146]Institute for Medical Informatics,
318 Biometrie and Epidemiology, University of Duisburg-Essen, Essen, Germany, [147]Centre for
319 Public Health, Queen's University of Belfast, Northern Ireland, [148]Genomic Oncology Area,
320 GENYO, Centre for Genomics and Oncological Research: Pfizer-University of Granada-
321 Andalusian Regional Government, Granada, Spain, [149]Hematology Department, Hospital
322 Universitario Virgen de las Nieves, Granada, Spain, [150]Instituto de Investigación Biosanitaria
323 de Granada (ibs.GRANADA), Granada, Spain, [151]Department of Epidemiology and

Biostatistics, School of Public Health, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China, [152]Genome Institute of Singapore, Agency for Science, Technology and Research, Singapore, [153]Public Health Informatics Unit, Department of Integrated Health Sciences, Nagoya University Graduate School of Medicine, Nagoya, 461-8673, Japan, [154]University of Alabama at Birmingham, Epidemiology, School of Public Health, AL, USA, [155]Tampere Centre for Skills Training and Simulation, Faculty of Medicine and Health Technology, Tampere University, Tampere, 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]Academy of Scientific and Innovative Research (AcSIR), New Delhi, India, [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]Health Services and Systems Research,

Duke-NUS Medical School, 169857, Singapore, [171]Division of Biomedical Informatics and Personalized Medicine, Department of Medicine, Anschutz Medical Campus, University of Colorado, Denver, Aurora, CO 80045, USA, [172]Genomics and Molecular Medicine Unit, CSIR-Institute of Genomics and Integrative Biology, New Delhi - 110020, India, [173]Academy of Scientific and Innovative Research, CSIR-Human Resource Development Centre, Ghaziabad, Uttar Pradesh, India, [174]Departments of Ophthalmology and Human Genetics, Radboud University Nijmegen Medical Center, Philips van Leydenlaan 15, Nijmegen, 6525 EX, the Netherlands, [175]Vanderbilt Epidemiology Center, Division of Epidemiology, Vanderbilt University Medical Center, USA, [176]Department of Pediatrics, University of California San Francisco, Oakland, CA 94609 USA, [177]National Center for Global Health and Medicine, Tokyo, 1628655, Japan, [178]Department of Genetics, University of North Carolina, Chapel Hill, NC 27599 USA, [179]Department of Biostatistics and Epidemiology, University of Massachusetts-Amherst, Amherst, MA 01003 USA, [180]Department of Cardiovascular Sciences, University of Leicester, Leicester, UK, [181]NIHR Leicester Biomedical Research Centre, Glenfield Hospital, Leicester, UK, [182]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, [183]Beijing Tongren Eye Center, Beijing Tongren Hospital, Capital Medical University, 1 Dong Jiao Min Xiang, Dong Cheng District, Beijing, 100730, China, [184]Institute of Genetics and Biophysics “Adriano Buzzati-Traverso" - CNR, Naples, Italy, [185]IRCCS Neuromed, Pozzilli, Isernia, Italy, [186]Division of Biostatistics, Washington University, St. Louis, MO 63110, USA, [187]Rush Alzheimer’s Disease Center, Rush University Medical Center, IL, USA, [188]Department of Neurological Sciences, Rush University Medical Center, IL, USA, [189]Dept of Nephrology, University Hospital Regensburg, Regensburg, Germany, [190]Institute for Maternal and Child Health—IRCCS, Burlo Garofolo, 34127 Trieste, Italy,

[191]Department of Molecular Epidemiology, German Institute of Human Nutrition Potsdam-Rehbruecke, Nuthetal, Germany, [192]German Center for Diabetes Research (DZD), München-Neuherberg, Germany, [193]Department of Genetics and Bioinformatics, Dasman Diabetes Institute, Kuwait, [194]Department of Nutrition and Dietetics, School of Health Science and Education, Harokopio University of Athens, Athens, Greece, [195]Department of Population Science and Experimental Medicine, University College London, London, UK, [196]Clinical Division of Neurogeriatrics, Department of Neurology, Medical University of Graz, Graz, Austria, [197]Institute for Medical Informatics, Statistics and Documentation, Medical University of Graz, Graz, Austria, [198]Massachusetts General Hospital Cancer Center, Charlestown, MA 02129, USA, [199]Institute for Maternal and Child Health, IRCCS Burlo Garofolo, Trieste, Italy, [200]Institute of Genetic and Biomedical Research, National Research Council of Italy, UOS of Sassari, Sassari, Italy, [201]Department of Epidemiology, University of Groningen, University Medical Center Groningen, Groningen, 9700 RB, the Netherlands, [202]Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, Turku, Finland, [203]Centre for Population Health Research, University of Turku and Turku University Hospital, Turku, Finland, [204]Sleep Medicine and Circadian Disorders, Brigham and Women's Hospital, Boston, Massachusetts 02115, USA, [205]Division of Sleep Medicine, Harvard Medical School, Boston, Massachusetts 02115, USA, [206]Central Diagnostics Laboratory, Division Laboratories, Pharmacy, and Biomedical genetics, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands, [207]Laboratory of Epidemiology and Population Science National Institute on Aging Intramural Research Program, NIH, USA, [208]Fels Cancer Institute for Personalized Medicine, Temple University Lewis Katz School of Medicine, Philadelphia, PA, USA, [209]Interfaculty Institute for Genetics and Functional Genomics, Department of Functional Genomics, University of Greifswald and University Medicine Greifswald, Greifswald, Germany, [210]Center for Research on Genomics and Global Health, National

399 Human Genome Research Institute, National Institutes of Health, 12 South Drive, Room
400 4047, Bethesda, MD, 20892, USA, [211]Oneomics. co. ltd. 2F, Soonchunhyang Mirai Medical
401 Center 173, Buheuyng-ro, Bucheon-si Gyeonggi-do, 14585, Korea, [212]Department of Clinical
402 Biochemistry and Immunology, Hospital of Southern Jutland, Kresten Philipsens Vej 15,
403 6200 Aabenraa, Denmark, [213]Department of Clinical Biochemistry, Lillebaelt Hospital,
404 Kolding, Denmark, [214]Department of Biomedical Science, Hallym University, Chuncheon,
405 Gangwon-do 24252, Korea, [215]Centre for Bone and Arthritis Research, Department of Internal
406 Medicine and Clinical Nutrition, Institute of Medicine, Sahlgrenska Academy, University of
407 Gothenburg, Gothenburg, Sweden, [216]Bioinformatics Core Facility, Sahlgrenska Academy,
408 University of Gothenburg, Gothenburg, Sweden, [217]Institute of Medical Informatics and
409 Statistics, Kiel University, Kiel, Germany, [218]Institute of Translational Genomics, Helmholtz
410 Zentrum München – German Research Center for Environmental Health, Neuherberg,
411 Germany, [219]Wellcome Centre for Human Genetics, University of Oxford, Oxford, UK,
412 [220]Oxford Centre for Diabetes Endocrinology and Metabolism, Oxford, UK, [221]School of
413 Medicine and Public Health, Faculty of Medicine and Health, University of Newcastle,
414 Newcastle, New South Wales, 2308, Australia, [222]Center for Geriatrics and Gerontology,
415 Division of Endocrinology and Metabolism, Department of Internal Medicine, Taichung
416 Veterans General Hospital, Taichung, Taiwan, [223]School of Medicine, National Yang-Ming
417 University, Taipei, Taiwan, [224]School of Medicine, National Defense Medical Center, Taipei,
418 Taiwan, [225]Division of Endocrinology and Metabolism, Department of Internal Medicine,
419 Taichung Veterans General Hospital, Taichung, Taiwan, [226]Department of Medicine, School
420 of Medicine, National Yang-Ming University, Taipei, Taiwan, [227]Department of Kinesiology,
421 Université Laval, Québec, Canada, [228]Department of Clinical Chemistry, Fimlab Laboratories,
422 Tampere 33520, Finland, [229]Department of Clinical Chemistry, Finnish Cardiovascular
423 Research Center - Tampere, Faculty of Medicine and Health Technology, Tampere

University, Tampere 33014, Finland, [230]Department of Cardiology, Heart Center, Tampere University Hospital, Tampere 33521, Finland, [231]Department of Cardiology, Finnish Cardiovascular Research Center - Tampere, Faculty of Medicine and Health Technology, Tampere University, Tampere 33014, Finland, [232]University of Queensland Diamantina Institute, Translational Research Institute, Kent St, Woolloongabba, Brisbane, QLD, 4102, Australia, [233]Department of Medicine, Bornholms Hospital, Rønne, Denmark, [234]School of Public Health, University of Alabama at Birmingham, AL, USA, [235]Cardiology, Division Heart & Lungs, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands, [236]Department of Population and Quantitative Health Sciences, Case Western Reserve University, Cleveland, OH, 44106, USA, [237]Department of Genetics, Stanford University School of Medicine Stanford, CA 94305, USA, [238]Department of Epidemiology - Erasmus MC - University Medical Center Rotterdam, Rotterdam, the Netherlands, [239]Ohio State University, Division of Endricinology, Columbus OH 43210, USA, [240]University of Washington, Department of Epidemiology, Seattle WA 98195, USA, [241]George Washington University, School of Medicine and Health Sciences, Washington DC 20037, USA, [242]Department of Epidemiology, School of Public Health, Peking University Health Science Center, Beijing, China, [243]Institute for Laboratory Medicine, University Hospital Leipzig, Paul-List-Strasse 13/15, 04103 Leipzig, Germany, [244]Laboratory of Epidemiology and Population Sciences, National Institute on Aging, NIH, Baltimore,MD, 20892-9205, USA, [245]Shanghai Institute of Nutrition and Health, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Shanghai, China, [246]Centre for Population Health Research, University of Turku and Turku University Hospital, Finland, [247]Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, Finland, [248]Department of Clinical Physiology and Nuclear Medicine, Turku University Hospital, Turku, Finland, [249]Department of Environmental and Preventive Medicine, Jichi Medical University School of

449 Medicine, Shimotsuke, 329-0498, Japan, [250]Centre for Population Health Sciences, Usher
450 Institute, University of Edinburgh, Teviot Place, Edinburgh, EH8 9AG, Scotland,
451 [251]Department of Functional Pathology, Shimane University School of Medicine, Izumo,
452 6938501, Japan, [252]Department of Pediatrics and Adolescent Medicine, Turku University
453 Hospital and University of Turku, Turku, Finland, [253]Department of Physiology, University of
454 Turku, Turku, Finland, [254]Faculty of Medicine, University of Split, Šoltanska 2, HR-21000,
455 Split, Croatia, [255]Medical Department III – Endocrinology, Nephrology, Rheumatology,
456 University of Leipzig Medical Center, Liebigstr. 21, 04103 Leipzig, Germany, [256]Department
457 of Nutrition-Dietetics, Harokopio University, Eleftheriou Venizelou, Athens, 17676, Greece,
458 [257]Department of Clinical Gene Therapy, Osaka University Graduate School of Medicine,
459 Suita, 5650871, Japan, [258]Department of Geriatric and General Medicine, Osaka University
460 Graduate School of Medicine, Suita, 5650871, Japan, [259]Department of Vascular Surgery,
461 Division of Surgical Specialties, University Medical Center Utrecht, Utrecht University,
462 Utrecht, the Netherlands, [260]Corneal Dystrophy Research Institute, Department of
463 Ophthalmology, Yonsei University College of Medicine, Seoul 03722, Korea, [261]Dept of
464 Radiology and Nuclear Medicine, Erasmus MC - University Medical Center Rotterdam,
465 Rotterdam, the Netherlands, [262]Julius Centre for Health Sciences and Primary Care, University
466 Medical Centre Utrecht, 3584CG, the Netherlands, [263]Second Department of Cardiology,
467 Medical School, National and Kapodistrian University of Athens, Attikon University
468 Hospital, Athens, Greece, [264]Center for Vision Research, Department of Ophthalmology and
469 The Westmead Institute, University of Sydney, Hawkesbury Rd, Sydney, New South Wales,
470 2145, Australia, [265]Menzies Institute for Medical Research, School of Medicine, University of
471 Tasmania, Liverpool St, Hobart, Tasmania, 7000, Australia, [266]Centre for Eye Research
472 Australia, University of Melbourne, Melbourne, Victoria, 3002, Australia, [267]Department of
473 Clinical Physiology, Tampere University Hospital, Tampere 33521, Finland, [268]Department of

Clinical Physiology, Finnish Cardiovascular Research Center - Tampere, Faculty of Medicine and Health Technology, Tampere University, Tampere 33014, Finland, [269]Centre Nutrition, santé et société (NUTRISS), Institute of Nutrition and Functional Foods (INAF), Québec, Canada, [270]Pennington Biomedical Research Center, Baton Rouge, LA 70808, USA, [271]Medical Department III – Endocrinology, Nephrology, Rheumatology, University of Leipzig Medical Center, Liebigstr. 18, 04103 Leipzig, Germany, [272]Discipline of Internal Medicine, Medical School, The University of Western Australia, Perth, WA, Australia, [273]Institute of Epidemiology, Kiel University, Kiel, Germany, [274]Institute of Clinical Molecular Biology, Kiel University, Kiel, Germany, [275]Sahlgrenska University Hospital, Department of Drug Treatment, Gothenburg, Sweden, [276]Geriatric Medicine, Institute of Medicine, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden, [277]Department of Internal Medicine, EwhaWomans University School of Medicine, Seoul, Korea, [278]Division of Cancer Control and Population Sciences, UPMC Hillman Cancer Center, University of Pittsburgh, Pittsburgh, PA 15232, USA, [279]Department of Epidemiology, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA 15232, USA, [280]Healthy Longevity Translational Research Programme, Yong Loo Lin School of Medicine, National University of Singapore, Singapore 117545, Singapore, [281]Singapore Institute for Clinical Sciences, Agency for Science Technology and Research (A*STAR), Singapore 117609, Singapore, [282]Department of Endocrinology and Metabolism, Kyung Hee University School of Medicine, Seoul 02447, Korea, [283]Institute for Community Medicine, University Medicine Greifswald, Germany, [284]Laboratory of Epidemiology and Population Science National Institute on Aging Intramural Research Program, NIH 251 Bayview Blvd, NIH Biomedical Research Center, Baltimore, MD 21224, USA, [285]Algebra University College, Ilica 242, Zagreb, Croatia, [286]Paavo Nurmi Centre, Sports and Exercise Medicine Unit, Department of Physical Activity and Health, University of Turku, Turku, Finland, [287]Interdisciplinary Center Psychopathology

and Emotion Regulation (ICPE), University of Groningen, University Medical Center Groningen, Groningen, 9700 RB, the Netherlands, [288]Institute of Molecular Genetics, National Research Council of Italy, Pavia, Italy, [289]Gottfried Schatz Research Center for Cell Signaling, Metabolism and Aging, Medical University of Graz, Graz, Austria, [290]Local Health Unit Toscana Centro, Firenze, Italy, [291]Institute of Nutritional Science, University of Potsdam, Nuthetal, Germany, [292]Department of Medicine, Surgery and Health Sciences, University of Trieste, Strada di Fiume 447, 34149, Trieste, Italy, [293]Dept of Nephrology, Diabetology, Rheumatology; Traunstein Hospital, Traunstein, Germany, [294]KfH Kidney Center Traunstein, Traunstein, Germany, [295]Center for Translational and Systems Neuroimmunology, Department of Neurology, Columbia University Medical Center, New York, NY, USA, [296]Program in Medical and Population Genetics, Broad Institute, Cambridge, MA, USA, [297]Medical School, National and Kapodistrian University Athens, 75 M. Assias Street, 115 27 Athens, Greece, [298]Dromokaiteio Psychiatric Hospital, 124 61 Athens, Greece, [299]Clinical Pharmacology, William Harvey Research Institute, Queen Mary University of London, London, EC1M 6BQ,UK, [300]Department of Ophthalmology, Medical Faculty Mannheim, Heidelberg University, Kutzerufer 1, Mannheim, 68167, Germany, [301]Institute of Molecular and Clinical Ophthalmology Basel, Switzerland, [302]Privatpraxis Prof Jonas und Dr Panda-Jonas, Heidelberg, Germany, [303]Department of Human Genetics, David Geffen School of Medicine at UCLA, University of California, Los Angeles, CA, USA, [304]Unidad de Biología Molecular y Medicina Genómica, Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán, Mexico 14080, Mexico, [305]Instituto de Investigaciones Biomédicas, UNAM, Mexico, [306]Departamento de Endocrinología y Metabolismo, Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán, Mexico 14080, Mexico, [307]Department of Nutrition, Gillings School of Global Public Health, University of North Carolina, Chapel Hill, North Carolina, 27599 USA, [308]Carolina Population Center, University of North Carolina, Chapel Hill, North

Carolina, 27516 USA, [309]USC–Office of Population Studies Foundation, University of San Carlos, Cebu City, 6000, Philippines, [310]Department of Anthropology, Sociology, and History, University of San Carlos, Cebu City, 6000 Philippines, [311]Department of Medicine, Faculty of Medicine, University of Kelaniya, Ragama, 11010, Sri Lanka, [312]Department of Public Health, Faculty of Medicine, University of Kelaniya, Ragama, 11010, Sri Lanka, [313]Children's Hospital Oakland Research Institute, Oakland, CA 94609 USA, [314]Institute of Biomedical Sciences, Academia Sinica, Taiwan, [315]Systems Genomics Laboratory, School of Biotechnology, Jawaharlal Nehru University, New Delhi - 110067, India, [316]Department of Medicine, University of Mississippi Medical Center, Jackson, MS, 39216, USA, [317]Department of Physiology and Biophysics, University of Mississippi Medical Center, Jackson, MS, 39216, USA, [318]Department of Medical Sciences, Uppsala University, Sweden, [319]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, Singapore, [320]Department of Medicine, University of North Carolina, Chapel Hill, NC, USA, [321]Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina, Chapel Hill, North Carolina, USA, [322]Injury Prevention Research Center, University of North Carolina, Chapel Hill, North Carolina, USA, [323]Division of Physical Therapy, University of North Carolina, Chapel Hill, North Carolina, USA, [324]Department of Psychiatry, Amsterdam UMC, Vrije Universiteit Amsterdam, the Netherlands, [325]Amsterdam Public Health research institute, VU medical center Amsterdam, the Netherlands, [326]Department of Biochemistry, College of Medicine, Ewha Womans University, Seoul 07804, Korea, [327]Faculty of Health and Medicine, University of Newcastle, Australia, [328]Washington University School of Medicine, Division of Biostatistics, MO, USA, [329]University of Kentucky, College of Public Health, KY, USA, [330]Institute of Cellular Medicine (Diabetes), The Medical School, Newcastle University,

549 Framlington Place, Newcastle upon Tyne, NE2 4HH, UK, [331]University of Helsinki and
550 Department of Medicine, Helsinki University Hospital, P.O.Box 340, Haartmaninkatu 4,
551 Helsinki, FI-00029, Finland, [332]Minerva Foundation Institute for Medical Research,
552 Biomedicum 2U, Tukholmankatu 8, Helsinki, FI-00290, Finland, [333]JSS Academy of Higher
553 Education and Research, Mysuru, India, [334]Programs in Metabolism and Medical and
554 Population Genetics, Broad Institute of MIT and Harvard, Cambridge, MA, USA, [335]Diabetes
555 Unit and Center for Genomic Medicine, Massachusetts General Hospital, Boston, MA,
556 USA10, [336]Harvard Medical School, Boston, Massachusetts, USA, [337]Unidad de Biología
557 Molecular y Medicina Genómica, Instituto de Investigaciones Bimédicas UNAM/ Instituto
558 Nacional de Ciencias Médicas y Nutrición Salvador Zubirán, Mexico City, Mexico,
559 [338]Dirección de Nutrición and Unidad de Estudios de Enfermedades Metabólicas, Instituto
560 Nacional de Ciencias Médicas y Nutrición Salvador Zubirán, Mexico City, Mexico,
561 [339]InstitutoNacional de Salud Publica y Centro de Estudios en Diabetes, Mexico, [340]Instituto
562 Nacional de Medicina Genómica, Mexico, [341]Human Genetics Center, School of Public
563 Health, University of Texas Health Science Center at Houston, Houston TX 77030, USA,
564 [342]Yong Loo Lin School of Medicine, National University of Singapore and National
565 University Health System, 119228, Singapore, [343]MRC Unit for Lifelong Health and Ageing at
566 UCL, 1-19 Torrington Place, London, WC1E 7HB, United Kingdom, [344]Kurume University
567 School of Medicine, Kurume, 830-0011, Japan, [345]Genetics, Merck Sharp & Dohme Corp.,
568 Kenilworth, NJ, 07033, USA, [346]Oxford Centre for Diabetes, Endocrinology & Metabolism,
569 University of Oxford, UK, [347]Population Health and Genomics, University of Dundee,
570 Ninwells Hospital and Medical School, Dundee, DD1 9SY, UK, [348]Intramural Research
571 Program, National Institute on Aging, 3001 S. Hanover St., Baltimore, MD 21225, [349]Beijing
572 Institute of Ophthalmology, Beijing Tongren Eye Center, Beijing Tongren Hospital, Capital
573 Medical University, Beijing Ophthalmology and Visual Sciences Key Laboratory, 100730

574 Beijing, China, [350]The Eye Hospital, School of Ophthalmology & Optometry, Wenzhou
575 Medical University, Wenzhou, Zhejiang 325027, China, [351]Synlab Academy, SYNLAB
576 Holding Deutschland GmbH, Mannheim and Augsburg, Germany, [352]Clinical Institute of
577 Medical and Chemical Laboratory Diagnostics, Medical University of Graz, Austria,
578 [353]Faculty of Medicine, University of Iceland, 101 Reykjavik, Iceland, [354]Leiden University
579 Medical Center, Department of Cell and Chemical Biology, Leiden, 2333ZA, The
580 Netherlands, [355]Leiden University Medical Center, Department of Biomedical Data Sciences,
581 Section Molecular Epidemiology, Leiden, 2333ZA, The Netherlands, [356]Amsterdam UMC,
582 Department of Epidemiology and Data Science, Amsterdam Public Health Research Institute,
583 Amsterdam, 1081HV, the Netherlands., [357]Amsterdam UMC, Department of General Practice
584 and Elderly Care, Amsterdam Public Health Research Institute, Amsterdam, 1081HV, The
585 Netherlands, [358]Department of Genetics, University of Pennsylvania, Philadelphia, PA, 19104,
586 USA, [359]Department of Surgery, University of Pennsylvania, Philadelphia, PA, 19104, USA,
587 [360]Corporal Michael Crescenz VA Medical Center, Philadelphia, Pennsylvania, PA, 19104,
588 USA, [361]Institute of Social and Economic Research, University of Essex, Wivenhoe Park, CO4
589 3SQ, United Kingdom, [362]Department of Cardiology, University of Groningen, University
590 Medical Center Groningen, 9700RB Groningen, The Netherlands, [363]Department of
591 Environmental Medicine and Public Health, Icahn School of Medicine at Mount Sinai, New
592 York, NY, USA, [364]Unidad de Investigacion Medica en Bioquimica, Hospital de
593 Especialidades, Centro Medico Nacional Siglo XXI, Instituto Mexicano del Seguro Social,
594 Mexico City, Mexico., [365]Department of Epidemiology, University of Washington, Seattle,
595 WA, USA, [366]Department of Health Services, University of Washington, Seattle, WA, USA,
596 [367]Institute of Regional Health Research, University of Southern Denmark, Odense, Denmark,
597 [368]16Center for Research on Genomics and Global Health, National Human Genome Research
598 Institute, National Institutes of Health, 12 South Drive, Room 4047, Bethesda, MD, 20892,

599 USA,, [369]Danish Aging Research Center, University of Southern Denmark; Odense C,
600 Denmark, [370]Public Health, Faculty of Medicine, University of Helsinki, Finland, [371]Broad
601 Institute of MIT and Harvard, Cambridge, MA, [372]Center for Applied Genomics, Children's
602 Hospital of Philadelphia, Philadelphia, PA, 19104 USA, [373]Department of Pediatrics, The
603 University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, 19104 USA,
604 [374]Division of Human Genetics, Children's Hospital of Philadelphia, Philadelphia, PA, 19104
605 USA, [375]Department of Genetics, University of Pennsylvania, Philadelphia, PA, 19104 USA,
606 [376]Radboud university medical center, Radboud Institute for Health Sciences, Nijmegen, The
607 Netherlands, [377]School of Medicine, Southern University of Science and Technology,
608 Shenzhen, China, [378]Institute for Cardiogenetics, University of Lübeck, DZHK (German
609 Research Centre for Cardiovascular Research), partner site Hamburg/Lübeck/Kiel, and
610 University Heart Center Lübeck, Lübeck, Germany, [379]Netherlands Heart Institute, Utrecht,
611 the Netherlands, [380]Division of Cardiology, Department of Medicine, Massachusetts General
612 Hospital, Boston, Massachusetts, USA, [381]Program of Medical and Population Genetics, Broad
613 Institute, Cambridge, Massachusetts, USA, [382]Center for Genomic Medicine, Massachusetts
614 General Hospital, Boston, Massachusetts, USA, [383]Department of Medicine, Harvard Medical
615 School, Boston, Massachusetts, USA, [384]Northern Finland Birth Cohorts, Infrastructure for
616 population studies, Faculty of Medicine, University of Oulu, Oulu, Finland, [385]Center for Life
617 Course Health Research, Faculty of Medicine, University of Oulu, Oulu, Finland, [386]Biocenter
618 of Oulu, University of Oulu, Oulu, Finland, [387]Institute for Genetic and Biomedical Research,
619 Italian National Council of Research (IRGB CNR), Cagliari, Italy, [388]University of Sassari,
620 Sassari, Italy, [389]Department of Clinical Epidemiology, Leiden University Medical Center,
621 Leiden, the Netherlands, [390]Department of Public Health and Primary Care, Leiden University
622 Medical Center, Leiden, the Netherlands, [391]Department of Internal Medicine, Division of
623 Endocrinology, Leiden University Medical Center, Leiden, the Netherlands, [392]Einthoven

624 Laboratory for Experimental Vascular Medicine, Leiden University Medical Center, Leiden,
625 the Netherlands, [393]Department of Human Genetics, Leiden University Medical Center,
626 Leiden, the Netherlands, [394]Population Health Research Institute, St George's, University of
627 London, London SW17 0RE, UK, [395]National Heart and Lung Institute, Imperial College
628 London, London, W2 1PG, UK, [396]School of Public Health, Imperial College London,
629 London, W12 7RH, UK, [397]Taichung Veterans General Hospital, Taichung, Taiwan; No. 1650,
630 Sec. 4, Taiwan Boulevard, Xitun District Taichung City 40705, Taiwan, [398]Division of
631 Endocrinology and Metabolism, Department of Medicine, Taipei Veterans General Hospital,
632 Taipei, Taiwan; No. 201, Sec. 2, Shipai Road, Beitou District, Taipei City, 112201, Taiwan,
633 [399]OCDEM, University of Oxford, Churchill Hospital, Oxford OX3 7LE, UK, [400]NIHR Oxford
634 Biomedical Research Centre, Churchill Hospital, Oxford, UK, [401]Ocular Epidemiology,
635 Singapore Eye Research Institute, Singapore National Eye Centre, 168751, Singapore,
636 [402]Ophthalmology & Visual Sciences Academic Clinical Program (Eye ACP), Duke-NUS
637 Medical School, 169857, Singapore, [403]Data Science, Singapore Eye Research Institute,
638 Singapore National Eye Centre, 168751, Singapore, [404]DZHK (German Centre for
639 Cardiovascular Research), Munich Heart Alliance partner site, Munich, Germany, [405]German
640 Center for Diabetes Research (DZD), Neuherberg, Germany, [406]University of Exeter Medical
641 School, University of Exeter, Exeter, EX2 5DW, UK, [407]Department of Medical Epidemiology
642 and Biostatistics, Karolinska Institutet, Stockholm, Sweden, [408]Amsterdam Public Health
643 research institute, Amsterdam UMC, the Netherlands, [409]Framingham Heart Study, National
644 Heart, Lung, and Blood Institute, US National Institutes of Health, Bethesda, MD, USA.,
645 [410]Department of Genetics, School of Medicine, Mashhad University of Medical Sciences,
646 Mashhad, Iran, [411]Department of Genetics, Shanghai-MOST Key Laboratory of Health and
647 Disease Genomics, Chinese National Human Genome Center at Shanghai, Shanghai, 201203
648 China, [412]Technical University of Munich (TUM) and Klinikum Rechts der Isar, TUM School

649 of Medicine, Munich, Germany, [413]Department of Public Health, University of Helsinki,
650 Helsinki, Finland, [414]Diabetes Research Group, King Abdulaziz University, Jeddah, Saudi
651 Arabia, [415]Institute of Clinical Medicine, Internal Medicine, University of Eastern Finland and
652 Kuopio University Hospital, Finland, [416]Stanford Cardiovascular Institute, Stanford
653 University, Stanford, CA 94305, USA, [417]Stanford Diabetes Research Center, Stanford
654 University, Stanford, CA 94305, USA, [418]Department of Medical Sciences, Molecular
655 Epidemiology and Science for Life Laboratory, Uppsala University, Uppsala, Sweden.,
656 [419]Regeneron Pharmaceuticals, Tarrytown, NY, USA, [420]Lee Kong Chian School of Medicine,
657 Nanyang Technological University, Singapore 308232, Singapore, [421]Imperial College
658 Healthcare NHS Trust, Imperial College London, London W12 0HS, UK, [422]MRC-PHE
659 Centre for Environment and Health, Imperial College London, London W2 1PG, UK,
660 [423]National Heart and Lung Institute, Imperial College London, London W12 0NN, UK,
661 [424]School of Electrical & Information Engineering, University of the Witwatersrand, South
662 Africa, [425]Department of Epidemiology, University of North Carolina, Chapel Hill, NC, USA,
663 [426]Institute for Minority Health Research, University of Illinois College of Medicine, Chicago,
664 Illinois, USA, [427]Department of Biostatistics, Harvard T.H. Chan School of Public Health, 677
665 Huntington Avenue, Boston, MA, 02115, USA, [428]QIMR Berghofer Medical Research
666 Institute, 300 Herston Road, Brisbane, Queensland 4006, Australia, [429]Center for Non-
667 Communicable Diseases, Karachi, Sindh, Pakistan & Faisalabad Institute of Cardiology,
668 Faislabad, Pakistan, [430]Department of Medicine, Columbia University Irving Medical Center,
669 New York, NY, USA, [431]Department of Cardiology, Columbia University Irving Medical
670 Center, New York, NY, USA, [432]Big Data Instutute, University of Oxford, Oxford OX3 7LF,
671 UK, [433]National Institute for Health Research Oxford Biomedical Research Centre, Oxford
672 University Hospitals, Oxford, UK, [434]Aberdeen Centre for Health Data Science,1:042
673 Polwarth Building School of Medicine, Medical Science and Nutrition University of

Aberdeen, Foresterhill, Aberdeen, AB25 2ZD, UK, [435]Division of Population Health and Genomics, Ninewells Hospital and Medical School, University of Dundee, Dundee DD1 9SY, United Kingdom, [436]Biomedical and Translational Informatics, Geisinger Health, Danville, PA 17822, USA, [437]Harvard Medical School, Boston, MA 02115, USA, [438]School of Basic and Medical Biosciences, Faculty of Life Sciences and Medicine, King's College London, London, UK, [439]Department of Biomedical Informatics, Vanderbilt University Medical Center, Nashville, TN, USA, [440]Departments of Medicine (Medical Genetics) and Genome Sciences, University of Washington, USA, [441]Center for Autoimmune Genomics and Etiology, Cincinnati Children's Hospital Medical Center (CCHMC), Cincinnati, OH, USA., [442]Division of Endocrinology, Metabolism, and Molecular Medicine, Department of Medicine, Northwestern University, Feinberg School of Medicine, Chicago, IL 60611, USA, [443]Department of Anthropology, Northwestern University, Evanston, IL 60208, USA, [444]Center for Genetic Medicine, Northwestern University, Feinberg School of Medicine, Chicago, IL 60611, USA, [445]HUNT Research Centre, Department of Public Health and Nursing, NTNU, Norwegian University of Science and Technology, Levanger, 7600 Norway, [446]Department of Medicine, Levanger Hospital, Nord-Trøndelag Hospital Trust, Levanger, 7600 Norway, [447]Department of Endocrinology, St. Olavs Hospital, Trondheim University Hospital, Trondheim, Norway, [448]RIKEN Center for Integrative Medical Sciences, Yokohama, Japan, [449]Laboratory of Complex Trait Genomics, Department of Computational Biology and Medical Sciences, Graduate School of Frontier Sciences, The University of Tokyo, Tokyo, Japan, [450]Laboratory of Statistical Immunology, WPI Immunology Frontier Research Center, Osaka University, Osaka, Japan, [451]Integrated Frontier Research for Medical Science Division, Institute for Open and Transdisciplinary Research Initiatives, Osaka University, Osaka, Japan, [452]Division of Molecular Pathology, Institute of Medical Science, The University of Tokyo, Tokyo, Japan, [453]Division of Genome Research, Center for Genome Science, National

699 Institute of Health, Chungcheongbuk-do, South Korea, [454]Faculty of Medicine, University of
700 Iceland, Sæmundargötu 2, Reykjavik, 102, Iceland, [455]VA Boston Healthcare System, Boston,
701 MA, USA, [456]VA Informatics and Computing Infrastructure, VA Salt Lake City Health Care
702 System, Salt Lake City, UT, USA, [457]University of Massachusetts, Boston, MA, USA,
703 [458]Department of Medicine, University of Pennsylvania Perelman School of Medicine,
704 Philadelphia, PA, USA, [459]Cardiovascular Institute, Stanford University School of Medicine,
705 Stanford, California, USA, [460]Corporal Michael J. Crescenz VA Medical Center, Philadelphia,
706 PA, USA, [461]Department of Medicine, Brigham Women's Hospital, Boston, MA, USA,
707 [462]Division of Cardiology, Emory University School of Medicine, Atlanta, GA, USA,
708 [463]Endocrinology, Boston Childrens Hospital, Boston 02115 MA, USA, [464]Departments of
709 Pediatrics and Genetics, Harvard Medical School, Boston, MA, USA, [465]Center for Genomic
710 Medicine, Department of Medicine, Massachusetts General Hospital, Boston, Massachusetts,
711 USA, [466]Cardiology Division, Massachusetts General Hospital, Harvard Medical School,
712 Boston, MA, USA, [467]Department of Medicine, Massachusetts General Hospital, Harvard
713 Medical School, Boston, MA, USA, [468]Cardiovascular Research Center and Center for
714 Genomic Medicine, Massachusetts General Hospital, Boston, MA, USA, [469]Centre for
715 Genetics and Genomics Versus Arthritis, Centre for Musculoskeletal Research, Division of
716 Musculoskeletal and Dermatological Sciences, The University of Manchester, Manchester,
717 UK, [470]Princess Al-Jawhara Al-Brahim Centre of Excellence in Research of Hereditary
718 Disorders (PACER-HD), King Abdulaziz University, Jeddah, Saudi Arabia, [471]Department of
719 Biostatistics, Boston University School of Public Health, Boston, MA, USA, [472]Department of
720 Human Genetics, University of Michigan, Ann Arbor, MI, 48109, USA, [473]Department of
721 Statistics, The Pennsylvania State University, University Park, PA, USA, [474]Huck Institutes of
722 the Life Sciences, The Pennsylvania State University, University Park, PA, USA, [475]VA Palo

723 Alto Health Care System, Palo Alto, CA, USA, [476]Department of Statistics, Stanford
724 University, Stanford, CA, USA

725 Corresponding Authors:

726 Xiang Zhu, PhD

727 Department of Statistics

728 Huck Institutes of the Life Sciences

729 The Pennsylvania State University

730 University Park, PA 16802

731 xiangzhu@psu.edu

732

733 Christopher D Brown, PhD

734 Department of Genetics

735 Perelman School of Medicine

736 University of Pennsylvania

737 Philadelphia PA 19104

738 chrbro@upenn.edu

739

740 **Abstract**

741 A major challenge of genome-wide association studies (GWAS) is to translate phenotypic
742 associations into biological insights. Here, we integrate a large GWAS on blood lipids
743 involving 1.6 million individuals from five ancestries with a wide array of functional
744 genomic datasets to discover regulatory mechanisms underlying lipid associations. We first
745 prioritize lipid-associated genes with expression quantitative trait locus (eQTL)
746 colocalizations, and then add chromatin interaction data to narrow the search for functional
747 genes. Polygenic enrichment analysis across 697 annotations from a host of tissues and cell
748 types confirms the central role of the liver in lipid levels, and highlights the selective
749 enrichment of adipose-specific chromatin marks in high-density lipoprotein cholesterol and
750 triglycerides. Overlapping transcription factor (TF) binding sites with lipid-associated loci
751 identifies TFs relevant in lipid biology. In addition, we present an integrative framework to
752 prioritize causal variants at GWAS loci, producing a comprehensive list of candidate causal
753 genes and variants with multiple layers of functional evidence. Two prioritized genes,
754 *CREBRF* and *RRBP1*, show convergent evidence across functional datasets supporting their
755 roles in lipid biology.
756
757

Introduction

Most GWAS findings have not directly led to mechanistic interpretations, largely because 90% of GWAS associations map to non-coding sequences [1,2]. Mechanistic interpretations in GWAS have proven challenging because the strongest signals identified in GWAS typically contain many variants in strong linkage disequilibrium (LD) [3] and functional mechanisms including genes of action are often not clear from GWAS data alone [4,5].

Linking trait-associated variants to genome function has emerged as a promising model for mechanistic interpretation of non-coding findings in GWAS. This 'variant-to-function' model is premised on recent observations that non-coding variants often affect a trait of interest through the regulation of genes and processes in trait-relevant cell types or tissues [2,6]. Implementing this functional model in GWAS has become more feasible as large-scale functional genomic resources, such as epigenomic [7] and transcriptomic [8] catalogues, have been systematically generated across a wide range of human cell types and tissues. The integration of functional genomics with GWAS has identified regulatory mechanisms in variants associated with some flagship disorders such as obesity [9] and schizophrenia [10], yielding important functional insights into the genetic architecture of human complex traits.

The history of the human genetics of lipids mirrors the successes and challenges of GWAS. Increasing sample size and genetic diversity has significantly boosted the power of discovery: the first lipid GWAS in 2008 with 8,816 European-descent individuals identified 29 lipid-associated loci[11] ; the latest study of 1.6 million individuals across five ancestries [12] found 941. Despite the dramatic increase in the number of associations, our biological understanding of many of these genetic discoveries remains limited. The causal gene has

783 been confidently assigned at only a small fraction of these loci [2], and the regulatory
784 mechanism connecting variant to phenotype has been conclusively characterized only for a
785 handful of genes [5]. Furthermore, systematic mapping of lipid-associated variants to their
786 biological functions has been missing in the literature at the time of this study.
787
788 Here we conduct a genome-scale integrative analysis on the largest GWAS to-date of five
789 lipid phenotypes (LDL, or low density lipoprotein; HDL, or high density lipoprotein; TC, or
790 total cholesterol; nonHDL, or non-high density lipoprotein; and TG, or triglycerides)
791 involving 1.65 million individuals from five ancestries [12]. Combining the lipid GWAS with a
792 wide array of functional genomic resources in diverse human tissues and cell types, we
793 identify regulatory mechanisms of noncoding genetic variation in lipids with a full suite of
794 computational approaches. Further, we develop a generalizable framework to understand how
795 tissue-specific gene regulation can explain GWAS findings, and demonstrate its real-world
796 value on lipid-associated loci.
797

798 **Material and Methods**

799 *GWAS*
800
801 We performed GWAS for five blood lipid traits (LDL, HDL, TC, TG, and nonHDL) in 1.65
802 million individuals from five ancestry groups [12](African and African-admixed, East Asian,
803 European, Hispanic, South Asian) at 91 million variants imputed primarily from the
804 Haplotype Reference Consortium [13] or 1000 Genomes Phase 3 [14]. The individual GWAS and
805 meta-analyses were performed using the hg19 version of the human reference genome. We
806 used MR-MEGA [15] for meta-analysis across cohorts.
807

We defined 'sentinel variants' as lead variants representing independent trait-associated loci in the genome. These windows are the greater of 500kb or 0.25cM around the sentinel variant; genetic distances were defined using reference maps from HapMap 3 [16]. We performed a second round of conditional analysis, conditioning on the sentinel variants to identify and remove any significant windows that are shadow signals (or dependent on) of a neighboring locus to enforce independence of associated loci.

Colocalization with eQTLs

We performed statistical colocalization of lipid GWAS with eQTLs obtained from GTEx v8 across 49 tissues [8]. For each of the five lipid traits, we used the same sentinel variants defined in the previous section to represent approximately independent GWAS-associated windows (also removing shadow signals as described before).

For each such window, we ran eQTL colocalization with GTEx v8 single-tissue cis-eQTL summary statistics [8]. For each of 49 GTEx tissues, we first identified all genes within 1Mb of the sentinel SNP, and then restricted analysis to those genes with significant eQTLs (i.e., 'eGenes' as defined by GTEx) in that tissue (FDR < 0.05). We used the R package 'coloc' (run on R version 3.4.3, coloc version 3.2.1) [17] with default parameters to run colocalization between the GWAS signal and the eQTL signal for each of these cis-eGenes, using as input those SNPs in the defined window (greater than 500kb or 0.25cM on either side of the lead variant), i.e. all SNPs present in both datasets. eQTL summary statistics were in GRCh38, so we first lifted over the GWAS summary statistics (in hg19) to GRCh38 using liftOver [18]. As in previous studies [19], we used a colocalization posterior probability of (PP3+PP4) > 0.8 to identify loci with enough colocalization power, and PP4/PP3 > 0.9 to define those loci that

show significant colocalization, where PP4 represents posterior probability of a single shared signal, and PP3 represents posterior probability of two unique signals in the GWAS and eQTL datasets.

Overlap with promoter Capture-C data

We used four promoter-focused Capture-C (henceforth Capture-C) datasets from three human cell/tissue types to capture physical interactions between gene promoters and their regulatory elements. We used three biological replicates of HepG2 liver carcinoma cells [20], another HepG2 dataset described in Selvarajan et al [21], hepatocyte-like cells (HLC) produced by differentiating three biological replicates of iPSCs (which in turn were generated from peripheral blood mononuclear cells using a previously published protocol [22]), and an adipose dataset obtained from Pan et al [23] that was produced using primary human white adipocytes.

The detailed protocol to prepare HepG2 or HLC cells for the Capture-C experiment is described in Chesi et al[20]. Briefly, for each dataset, 10 million cells were used for promoter Capture-C library generation. Custom capture baits were designed using an Agilent SureSelect library design targeting both ends of DpnII restriction fragments encompassing promoters (including alternative promoters) of all human coding genes, noncoding RNA, antisense RNA, snRNA, miRNA, snoRNA, and lincRNA transcripts, totaling 36,691 RNA baited fragments. Each library was then sequenced on an Illumina HiSeq 4000 (HepG2) or Illumina NovoSeq (HLC), generating 1.6 billion read pairs per sample (50 base pair read length.) We used HiCUP v0.7.2 [24] to process the raw FastQ files into loop calls and CHiCAGO v1.6.0 [24,25] to define significant looping interactions; we defined a CHiCAGO score of 5 as significant, as specified in the default parameters.

Starting with Capture-C maps processed as described above, we re-annotated the baits to gene IDs from Gencode v19 [26] to ensure uniformity of gene annotations with the rest of our pipeline. For each bait, we identified any gene whose transcription start site (TSS) from any transcript in Gencode v19 was within 175 base pair distance from the bait (to account for differing bait designs for external datasets which may not directly overlap the canonical TSS). We filtered all datasets to only include interactions in which the interacting end was not another bait. Enrichment with colocalized genes was robust to our choice of distance between bait and gene (enrichment with eQTL colocalized genes ranging from 2.94-2.96 for bait distances from 0-350 base pairs).

To identify genetic variants associated with any of the five lipid traits that physically interact with locations in the genome, we used the R package 'Genomic Ranges' version 1.30.3 [27] to find overlap between credible sets for each trait's GWAS and the previously annotated promoter Capture-C data; we refer to these as Capture-C/GWAS interactions. Each credible set was defined as the set of variants with a 95% posterior probability of being the causal variant. For all individual variants within all GWAS-associated loci for the five lipid traits, we identified which variants overlapped any interacting end of the four previously annotated promoter Capture-C data.

Presence of gene-variant pairs in same topologically associated domains

To estimate the frequency of colocalized gene-sentinel pairs in the same topologically associated domain (TAD), we used publicly-available TADs from human liver [28]. We compared the number of colocalizations with the sentinel variant and colocalized gene in the

same TAD divided by all colocalizations in which the sentinel variant lies in a TAD. To test if this ratio was statistically significant, we generated random TAD boundaries using 'bedtools shuffle' 1000 times, and calculated the same ratio for these randomly-generated TAD boundaries.

Pathway enrichment

We used ClusterProfiler v3.6.0 [29] to look for pathways over-represented in each gene list: genes with eQTL colocalization and genes interacting with variants in GWAS credible sets. We used the enrichKEGG function to look for pathway enrichment in KEGG pathways (using the latest version of the KEGG database [30]). We first re-mapped gencode IDs to gene symbols using the Gencode v24 annotation and then used the biomaRt R package v2.34.2 [31] to convert gene symbols to Entrez IDs. We ran enrichKEGG to identify enriched pathways significant at a Benjamini-Hochberg threshold of 0.05.

Enrichment in known lipid-associated genes

We calculated enrichment odds ratio of genes identified in our analysis with three known sets of lipid-associated genes using the Fisher's exact test (R function 'fisher.test'). First, we identified a list of 33 Mendelian genes from ClinVar [32] with lipidemia-associated ICD10 codes (E78). Second, we used the set of genes identified from a transcriptome-wide association study (TWAS) on the same GWAS and GTEx v8 summary statistics using the S-PrediXcan software [33] default setup. Third, we used 35 genes with rare-coding variants associated with lipid levels [34].

908 *Stratified LD score regression*

909

910 We used LDSC version 1.0.1 [35] to estimate the enrichment of heritability using GWAS
911 summary statistics in different epigenetic and transcriptomic annotations, including gene
912 expression, chromatin marks and TF binding sites. The gene expression and chromatin mark
913 annotations and the corresponding LD scores were provided as
914 'Multitissuegeneexpr1000Gv3' and 'Multitissuechromatin1000Gv3' databases in LDSC
915 software. The TF binding site annotations were extracted from ChIP-seq data of 161 TFs
916 from ENCODE, and their LD scores were estimated from 1000 Genomes Phase 3 European
917 samples using 'ldsc.py --l2'. We first converted the summary statistics for each phenotype to
918 LDSC-formatted summary statistics using 'munge_sumstats.py'. Second, we ran 'ldsc.py'
919 using the baseline_v1.2 baseline model on each annotation to estimate enrichment of
920 heritability. For primary analyses, we used multi-population GWAS summary statistics and
921 LD scores estimated from 1000 Genomes Phase 3 European samples. For secondary analyses
922 on East Asian GWAS alone, we obtained EAS-specific LD scores for the same epigenomic
923 annotations [36].

924

925 *GREGOR analysis*

926

927 We used GREGOR [37] to estimate enrichment of sentinel variants for each lipid phenotype in
928 TF binding sites for 161 TFs from ENCODE compared to a null distribution of variants
929 matched for allele frequency. We ran GREGOR with default parameters, specifying 0.8 as
930 the R^2 threshold, window size of 1Mb, and 'EUR' as the population. Annotations with FDR-
931 adjusted P-value < 0.05 were considered significant.

932

Enrichment in single-cell expression data

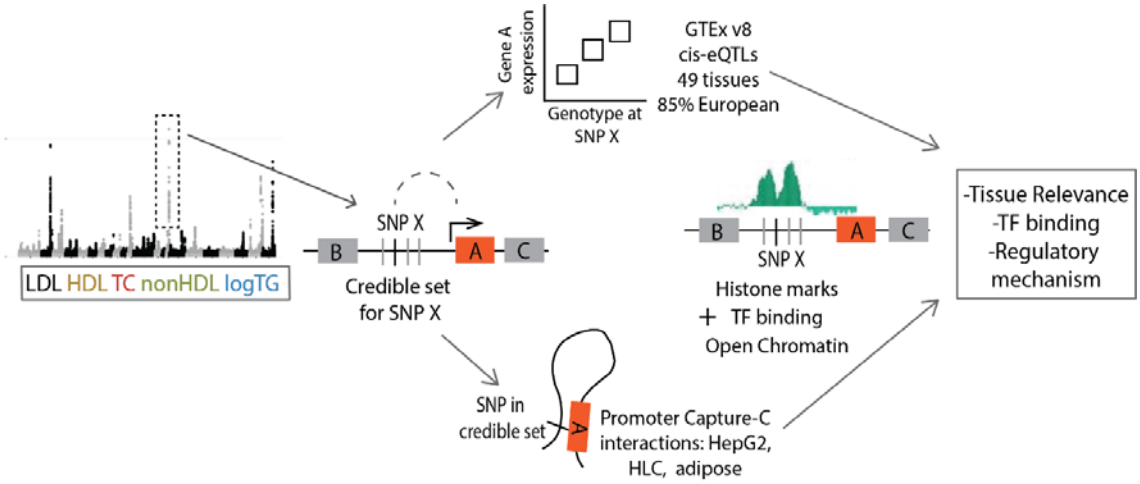
We overlapped our list of colocalized genes with publicly available single-cell RNA-sequencing data of 8,444 cells from liver [38] and 38,408 cells from adipose (Web resources) in humans. For both datasets, we downloaded normalized TPM data and existing tSNE cluster annotations for each cell. For each cluster, we defined median expression for each gene across all cells in that cluster. Then for each cluster, we calculated the enrichment P-value for our list of colocalized genes using the 'fgsea' R package v1.4.1, which looks for overrepresentation of our gene list in ranked genes for each cluster [39], implemented in R 3.4.3.

Results

We systematically integrated lipid GWAS results [12] with multiple layers of functional genomic data from diverse tissues and cell types to understand regulatory mechanisms at lipid-associated loci (Figure 1). Specifically, we overlaid GWAS loci with eQTL and chromatin-chromatin interactions to identify causal genes. We assessed polygenic enrichments of tissue-specific histone marks to prioritize relevant tissues and examined GWAS loci at transcription factor (TF) binding sites to detect lipid-relevant TFs. Finally, we combined all these layers to prioritize functional variants at GWAS loci, providing a holistic view of gene regulation at lipid loci in relevant tissue and cell types.

Figure 1: Schematic overview of the multi-layer functional genomic analysis. We first integrate GWAS summary statistics for five lipid phenotypes with eQTL and chromatin interaction data to identify potential genes mediating the GWAS association, and then

incorporate epigenomic annotations to identify regulatory mechanisms at these loci. For any lead variant 'X', A, B, and C represent nearby eGenes, and SNPs around SNP X represent variants in the credible set.



Colocalization with eQTLs identifies candidate lipid-relevant genes

First, we identified shared association signals between lipid levels and expression of nearby genes, since most GWAS signals are presumed to influence complex traits through impact on gene expression [40]. To do so, we tested for colocalization of each of the 1,750 significant lipid GWAS association signals across the five traits examined with significant cis-eQTL data across 49 human tissues from the GTEx consortium [8]. Here, we defined significant GWAS signals as 1,750 loci reaching genome-wide significance and corrected for shadow signals (Methods) in our multi-population meta-analysis for at least one of five lipid traits.

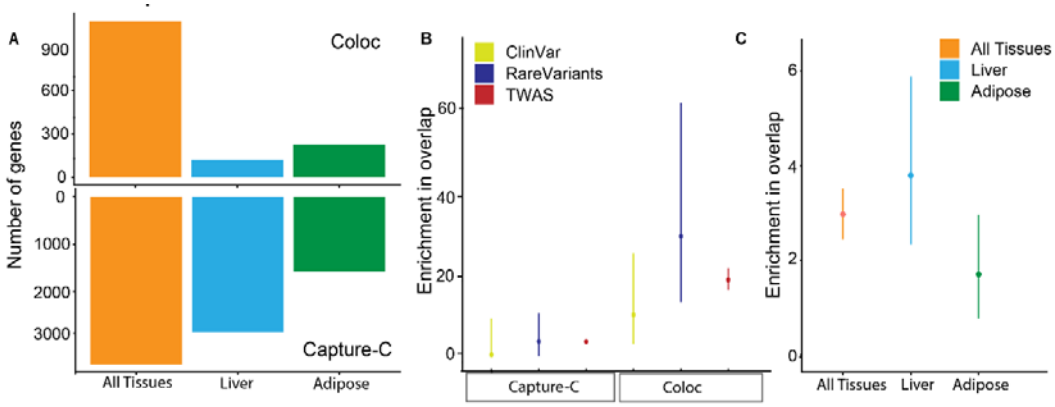
Second, we restricted our analysis to those loci likely mediated through regulatory mechanisms as opposed to coding variation. In particular, we excluded all loci with credible

sets containing at least one missense variant (369 of 1,750 loci, 21% of credible sets). Of the remaining 1,381 GWAS loci, 696 significantly colocalized with eQTLs (the ratio of posterior probability of a shared signal to the posterior probability of two signals being > 0.9 [19]; Methods) in at least one of 49 tissues for at least one lipid phenotype. This resulted in 1,076 colocalized eGenes ranging from 1 to 16 genes per locus (Figure 2A; Table S1). Since with eQTL data alone it is difficult to disentangle a single functional gene from multiple functional (and likely coregulated) genes at a locus [41] we performed all downstream analyses with all 1,076 colocalized genes, to further prioritize functional genes at loci with multiple eGenes.

To acquire additional functional insights into the 1,076 colocalized genes, we assessed their enrichments across existing biological and clinical gene sets. Colocalized genes showed enrichments in (a) 20 KEGG pathways [30] at FDR 5% (Table S2), including known lipid-related processes such as cholesterol metabolism, PPAR signaling, and bile secretion; (b) 33 Mendelian genes from ClinVar [32] associated with lipid-related ICD10 codes, (11 fold enrichment at P=2.08e-06, including *APOB, LPL,* and *APOE;* Figure 2B), suggesting the shared genetic basis of Mendelian and complex lipid phenotypes [42]; (c) 35 genes with rare-variant burden for lipid phenotypes in a recent multi-ancestry analysis [34] (30-fold enrichment, P = 1.77e-16, including *APOB, LPL, LIPG* and *ANGPTL4*), confirming shared mechanisms of rare and common variation underlying lipid traits [42,43]. Colocalized genes also showed enrichment with genes implicated in TWAS run on the same GWAS and eQTL summary statistics (20-fold enrichment, P<2.22e-308). These enrichment results demonstrate the biological relevance of candidate functional genes prioritized by our approach.

Figure 2: Overlap between eQTL colocalized genes and capture-C prioritized genes, and their enrichment in known lipid-associated genes. A. Numbers of genes identified by two

approaches: eQTL colocalization (upper half) and promoter capture-c interactions (lower half) B. Overlap between our list of prioritized genes (left: capture-C prioritized genes; right: eQTL colocalized genes) with three sets of genes previously associated with lipid biology (ClinVar lipidemia-associated genes, genes implicated in rare burden of lipids, and genes from a lipid TWAS). C. Enrichment in overlap between eQTL colocalized genes and capture-C prioritized genes against what is expected by chance, assuming both gene sets are independent. Enrichment estimates and confidence intervals shown in Panels B and C were obtained using Fisher's exact test.



Chromatin-chromatin interactions improve eQTL-based colocalization

Our eQTL-based colocalization analysis uses a linear sequence of DNA, and ignores physical interaction between non-adjacent DNA segments, another regulatory layer underlying complex human traits [44]. To add this layer to our analysis, we generated Capture-C data from HepG2 liver carcinoma cells (denoted as HepG2.1) and hepatocyte-like cells (HLC) derived from differentiating iPSCs (the latter is described in [22]), as well as publicly-available Capture-C datasets from HepG2 [21,43] (denoted as HepG2.2) and adipose tissue [23]. We defined a GWAS-relevant interaction as any Capture-C interaction between any gene and a variant in the 95% credible set for a GWAS locus[45]. Credible set sizes ranged from 1 to 417 variants at

the 1,750 examined loci, with a median size of 5 variants per credible set. In total, 1,079 GWAS loci had at least one variant in the credible set with a physical interaction with a gene promoter and 3,543 of 26,621 genes with promoter-interactions had promoters physically interacting with at least one GWAS credible set variant (Figure 2A; Table S3). Unlike eQTL-colocalized genes, these genes interacting with their credible sets showed limited enrichment in relevant KEGG pathways (Table S2) and lipid-related genes from ClinVar (Figure 2B), though we see 5-fold enrichment (compared to greater than 10-fold enrichment for eQTL-colocalized genes) in genes with rare-variant lipid associations (P =2.8e-05) and TWAS genes (P=2.5e-288).

Genes physically interacting with GWAS loci helped shortlist functional genes from eQTL colocalization despite their reduced enrichments in known gene sets. Of 1,079 credible sets with promoter interactions, 224 also colocalized with eQTLs for the same gene. At the gene level, 233 genes were implicated in both eQTL colocalization and Capture-C interactions (Figure 2C), representing an enrichment of 3-fold compared to random chance (P =3.11e-38). Among these loci with concordant eQTL colocalizations and Capture-C interactions, only 39% of them mapped to a single gene using eQTL data alone, whereas adding Capture-C information increased this fraction to 80%. These results showcase the potential value of combining eQTLs with physical chromatin interactions to prioritize functional genes at GWAS loci.

Since eQTLs are likely to reside in the same topologically associated domain (TADs) as the genes they regulate [46], we examined TAD structure from independent datasets at lipid GWAS loci with eQTL colocalizations. Of eQTL-GWAS colocalizations in which the sentinel variant resided within a liver TAD [28], the colocalized gene resided in the same liver TAD

84.8% of the time (P < 0.001 with 1000 permutations; Methods). When we restricted colocalizations to those supported by Capture-C data in any cell type, 91.2% fall in the same TAD. These results add to the existing evidence for TAD boundaries being regulatory insulators in the cell [47] and confirm our integration of chromatin interactions with eQTL colocalizations as an effective strategy to hone in on functional genes.

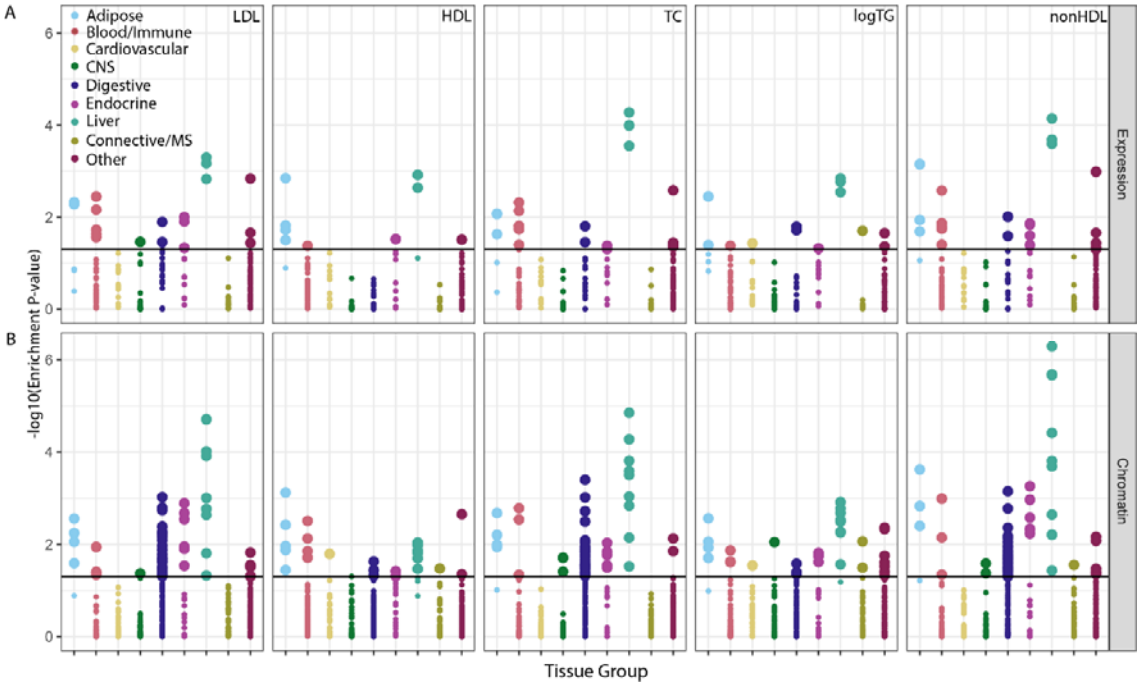
Tissue-specific enrichment of GWAS signals differentiates lipid traits

Regulatory variants often affect complex traits in a tissue-specific manner [6], as shown in our eQTL colocalization analysis. Specifically, by computing the ratio of the number of colocalizations in a tissue to eQTL sample size in that tissue, we found that the liver was universally enriched for colocalized eGenes with respect to sample size across all lipid traits whereas adipose was selectively enriched in HDL and TG only (Figure S1). Motivated by these findings, we leveraged systematic approaches and additional data to identify relevant tissues and cell types for each lipid trait.

We implemented stratified LD score regression (S-LDSC), a polygenic approach not restricted to genome-wide significant variants, on tissue-specific transcriptomic and epigenomic annotations across 204 datasets from more than 170 tissues and cell types, to identify relevant tissues for each lipid trait (Methods). Consistent with previous studies and our eQTL-based analysis, liver-related tissues (Table S4) showed strong enrichments across all lipid traits (S-LDSC enrichment p-values ranging from .001 in TG to .0001 in TC), for both expression (Figure 3A) and chromatin annotations (Figure 3B). This result was confirmed by analysis using two other approaches: DEPICT [48] (Figure S2) and RSS-NET [49] (Table S5). To assess the robustness of our S-LDSC results based on multi-population

GWAS, we applied S-LDSC to population-specific GWAS in European and East Asian ancestry participants together with population-specific LD scores (Methods) and obtained similar results (Table S6).

Figure 3: Tissue relevance based on lipid GWAS and functional annotations. Partitioning heritability of GWAS summary statistics for five lipid traits on gene expression (A) and chromatin mark (B) annotations across tissues. Each plotted point represents a tested dataset for enrichment of heritability, with larger dots representing datasets with P-value < 0.05; multiple annotation datasets are tested for the same tissue group. Each color represents a tissue group, and the y-axis represents -log10 P-value of enrichment of heritability.



The S-LDSC results also highlighted tissues selectively enriched in certain lipid traits as shown in the eQTL-based analysis. The most enriched category for HDL using chromatin annotation is 'Adipose H3K4me3' (P-value 7.6e-04); for TG, enrichment in liver-related

tissues (P-value 1.2e-03) is similar to enrichment in adipose (P-value 2.7e-03). For LDL, TC, and non-HDL, enrichment P-values for the liver were much more significant than for all other tissues including adipose (Figure 3B). We observed the same pattern in S-LDSC results based on gene expression (Figure 3A). This finding is consistent with the known influence of adipose on plasma HDL levels [50], and the role of adipose as TG deposits [51]. These results were corroborated by eQTL colocalizations stratified by phenotype (Figure S1) and DEPICT analysis on gene expression [48] (Figure S2). Together, these results confirm the liver as the tissue of action for all five lipid traits, and highlight the additional role of adipose in HDL and TG only.

Given the importance of the liver and adipose in modulating lipid levels, we further identified the relevant cell types within these tissues. Using existing single-cell data from adipose and liver, we performed gene-set enrichment analysis [52] to identify cell-type clusters enriched for genes colocalized with any lipid trait. Out of 11 identified cell types in 20 clusters in the liver, only hepatocytes were enriched at FDR-adjusted $P < 0.05$ (Figure S3), consistent with previous results[21]. In adipose, only adipocyte clusters and macrophage-monocyte clusters showed suggestive enrichment (nominal $P < 0.05$) in colocalized genes (Figure S4). Of note, the enrichment in adipocytes was significant when we restricted this analysis to genes that were colocalized only with HDL and TG (FDR-corrected $P < 0.05$), consistent with the selective enrichments of adipose in HDL and TG (but not the other lipid traits) from our S-LDSC analysis. Evaluations at cellular resolution are required to understand the cell-type specific mechanisms underlying lipid GWAS loci, but our results could form a useful basis for future studies.

Overlapping GWAS signals with binding sites highlights lipid-relevant TFs

TFs have been implicated as a key mediator of linking genetic variation to complex traits [53]. To understand lipid GWAS in the context of TF activity, we assessed enrichment of genome-wide significant variants at TF binding sites using GREGOR [37] and performed polygenic enrichment analysis of TF binding sites using S-LDSC.

Using ChIP-Seq data from 161 TFs across 91 cell types from the ENCODE project [7], 70.7% of lipid credible sets overlapped with at least one TF binding site. Using GREGOR [37], we identified 137 TFs whose binding sites were significantly enriched in GWAS lead SNPs for at least one lipid phenotype (enrichment > 2; FDR adjusted P-value < 0.05; Figure S5; Table S7). Among these 137 enriched TFs, 69 of them (50%) showed significant enrichments across all five lipid phenotypes, suggesting a potential core regulatory circuit shared by all lipid traits (Figure S5). The TF with the strongest enrichment in all phenotypes was ESRRA (estrogen-related receptor alpha), a nuclear receptor active in metabolic tissues [54]; ESRRA has been implicated in adipogenesis and lipid metabolism, and ESRRA-null mice display an increase in fat mass and obesity [54].

The GREGOR analysis also highlighted 68 TFs significantly enriched in specific subsets of (but not all five) lipid phenotypes (Figure S8). For example, we found 4 TFs (FOXM1, PBX3, ZKSCAN1, ZEB1) enriched in HDL and TG only, 4 TFs (EZH2, NFE2, NFATC1, KDM5A) enriched in HDL only and 11 TFs (FOSL1, IRF3, JUN, MEF2C, NANOG, PRDM1, RUNX3, SIRT6, SMC3, STAT3, ZNF217) enriched in TG only. Of these TFs, the central role of ZEB1 in adiposity [55] and fat cell differentiation has been demonstrated [56]. Taken together, these TF-centric findings corroborate the selective enrichments of adipose in

HDL and TG (but not the other lipid traits) identified in our previous tissue prioritization analyses.

Similar to tissue prioritization, we also performed polygenic enrichment analysis of TF binding sites using S-LDSC (Table S8), which differed from GREGOR analysis by looking at not only the genome-wide significant associations but also the polygenic signal irrespective of GWAS P-values. On the same 161 ENCODE TFs, this polygenic analysis identified 25 TFs whose binding sites were significantly enriched in heritability (nominal $P < 0.05$) for at least one lipid phenotype (Figure S6); reassuringly, 24 of 25 TFs were also significant in GREGOR analysis. Among these enriched TFs, eight (34%) were significantly enriched in all five lipid traits *(*CEBPB, CEBPD, FOXA2, HDAC2, HNF4G, NFYA, RXRA, SP1; $P < 0.05$). Of those TFs significant in both analyses, RXRA (retinoid X receptor alpha) is also encoded by a colocalized gene (*RXRA*) near a GWAS hit (chr9:137,268,682). RXRA is a ligand-activated transcription factor that forms heterodimers with other receptors (including PPARG) and is involved in lipid metabolism [57] and homeostasis. Moreover, 145 GWAS loci (Table S9) overlap RXRA binding peaks, suggesting that the GWAS variants might affect lipids (partially) through affecting the binding activity of RXRA. While the *RXRA*-associated variant has been previously implicated as a GWAS locus [58], our study demonstrates its role in lipid biology through its regulatory influence on other lipid-associated genes.

Multi-layer functional integration reveals regulatory mechanisms at GWAS loci

Motivated by our finding that integrating chromatin interaction improved eQTL colocalizations, we further brought together multiple lines of functional evidence at each GWAS locus for mechanistic inference. We started with the list of genes with evidence for

both eQTL colocalization in the liver or adipose and credible set physical interactions. We next annotated each variant in the 95% credible set with various indicators of regulatory function, including its open chromatin status in liver or adipose-related cell types, its proximity to a promoter or an enhancer, and its RegulomeDB regulation probability [59] (see Table S10 for the complete list of annotations used). To account for complexities of regulatory mechanisms and limitations of functional datasets, we combined evidence across these datasets to prioritize variants at GWAS loci (Figure 4A). Specifically, we prioritized variants with at least three independent lines of functional evidence (chromatin openness, physically interaction with target genes, and promoter/enhancer status in liver or adipose), with at least two being in the same tissue with colocalization with the target gene, and with a RegulomeDB score > 0.5. Applying this simple procedure to lipid GWAS we identified 13 candidate loci, each with the strongest multi-layer evidence pointing to a single functional variant (Table 1). Below we describe two examples to highlight key features of this multi-layer integration framework.

RRBP1 (ribosomal binding protein 1) could be identified from eQTL colocalization alone, but our multi-layer integration approach strengthened the conclusion via convergent evidence from various sources (Figure 4B). The *RRBP1* eQTL signals in the liver colocalize with LDL, TC, and nonHDL GWAS signals. The 'T' allele of the lead variant (chr20:17,844,684, hg19) decreases *RRBP1* expression levels and increases LDL, TC, and nonHDL levels. This lead variant is in open chromatin in HLC, and physically interacts with the *RRBP1* promoter (250kb away) in adipose and HepG2. All these data consistently point to *RRBP1* as the functional gene underlying this locus. RRBP1 specifically tethers the endoplasmic reticulum to the mitochondria in the liver--an interaction that is enriched in hepatocytes--and regulates

1184 very low density lipoprotein (vLDL) levels [60]. Rare variants in *RRBP1* are associated with

1185 LDL in humans [61] and silencing *RRBP1* in liver affects lipid homeostasis in mice [60].

1186

1187 *Figure 4. An easy-to-implement multi-layer framework to prioritize functional variants at*

1188 *GWAS loci. A. Variant annotation and prioritization scheme at each credible set. B. Evidence*

1189 *for gene RRBP1 from functional genomics data. The LDL GWAS locus at this region is an*

1190 *eQTL for gene RRBP1 in the liver (second row). Variants in the credible set of this locus*

1191 *interact with the gene promoter in both adipose and HepG2 Capture-C data. The interacting*

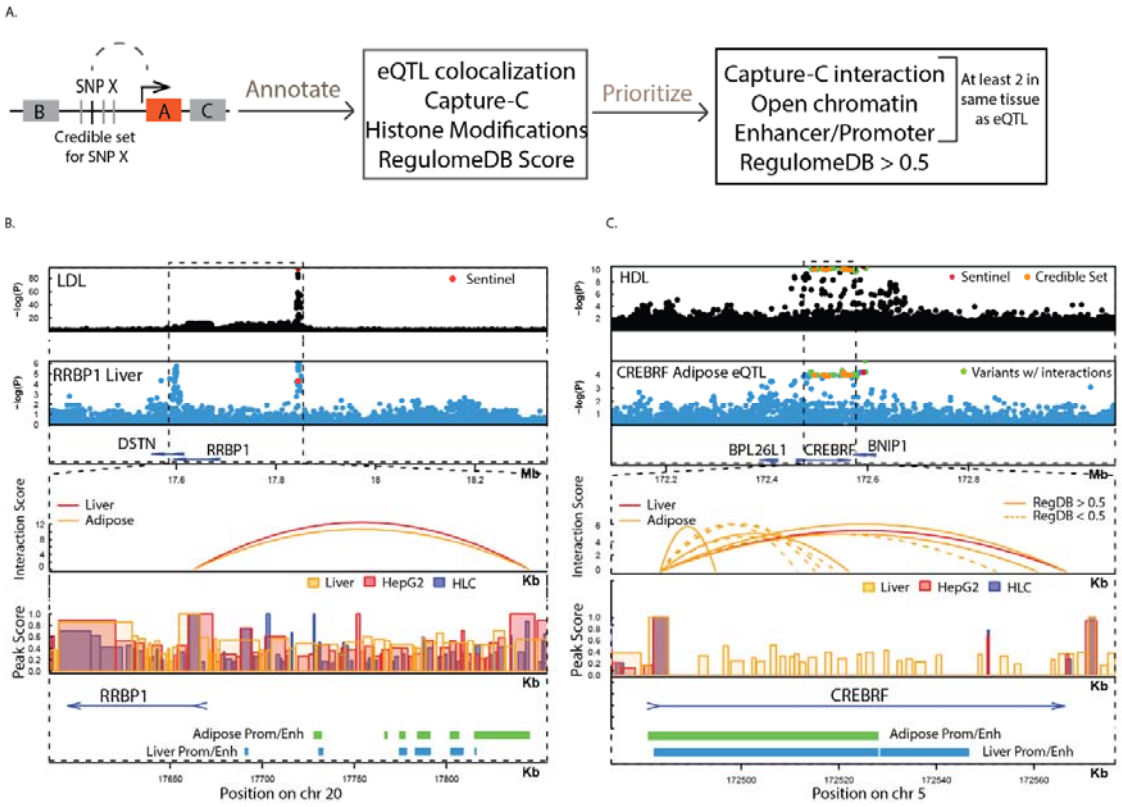
1192 *variant is also in an open chromatin peak in three liver-related cell types. C. Multiple*

1193 *sources of functional genomics data support CREBRF as a gene contributing to HDL levels.*

1194 *The HDL GWAS locus at this region is an eQTL for gene CREBRF in adipose (second row).*

1195 *Variants in the credible set at this locus interact with the CREBRF promoter in adipose. The*

1196 *interacting variant is also in open chromatin in liver-related cell types.*



1197

CREBRF (CREB3 regulatory factor) demonstrates the power of our multi-layer integration framework in prioritizing functional variants (Figure 4C). The eQTL signals of *CREBRF* colocalized with a GWAS locus for HDL with 30 candidate variants. In contrast, our multi-layer approach identified a single candidate variant (chr5:172,566,698) at this locus that physically interacts with the *CREBRF* promoter in adipose, was predicted to be a regulatory element (RegulomeDB score=0.91). Consistent with the index variant (chr5:172,591,337), the allele 'A' at this functional variant increased HDL levels and increased *CREBRF* expression in adipose. Missense variants in *CREBRF* have been linked to body mass index, and the gene has been linked to obesity risk in Samoans [62].

Finally, to compare the power of functional fine-mapping with trans-ancestry fine-mapping, we applied our prioritization rule to credible sets derived from European-only meta-analysis. The 111 variants prioritized by our rule described above (including multiple variants in the same credible set) were all found in the multi-ancestry credible sets, representing a 3.7 fold enrichment ($P < 1e-04$ derived from 10000 permutations randomly sampling variants from the European-only credible sets). This convergence of complementary approaches to the same smaller set of variants highlights the power of multi-ancestry datasets as an approach to narrow in on functional variants.

Discussion

Here we integrate the largest multi-population lipid GWAS to date with a wide array of functional genomic resources to understand how noncoding genetic variation affects lipids through gene regulation. Specifically, we identify 1,076 genes whose eQTL signals

colocalize with lipid GWAS signals and demonstrate how physical chromatin interaction can improve standard eQTL-based colocalization. We assess tissue-specific enrichments of lipid GWAS signals and demonstrate the selective importance of adipose in HDL and triglyceride biology. We examine binding site enrichments of 161 TFs in lipid GWAS and expand our understanding of lipid GWAS loci (e.g., *RXRA*) in the context of TF activity. Finally, we build a simple and interpretable prioritization framework that automatically combines multiple lines of evidence from orthogonal datasets, pinpointing a single functional variant at each of 13 lipid-associated loci (e.g., *RRBP1* and *CREBRF*). While there are studies that interpret lipid GWAS associations [21,63,64], the size of our multi-population GWAS and multi-layer functional integration represent a comprehensive effort and an important step forward in this direction.

Our multi-layer analysis has two key strengths. First, despite a large array of functional genomic resources being embedded, our analysis produces results with high consistency. For example, the selective enrichment of adipose in HDL and TG identified by S-LDSC is confirmed by our eQTL-based colocalization and TF binding site overlap. Another example of consistency is the multi-layer prioritization of *RRBP1*, which can be identified from eQTL-based colocalization alone and it is further validated by chromatin openness and interaction. Such convergent evidence from various sources improves the confidence of our findings. Second, our analysis highlights that combining multiple layers of regulatory information can improve sensitivity to prioritize functional genes and variants. For example, we refined eQTL colocalized genes (1,076) to a smaller set of functional genes (233) through integration with promoter Capture-C data. Another example of sensitivity is *CREBRF*, where eQTL-based colocalization implicates 30 candidate variants and adding other regulatory layers points to a

single functional variant. Moving forward, we expect these two features will serve as useful guidelines for future integrative genomic analyses of other traits.

Our results rely on the breadth and accuracy of functional genomic datasets used in our analyses. First, unlike our lipid GWAS, current functional datasets [65] are limited both in sample size and ancestral diversity, which can affect discovery and replication of regulatory mechanisms in diverse populations. Second, some functional datasets are generated at limited resolution. For example, our colocalizations are based on eQTLs from bulk tissue RNA-seq [8], which may miss detailed cell types and biological processes in which lipid-associated SNPs regulate gene expression [66]. Third, some functional datasets are not available across the full spectrum of human tissues and cell types. For example, our chromatin-chromatin interaction analysis only examines a few cell types in two known lipid-related tissues, producing results that may be biased towards known lipid biology. As more comprehensive and accurate functional genomic resources are becoming publicly available in diverse cellular contexts and ancestry groups, the resolution and power of integrative analyses like ours will be markedly increased.

Other limitations of this study stem from computational methods embedded in our framework. First, the colocalization approach 'coloc' assumes one causal variant per locus, whereas recent studies suggest extensive allelic heterogeneity [67] consistent with a model of a milieu of related transcription factors binding within a single locus. Accounting for allelic heterogeneity in summary statistics-based colocalization typically requires modelling multiple correlated SNPs through LD matrix [68], which is computationally intensive in large-scale analyses derived from many cohorts with diverse ancestries, like the multi-population GWAS examined here. Second, due to restricted access to individual genotypes of 201

cohorts, we cannot produce multi-population LD scores within GLGC but have to use European-based LD scores in all S-LDSC analyses. This approach, though less rigorous in principle, provides robust results in practice (as confirmed by our ancestry-specific analysis), largely because 79% of cohorts in GLGC are of European descent [12]. That said, we caution that the same approach might fall short in ancestrally diverse studies with few European individuals [69]. Third, our multi-layer variant prioritization framework is built on a series of simple rules that are easy to implement on large datasets. This approach could possibly be formalized as statistical models (e.g., priors in Bayesian methods [49]), but certainly simplify computation and improve scalability of our framework. Despite the technical limitations, our approach here can serve as a useful benchmark for future development of methods with improved statistical rigor and computation efficiency.

In summary, mapping noncoding genetic variation of complex traits to biological functions can benefit greatly from thorough integration of multiple layers of functional genomics, as demonstrated in the present study. Although tested on lipids only, our integrative framework is straightforward to implement more broadly on many other phenotypes, yielding functional insights of heritable traits and diseases in humans.

Description of Supplemental Data

Supplemental data include seven figures and ten tables, and study-specific acknowledgements.

Declaration of 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.

Acknowledgments

GMP, PN and CW are supported by NHLBI R01HL127564. GMP and PN are supported by R01HL142711. AG acknowledge support from the Wellcome Trust (201543/B/16/Z), European Union Seventh Framework Programme FP7/2007-2013 under grant agreement no. HEALTH-F2-2013-601456 (CVGenes@Target) & the TriPartite Immunometabolism Consortium [TrIC]-Novo Nordisk Foundation's Grant number NNF15CC0018486. JMM is supported by American Diabetes Association Innovative and Clinical Translational Award 1-19-ICTS-068. SR was supported by the Academy of Finland Center of Excellence in Complex Disease Genetics (Grant No 312062), the Finnish Foundation for Cardiovascular Research, the Sigrid Juselius Foundation and University of Helsinki HiLIFE Fellow and Grand Challenge grants. EW was supported by the Finnish innovation fund Sitra (EW) and Finska Läkaresällskapet. CNS was supported by American Heart Association Postdoctoral Fellowships 15POST24470131 and 17POST33650016. Charles N Rotimi is supported by

Z01HG200362. Zhe Wang, Michael Presuss, and Ruth JF Loos are supported by R01HL142302. NJT is a Wellcome Trust Investigator (202802/Z/16/Z), is the PI of the Avon Longitudinal Study of Parents and Children (MRC & WT 217065/Z/19/Z), is supported by the University of Bristol NIHR Biomedical Research Centre (BRC-1215-2001), the MRC Integrative Epidemiology Unit (MC_UU_00011) and works within the CRUK Integrative Cancer Epidemiology Programme (C18281/A19169). Ruth E Mitchell is a member of the MRC Integrative Epidemiology Unit at the University of Bristol funded by the MRC (MC_UU_00011/1). Simon Haworth is supported by the UK National Institute for Health Research Academic Clinical Fellowship. Paul S. de Vries was supported by American Heart Association grant number 18CDA34110116. Julia Ramierz acknowledges support by the People Programme of the European Union's Seventh Framework Programme grant n° 608765 and Marie Sklodowska-Curie grant n° 786833. Maria Sabater-Lleal is supported by a Miguel Servet contract from the ISCIII Spanish Health Institute (CP17/00142) and co-financed by the European Social Fund. Jian Yang is funded by the Westlake Education Foundation. Olga Giannakopoulou has received funding from the British Heart Foundation (BHF) (FS/14/66/3129). Dr. Sander W. van der Laan is funded through grants from the Netherlands CardioVascular Research Initiative of the Netherlands Heart Foundation (CVON 2011/B019 and CVON 2017-20: Generating the best evidence-based pharmaceutical targets for atherosclerosis [GENIUS I&II]). We are thankful for the support of the FP7 EU project CVgenes@target (HEALTH-F2-2013-601456), ERA-CVD program 'druggable-MI-targets' (grant number: 01KL1802) and the Leducq Fondation 'PlaqOmics'. Study-specific acknowledgements are available in the **Supplementary Material**.

Web Resources

GLGC 2021 summary statistics: http://csg.sph.umich.edu/willer/public/glgc-lipids2021/

GTEx v8 summary statistics: https://www.gtexportal.org/home/datasets

coloc: https://cran.r-project.org/web/packages/coloc

liftOver: https://genome.ucsc.edu/cgi-bin/hgLiftOver

HiCUP: https://www.bioinformatics.babraham.ac.uk/projects/hicup/

CHiCAGO: https://www.bioconductor.org/packages/release/bioc/html/Chicago.html

GenomicRanges: https://bioconductor.org/packages/release/bioc/html/GenomicRanges.html

bedtools: https://bedtools.readthedocs.io/en/latest/

ClusterProfiler: https://guangchuangyu.github.io/clusterProfiler

biomaRt: https://bioconductor.org/packages/release/bioc/html/biomaRt.html

ClinVar: https://www.ncbi.nlm.nih.gov/clinvar/

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

LDSC software: https://github.com/bulik/ldsc

LD scores and related annotations: https://data.broadinstitute.org/alkesgroup/LDSCORE/

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

RSS-NET: https://github.com/SUwonglab/rss-net

Adipose single cell data:

https://singlecell.broadinstitute.org/single_cell/study/SCP133/human-adipose-svf-single-cell

fgsea: http://bioconductor.org/packages/release/bioc/html/fgsea.html

GREGOR: https://genome.sph.umich.edu/wiki/GREGOR

RegulomeDB: https://regulomedb.org/regulome-search/

Data and Code Availability

HLC Capture-C data is available at

https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE189026

References

1. Gallagher, M.D., and Chen-Plotkin, A.S. (2018). The Post-GWAS Era: From Association to Function. Am. J. Hum. Genet. *102*, 717–730.

2. Cano-Gamez, E., and Trynka, G. (2020). From GWAS to Function: Using Functional Genomics to Identify the Mechanisms Underlying Complex Diseases. Front. Genet. *11*, 424.

3. Schaid, D.J., Chen, W., and Larson, N.B. (2018). From genome-wide associations to candidate causal variants by statistical fine-mapping. Nat. Rev. Genet. *19*, 491–504.

4. Smemo, S., Tena, J.J., Kim, K.-H., Gamazon, E.R., Sakabe, N.J., Gómez-Marín, C., Aneas, I., Credidio, F.L., Sobreira, D.R., Wasserman, N.F., et al. (2014). Obesity-associated variants within FTO form long-range functional connections with IRX3. Nature *507*, 371–375.

5. Musunuru, K., Strong, A., Frank-Kamenetsky, M., Lee, N.E., Ahfeldt, T., Sachs, K.V., Li, X., Li, H., Kuperwasser, N., Ruda, V.M., et al. (2010). From noncoding variant to phenotype via SORT1 at the 1p13 cholesterol locus. Nature *466*, 714–719.

6. Hekselman, I., and Yeger-Lotem, E. (2020). Mechanisms of tissue and cell-type specificity in heritable traits and diseases. Nat. Rev. Genet. *21*, 137–150.

7. ENCODE Project Consortium (2012). An integrated encyclopedia of DNA elements in the human genome. Nature *489*, 57–74.

8. The GTEx Consortium (2020). The GTEx Consortium atlas of genetic regulatory effects

1412 across human tissues. Science *369*, 1318–1330.

1413 9. Loos, R.J.F., and Yeo, G.S.H. (2021). The genetics of obesity: from discovery to biology.
1414 Nat. Rev. Genet. doi: 10.1038/s41576-021-00414-z

1415 10. Huo, Y., Li, S., Liu, J., Li, X., and Luo, X.-J. (2019). Functional genomics reveal gene
1416 regulatory mechanisms underlying schizophrenia risk. Nat. Commun. *10*, 1–19.

1417 11. Willer, C.J., Sanna, S., Jackson, A.U., Scuteri, A., Bonnycastle, L.L., Clarke, R., Heath,
1418 S.C., Timpson, N.J., Najjar, S.S., Stringham, H.M., et al. (2008). Newly identified loci that
1419 influence lipid concentrations and risk of coronary artery disease. Nat. Genet. *40*, 161–169.

1420 12. Sarah E Graham, Shoa L Clarke, Kuan-Han H Wu, Stavroula Kanoni, Greg JM Zajac,
1421 Shweta Ramdas, Ida Surakka, Ioanna Ntalla, Sailaja Vedantam,, Thomas W Winkler, et al.
1422 (in press). The power of genetic diversity in genome-wide association studies of lipids.
1423 Nature. doi: 10.1038/s41586-021-04064-3

1424 13. McCarthy, S., Das, S., Kretzschmar, W., Delaneau, O., Wood, A.R., Teumer, A., Kang,
1425 H.M., Fuchsberger, C., Danecek, P., Sharp, K., et al. (2016). A reference panel of 64,976
1426 haplotypes for genotype imputation. Nat. Genet. *48*, 1279–1283.

1427 14. 1000 Genomes Project Consortium, and Adam Auton, Lisa D Brooks, Richard M Durbin,
1428 Erik P Garrison, Hyun Min Kang, Jan O Korbel, Jonathan L Marchini, Shane McCarthy, Gil
1429 A McVean, Gonçalo R Abecasis (2015). A global reference for human genetic variation.
1430 Nature *526*, 68–74.

1431 15. Mägi, R., Horikoshi, M., Sofer, T., Mahajan, A., Kitajima, H., Franceschini, N.,
1432 McCarthy, M.I., COGENT-Kidney Consortium, T2D-GENES Consortium, and Morris, A.P.
1433 (2017). Trans-ethnic meta-regression of genome-wide association studies accounting for
1434 ancestry increases power for discovery and improves fine-mapping resolution. Hum. Mol.
1435 Genet. *26*, 3639–3650.

1436 16. The International HapMap 3 Consortium (2010). Integrating common and rare genetic
1437 variation in diverse human populations. Nature *467*, 52–58.

1438 17. Giambartolomei, C., Vukcevic, D., Schadt, E.E., Franke, L., Hingorani, A.D., Wallace,
1439 C., and Plagnol, V. (2014). Bayesian test for colocalisation between pairs of genetic
1440 association studies using summary statistics. PLoS Genet. *10*, e1004383.

1441 18. Kuhn, R.M., Haussler, D., and Kent, W.J. (2013). The UCSC genome browser and
1442 associated tools. Brief. Bioinform. *14*, 144–161.

1443 19. Çalışkan, M., Manduchi, E., Rao, H.S., Segert, J.A., Beltrame, M.H., Trizzino, M., Park,
1444 Y., Baker, S.W., Chesi, A., Johnson, M.E., et al. (2019). Genetic and Epigenetic Fine
1445 Mapping of Complex Trait Associated Loci in the Human Liver. Am. J. Hum. Genet. *105*,
1446 89–107.

1447 20. Chesi, A., Wagley, Y., Johnson, M.E., Manduchi, E., Su, C., Lu, S., Leonard, M.E.,
1448 Hodge, K.M., Pippin, J.A., Hankenson, K.D., et al. (2019). Genome-scale Capture C
1449 promoter interactions implicate effector genes at GWAS loci for bone mineral density. Nat.
1450 Commun. *10*, 1260.

21. Selvarajan, I., Toropainen, A., Garske, K.M., López Rodríguez, M., Ko, A., Miao, Z., Kaminska, D., Õunap, K., Örd, T., Ravindran, A., et al. (2021). Integrative analysis of liver-specific non-coding regulatory SNPs associated with the risk of coronary artery disease. Am. J. Hum. Genet. *108*, 411–430.

22. Pashos, E.E., Park, Y., Wang, X., Raghavan, A., Yang, W., Abbey, D., Peters, D.T., Arbelaez, J., Hernandez, M., Kuperwasser, N., et al. (2017). Large, Diverse Population Cohorts of hiPSCs and Derived Hepatocyte-like Cells Reveal Functional Genetic Variation at Blood Lipid-Associated Loci. Cell Stem Cell *20*, 558–570.e10.

23. Pan, D.Z., Garske, K.M., Alvarez, M., Bhagat, Y.V., Boocock, J., Nikkola, E., Miao, Z., Raulerson, C.K., Cantor, R.M., Civelek, M., et al. (2018). Integration of human adipocyte chromosomal interactions with adipose gene expression prioritizes obesity-related genes from GWAS. Nat. Commun. *9*, 1512.

24. Wingett, S., Ewels, P., Furlan-Magaril, M., Nagano, T., Schoenfelder, S., Fraser, P., and Andrews, S. (2015). HiCUP: pipeline for mapping and processing Hi-C data. F1000Res. *4*, 1310.

25. Cairns, J., Freire-Pritchett, P., Wingett, S.W., Várnai, C., Dimond, A., Plagnol, V., Zerbino, D., Schoenfelder, S., Javierre, B.-M., Osborne, C., et al. (2016). CHiCAGO: robust detection of DNA looping interactions in Capture Hi-C data. Genome Biol. *17*, 127.

26. Harrow, J., Frankish, A., Gonzalez, J.M., Tapanari, E., Diekhans, M., Kokocinski, F., Aken, B.L., Barrell, D., Zadissa, A., Searle, S., et al. (2012). GENCODE: the reference human genome annotation for The ENCODE Project. Genome Res. *22*, 1760–1774.

27. Lawrence, M., Huber, W., Pagès, H., Aboyoun, P., Carlson, M., Gentleman, R., Morgan, M.T., and Carey, V.J. (2013). Software for computing and annotating genomic ranges. PLoS Comput. Biol. *9*, e1003118.

28. Leung, D., Jung, I., Rajagopal, N., Schmitt, A., Selvaraj, S., Lee, A.Y., Yen, C.-A., Lin, S., Lin, Y., Qiu, Y., et al. (2015). Integrative analysis of haplotype-resolved epigenomes across human tissues. Nature *518*, 350–354.

29. Yu, G., Wang, L.-G., Han, Y., and He, Q.-Y. (2012). clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. OMICS: A Journal of Integrative Biology *16*, 284–287.

30. Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M., and Tanabe, M. (2016). KEGG as a reference resource for gene and protein annotation. Nucleic Acids Res. *44*, D457–D462.

31. Durinck, S., Spellman, P.T., Birney, E., and Huber, W. (2009). Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. Nat. Protoc. *4*, 1184–1191.

32. Landrum, M.J., Chitipiralla, S., Brown, G.R., Chen, C., Gu, B., Hart, J., Hoffman, D., Jang, W., Kaur, K., Liu, C., et al. (2020). ClinVar: improvements to accessing data. Nucleic Acids Res. *48*, D835–D844.

33. Barbeira, A.N., Dickinson, S.P., Bonazzola, R., Zheng, J., Wheeler, H.E., Torres, J.M., Torstenson, E.S., Shah, K.P., Garcia, T., Edwards, T.L., et al. (2018). Exploring the

1491 phenotypic consequences of tissue specific gene expression variation inferred from GWAS
1492 summary statistics. Nat. Commun. *9*, 1825.

1493 34. Hindy, G., Dornbos, P., Chaffin, M.D., Liu, D.J., Wang, M.X., Selvaraj, M.S., Zhang, D.,
1494 Park, J., Aguilar-Salinas, C.A., Antonacci-Fulton, L., et al. (2021). Rare coding variants in 35
1495 genes associate with circulating lipid levels – a multi-ancestry analysis of 170,000 exomes.
1496 bioRxiv 10.1101/2020.12.22.423783

1497 35. Finucane, H.K., Bulik-Sullivan, B., Gusev, A., Trynka, G., Reshef, Y., Loh, P.-R.,
1498 Anttila, V., Xu, H., Zang, C., Farh, K., et al. (2015). Partitioning heritability by functional
1499 annotation using genome-wide association summary statistics. Nat. Genet. *47*, 1228.

1500 36. Kanai, M., Akiyama, M., Takahashi, A., Matoba, N., Momozawa, Y., Ikeda, M., Iwata,
1501 N., Ikegawa, S., Hirata, M., Matsuda, K., et al. (2018). Genetic analysis of quantitative traits
1502 in the Japanese population links cell types to complex human diseases. Nat. Genet. *50*, 390–
1503 400.

1504 37. Schmidt, E.M., Zhang, J., Zhou, W., Chen, J., Mohlke, K.L., Eugene Chen, Y., and
1505 Willer, C.J. (2015). GREGOR: evaluating global enrichment of trait-associated variants in
1506 epigenomic features using a systematic, data-driven approach. Bioinformatics *31*, 2601–
1507 2606.

1508 38. MacParland, S.A., Liu, J.C., Ma, X.-Z., Innes, B.T., Bartczak, A.M., Gage, B.K., Manuel,
1509 J., Khuu, N., Echeverri, J., Linares, I., et al. (2018). Single cell RNA sequencing of human
1510 liver reveals distinct intrahepatic macrophage populations. Nat. Commun. *9*, 4383.

1511 39. Korotkevich, G., Sukhov, V., Budin, N., Shpak, B., Artyomov, M.N., and Sergushichev,
1512 A. Fast gene set enrichment analysis. bioRxiv https://doi.org/10.1101/060012

1513 40. Neumeyer, S., Hemani, G., and Zeggini, E. (2020). Strengthening Causal Inference for
1514 Complex Disease Using Molecular Quantitative Trait Loci. Trends Mol. Med. *26*, 232–241.

1515 41. Y C Loraine Tung, Giles S H Yeo, Stephen O'Rahilly, Anthony P Coll (2014). Obesity
1516 and FTO: Changing Focus at a Complex Locus. Cell Metab. *20*, 710–718.

1517 42. Blair, D.R., Lyttle, C.S., Mortensen, J.M., Bearden, C.F., Jensen, A.B., Khiabanian, H.,
1518 Melamed, R., Rabadan, R., Bernstam, E.V., Brunak, S., et al. (2013). A nondegenerate code
1519 of deleterious variants in Mendelian loci contributes to complex disease risk. Cell 155(1):70-
1520 80.

1521 43. Teslovich, T.M., Musunuru, K., Smith, A.V., Edmondson, A.C., Stylianou, I.M., Koseki,
1522 M., Pirruccello, J.P., Ripatti, S., Chasman, D.I., Willer, C.J., et al. (2010). Biological, clinical
1523 and population relevance of 95 loci for blood lipids. Nature *466*, 707–713.

1524 44. David U Gorkin, Danny Leung, Bing Ren (2014). The 3D Genome in Transcriptional
1525 Regulation and Pluripotency. Cell Stem Cell *14*, 762–775.

1526 45. Wakefield, J. (2009). Bayes factors for genome-wide association studies: comparison
1527 with P-values. Genet. Epidemiol. *33*, 79–86.

1528 46. Yu, J., Hu, M., and Li, C. (2019). Joint analyses of multi-tissue Hi-C and eQTL data
1529 demonstrate close spatial proximity between eQTLs and their target genes. BMC Genet. *20*,

43.

47. Matharu, N.K., and Ahanger, S.H. (2015). Chromatin Insulators and Topological Domains: Adding New Dimensions to 3D Genome Architecture. Genes *6*, 790–811.

48. Pers, T.H., Karjalainen, J.M., Chan, Y., Westra, H.-J., Wood, A.R., Yang, J., Lui, J.C., Vedantam, S., Gustafsson, S., Esko, T., et al. (2015). Biological interpretation of genome-wide association studies using predicted gene functions. Nat. Commun. *6*, 5890.

49. Zhu, X., Duren, Z., and Wong, W.H. (2021). Modeling regulatory network topology improves genome-wide analyses of complex human traits. Nat. Commun. *12*, 2851.

50. Zhang, T., Chen, J., Tang, X., Luo, Q., Xu, D., and Yu, B. (2019). Interaction between adipocytes and high-density lipoprotein:new insights into the mechanism of obesity-induced dyslipidemia and atherosclerosis. Lipids Health Dis. *18*, 223.

51. A. D. Sniderman, K. Cianflone, P. Arner, L. K. M. Summers, and K. N. Frayn (1998) The Adipocyte, Fatty Acid Trapping, and Atherogenesis. Arteriosclerosis, Thrombosis, and Vascular Biology. 18:147–151.

52. Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., et al. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. U. S. A. *102*, 15545–15550.

53. Degtyareva, A.O., Antontseva, E.V., and Merkulova, T.I. (2021). Regulatory SNPs: Altered Transcription Factor Binding Sites Implicated in Complex Traits and Diseases. Int. J. Mol. Sci. *22*,.

54. Tripathi, M., Yen, P.M., and Singh, B.K. (2020). Estrogen-Related Receptor Alpha: An Under-Appreciated Potential Target for the Treatment of Metabolic Diseases. Int. J. Mol. Sci. *21*,.

55. Saykally, J.N., Dogan, S., Cleary, M.P., and Sanders, M.M. (2009). The ZEB1 Transcription Factor Is a Novel Repressor of Adiposity in Female Mice. PLoS One *4*, e8460.

56. Gubelmann, C., Schwalie, P.C., Raghav, S.K., Röder, E., Delessa, T., Kiehlmann, E., Waszak, S.M., Corsinotti, A., Udin, G., Holcombe, W., et al. (2014). Identification of the transcription factor ZEB1 as a central component of the adipogenic gene regulatory network. Elife *3*, e03346.

57. Neuschwander-Tetri, B.A. (2015). Retinoid X receptor: the forgotten partner in regulating lipid metabolism? Am. J. Clin. Nutr. *102*, 5–6.

58. Peloso, G.M., Demissie, S., Collins, D., Mirel, D.B., Gabriel, S.B., Cupples, L.A., Robins, S.J., Schaefer, E.J., and Brousseau, M.E. (2010). Common genetic variation in multiple metabolic pathways influences susceptibility to low HDL-cholesterol and coronary heart disease. J. Lipid Res. *51*, 3524–3532.

59. Boyle, A.P., Hong, E.L., Hariharan, M., Cheng, Y., Schaub, M.A., Kasowski, M., Karczewski, K.J., Park, J., Hitz, B.C., Weng, S., et al. (2012). Annotation of functional variation in personal genomes using RegulomeDB. Genome Research *22*, 1790–1797.

60. Anastasia, I., Ilacqua, N., Raimondi, A., Lemieux, P., Ghandehari-Alavijeh, R., Faure, G., Mekhedov, S.L., Williams, K.J., Caicci, F., Valle, G., et al. (2021). Mitochondria-rough-ER contacts in the liver regulate systemic lipid homeostasis. Cell Rep. *34*, 108873.

61. Jurgens, S.J., Choi, S.H., Morrill, V.N., Chaffin, M., Pirruccello, J.P., Halford, J.L., Weng, L.-C., Nauffal, V., Roselli, C., Hall, A.W., et al. (2020). Rare Genetic Variation Underlying Human Diseases and Traits: Results from 200,000 Individuals in the UK Biobank.

62. Minster, R.L., Hawley, N.L., Su, C.-T., Sun, G., Kershaw, E.E., Cheng, H., Buhule, O.D., Lin, J., Reupena, M.S., Viali, S. 'itea, et al. (2016). A thrifty variant in CREBRF strongly influences body mass index in Samoans. Nat. Genet. *48*, 1049–1054.

63. Klarin, D., Damrauer, S.M., Cho, K., Sun, Y.V., Teslovich, T.M., Honerlaw, J., Gagnon, D.R., DuVall, S.L., Li, J., Peloso, G.M., et al. (2018). Genetics of blood lipids among ~300,000 multi-ethnic participants of the Million Veteran Program. Nat. Genet. *50*, 1514–1523.

64. Li, Z., Votava, J.A., Zajac, G.J.M., Nguyen, J.N., Leyva Jaimes, F.B., Ly, S.M., Brinkman, J.A., De Giorgi, M., Kaul, S., Green, C.L., et al. (2020). Integrating Mouse and Human Genetic Data to Move beyond GWAS and Identify Causal Genes in Cholesterol Metabolism. Cell Metab. *31*, 741–754.e5.

65. Varshney, A., VanRenterghem, H., Orchard, P., Boyle, A.P., Stitzel, M.L., Ucar, D., and Parker, S.C.J. (2019). Cell Specificity of Human Regulatory Annotations and Their Genetic Effects on Gene Expression. Genetics *211*, 549–562.

66. van der Wijst, M.G.P., de Vries, D.H., Groot, H.E., Trynka, G., Hon, C.C., Bonder, M.J., Stegle, O., Nawijn, M.C., Idaghdour, Y., van der Harst, P., et al. (2020). Science Forum: The single-cell eQTLGen consortium. eLife 2020;9:e52155.

67. Arvanitis, M., Tayeb, K., Strober, B.J., and Battle, A. (2021). Redefining tissue specificity of genetic regulation of gene expression in the presence of allelic heterogeneity. medRxiv 2021.06.28.21259545.

68. Zhu, X., and Stephens, M. (2017). Bayesian large-scale multiple regression with summary statistics from genome-wide association studies. Ann. Appl. Stat. *11*, 1561–1592.

69. Wojcik, G.L., Graff, M., Nishimura, K.K., Tao, R., Haessler, J., Gignoux, C.R., Highland, H.M., Patel, Y.M., Sorokin, E.P., Avery, C.L., et al. (2019). Genetic analyses of diverse populations improves discovery for complex traits. Nature *570*, 514–518.

Tables

Table 1. Thirteen prioritized loci with highest confidence of a single functional variant in the credible set. The 'sentinel' column represents the lead variant at the locus. 'Prioritized var' represents the prioritized variant in the credible set. Columns 5-8 represent overlap of the functional variant with open chromatin ('Open'), capture-C ('CapC') interactions with the candidate gene, enhancer and promoter marks from Roadmap in liver ('Liver'), adipose ('Ad'), both or none of these datasets. The 'RegDB' column represents the RegulomeDB score of the prioritized variant.

Gene Name	**Tissue**	**Sentinel**	**Prioritized Var**	**Open**	**CapC**	**Enhancer**	**Promoter**	**RegDB**
CEP68	Adipose	2:65284231	65279414	Liver	Liver	None	Ad	0.5896
TIPARP	Adipose	3:156797941	156795408	Both	Both	Ad	Liver	0.705
CREBRF	Adipose	5:172591337	172566698	Liver	Ad	None	Both	0.9124
PALM2	Adipose	9:112556911	112556911	Both	Ad	Both	None	0.6091
MEGF9	Adipose	9:123481206	123421556	Liver	Ad	None	Liver	0.9933
GBF1	Liver	10:104142294	104107191	Ad	Ad	Both	Ad	0.705
MICAL2	Liver	11:12071855	12221016	Liver	Liver	Liver	Ad	0.6018
ACP2	Liver	11:47278917	47276350	Ad	Liver	Liver	Ad	0.6091
PTPRJ	Adipose	11:48021778	48011180	Liver	Ad	Liver	Ad	0.8797
NFATC2IP	Adipose	16:28899411	28883327	Liver	Liver	None	Both	0.6091
HELZ	Liver	17:65109591	65156919	Liver	Liver	Both	Ad	0.60906
FAM210A	Liver	18:13725674	13725674	Liver	Liver	Both	Ad	0.7571
RRBP1	Liver	20:17844684	17844684	Both	Ad	Both	Ad	0.6091