Universiteit Leiden
The Netherlands

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

Ramdas, S.; Judd, J.; Graham, S.E.; Kanoni, S.; Wang, Y.X.; Surakka, I.; ... ; Lee, H.

Citation
Ramdas, S., Judd, J., Graham, S. E., Kanoni, S., Wang, Y. X., Surakka, I., ... Lee, H. (2022). A multi-layer functional genomic analysis to understand noncoding genetic variation in lipids. *American Journal Of Human Genetics*, *109*(8), 1366-1387. doi:10.1016/j.ajhg.2022.06.012

Version: Publisher's Version
License: Licensed under Article 25fa Copyright Act/Law (Amendment Taverne)
Downloaded from: https://hdl.handle.net/1887/3479845

Note: To cite this publication please use the final published version (if applicable).

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

Authors

Shweta Ramdas, Jonathan Judd, Sarah E. Graham, ..., Cristen J. Willer, Xiang Zhu, Christopher D. Brown

Correspondence

xiangzhu@psu.edu (X.Z.), chrbro@upenn.edu (C.D.B.)

In this study, we present a multi-layer framework to combine the largest multi-ancestry GWAS to date on lipid levels with both transcriptomic and epigenomic datasets to prioritize regulatory variants, effector genes, cell types, and tissues with strong functional relevance to lipid biology.

Ramdas et al., 2022, The American Journal of Human Genetics *109*, 1366–1387
August 4, 2022 © 2022 American Society of Human Genetics.
https://doi.org/10.1016/j.ajhg.2022.06.012

CellPress

ARTICLE

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

Shweta Ramdas,[1,420] Jonathan Judd,[2,420] Sarah E. Graham,[3,420] Stavroula Kanoni,[4,420] Yuxuan Wang,[5,420] 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 J.M. Zajac,[13] Kuan-Han H. Wu,[14] Ioanna Ntalla,[15] Qin Hui,[16,17] Derek Klarin,[10,18,19] Austin T. Hilliard,[6] Zeyuan Wang,[16,17] Chao Xue,[3] Gudmar Thorleifsson,[20] Anna Helgadottir,[20] Daniel F. Gudbjartsson,[20,21] Hilma Holm,[20] Isleifur Olafsson,[22] Mi Yeong Hwang,[23] Sohee Han,[23] Masato Akiyama,[24,25] Saori Sakaue,[26,27,28] Chikashi Terao,[29] Masahiro Kanai,[10,24,30] Wei Zhou,[10,14,31] Ben M. Brumpton,[32,33,34] Humaira Rasheed,[32,33,35] Aki S. Havulinna,[36,37] Yogasudha Veturi,[38]

(Author list continued on next page)

Summary

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

Introduction

Most GWAS findings have not directly led to mechanistic interpretations, largely because approximately 90% of GWAS associations map to noncoding 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]

[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 94305, 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, John Vane Science Centre, Charterhouse Square, London EC1M 6BQ, UK; [5]Department of Biostatistics, Boston University School of Public Health, 801 Massachusetts Avenue, Boston, MA 02118, USA; [6]VA Palo Alto Health Care Systems, Palo Alto, CA, 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, MA 02115, USA; [10]Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, 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, EC1M 6BQ London, UK; [16]Department of Epidemiology, Emory University Rollins School of Public Health, Atlanta, GA, 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]deCODE genetics/Amgen, Inc., Sturlugata 8, Reykjavik 102, Iceland; [21]School of Engineering and Natural Sciences, University of Iceland, Sæmundargötu 2, Reykjavik 102, Iceland; [22]Department of Clinical Biochemistry, Landspitali - National University Hospital of Iceland, Hringbraut, Reykjavik 101, Iceland; [23]Division of Genome Science, Department of Precision Medicine, National Institute of Health, Chungbuk, South Korea; [24]Laboratory for Statistical Analysis, RIKEN Center for Integrative Medical Sciences, Yokohama, Japan; [25]Department of Ophthalmology, Graduate School of Medical Sciences, Kyushu University, Fukuoka, Japan; [26]Department of Statistical Genetics, Osaka University Graduate School of Medicine, Osaka, Japan; [27]Laboratory for Statistical Analysis, RIKEN Center for Integrative Medical Sciences, The University of Tokyo, Tokyo, Japan; [28]Department of Allergy and Rheumatology, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan; [29]Laboratory for Statistical and Translational Genetics, RIKEN Center for Integrative Medical Sciences, Yokohama, Japan; [30]Department of Biomedical Informatics, Harvard Medical School, Boston, MA, USA; [31]Analytic and Translational Genetics Unit, Massachusetts General Hospital, Boston, MA, USA; [32]K.G. Jebsen Center for Genetic Epidemiology, Department of Public Health and Nursing, NTNU, Norwegian University of Science and Technology, Trondheim, Norway; [33]MRC Integrative Epidemiology Unit (IEU), Bristol Medical School, University of Bristol, Oakfield House, Oakfield Grove, BS8 2BN Bristol, UK; [34]Clinic of Medicine, St. Olavs Hospital, Trondheim University Hospital, Trondheim, Norway; [35]Division of Medicine and Laboratory Sciences,

(Affiliations continued on next page)

© 2022 American Society of Human Genetics.

Jennifer Allen Pacheco,[39] Elisabeth A. Rosenthal,[40] Todd Lingren,[41] QiPing Feng,[42] Iftikhar J. Kullo,[43] Akira Narita,[44] Jun Takayama,[44] Hilary C. Martin,[45] Karen A. Hunt,[46] Bhavi Trivedi,[46] Jeffrey Haessler,[47] Franco Giulianini,[48] Yuki Bradford,[38] Jason E. Miller,[38] Archie Campbell,[49,50] Kuang Lin,[51] Iona Y. Millwood,[51,52] Asif Rasheed,[53] George Hindy,[54] Jessica D. Faul,[55] Wei Zhao,[56] David R. Weir,[55] Constance Turman,[57] Hongyan Huang,[57] Mariaelisa Graff,[58] Ananyo Choudhury,[59] Dhriti Sengupta,[59] Anubha Mahajan,[60] Michael R. Brown,[61] Weihua Zhang,[62,63,64] Ketian Yu,[13] Ellen M. Schmidt,[13] Anita Pandit,[13] Stefan Gustafsson,[65] Xianyong Yin,[13] Jian'an Luan,[66] Jing-Hua Zhao,[67] Fumihiko Matsuda,[68] Hye-Mi Jang,[23] Kyungheon Yoon,[23] Carolina Medina-Gomez,[69] Achilleas Pitsillides,[5] Jouke Jan Hottenga,[70,71] Andrew R. Wood,[72] Yingji Ji,[72] Zishan Gao,[73,74,75] Simon Haworth,[33,76] Ruth E. Mitchell,[33,77] Jin Fang Chai,[78] Mette Aadahl,[79,80] Anne A. Bjerregaard,[79] Jie Yao,[81] Ani Manichaikul,[82] Wen-Jane Lee,[83] Chao Agnes Hsiung,[84] Helen R. Warren,[4,85] Julia Ramirez,[4] Jette Bork-Jensen,[86] Line L. Kårhus,[79] Anuj Goel,[60,87] Maria Sabater-Lleal,[88,89] Raymond Noordam,[90] Pala Mauro,[91] Floris Matteo,[91,92] Aaron F. McDaid,[93,94] Pedro Marques-Vidal,[95] Matthias Wielscher,[62] Stella Trompet,[90,96] Naveed Sattar,[97] Line T. Møllehave,[79] Matthias Munz,[98] Lingyao Zeng,[99,100] Jianfeng Huang,[101] Bin Yang,[101] Alaitz Poveda,[102] Azra Kurbasic,[102] Sebastian Schönherr,[103] Lukas Forer,[103] Markus Scholz,[104,105] Tessel E. Galesloot,[106] Jonathan P. Bradfield,[107] Sanni E. Ruotsalainen,[36] E. Warwick Daw,[108] Joseph M. Zmuda,[109] Jonathan S. Mitchell,[110] Christian Fuchsberger,[110] Henry Christensen,[111] Jennifer A. Brody,[112] Phuong Le,[113,114] Mary F. Feitosa,[108] Mary K. Wojczynski,[108] Daiane Hemerich,[115] Michael Preuss,[115] Massimo Mangino,[116,117] Paraskevi Christofidou,[116]

(Author list continued on next page)

University of Oslo, Oslo, Norway; [36]Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Tukholmankatu 8, 00014 Helsinki, Finland; [37]Department of Public Health and Welfare, Finnish Institute for Health and Welfare, Helsinki, Finland; [38]Department of Genetics, Institute for Biomedical Informatics, University of Pennsylvania, Perelman School of Medicine, Philadelphia, PA 19104, USA; [39]Center for Genetic Medicine, Northwestern University, Feinberg School of Medicine, Chicago, IL 60611, USA; [40]Department of Medicine (Medical Genetics), University of Washington, Seattle, WA, USA; [41]Division of Biomedical Informatics, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA; [42]Division of Clinical Pharmacology, Department of Medicine, Vanderbilt University Medical Center, Nashville, TN, USA; [43]Department of Cardiovascular Medicine and the Gonda Vascular Center, Mayo Clinic, Rochester, MN, USA; [44]Tohoku Medical Megabank Organization, Tohoku University, Sendai 980-8573, Japan; [45]Wellcome Trust Sanger Institute, CB10 1SA Hinxton, UK; [46]Blizard Institute, Barts and the London School of Medicine and Dentistry, Queen Mary University of London, London, UK; [47]Fred Hutchinson Cancer Research Center, Division of Public Health Sciences, Seattle, WA 98109, USA; [48]Division of Preventive Medicine, Brigham and Women's Hospital, Boston, MA 02215, USA; [49]Centre for Genomic and Experimental Medicine, Institute of Genetics and Cancer, University of Edinburgh, Western General Hospital, EH4 2XU Edinburgh, UK; [50]Usher Institute, The University of Edinburgh, Nine, Edinburgh Bioquarter, 9 Little France Road, EH16 4UX Edinburgh, UK; [51]Clinical Trial Service Unit and Epidemiological Studies Unit, Nuffield Department of Population Health, University of Oxford, OX3 7LF Oxford, UK; [52]Medical Research Council Population Health Research Unit, Nuffield Department of Population Health, University of Oxford, OX3 7LF Oxford, UK; [53]Center for Non-Communicable Diseases, Karachi, SD, Pakistan & Faisalabad Institute of Cardiology, Faislabad, Pakistan; [54]Department of Population Medicine, Qatar University College of Medicine, QU Health, Doha, Qatar; [55]Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor, MI 48104, USA; [56]Department of Epidemiology, School of Public Health, University of Michigan, Ann Arbor, MI 48109, USA; [57]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; [58]Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA; [59]Sydney Brenner Institute for Molecular Bioscience, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa; [60]Wellcome Centre for Human Genetics, University of Oxford, Oxford OX3 7BN, UK; [61]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, TX 77030, USA; [62]Department of Epidemiology and Biostatistics, MRC-PHE Centre for Environment and Health, School of Public Health, Imperial College London, W2 1PG London, UK; [63]Department of Cardiology, Ealing Hospital, London North West University Healthcare NHS Trust, UB1 3HW Middlesex, UK; [64]Imperial College Healthcare NHS Trust, Imperial College London, W12 0HS London, UK; [65]Department of Medical Sciences, Molecular Epidemiology and Science for Life Laboratory, Uppsala University, Uppsala, Sweden; [66]MRC Epidemiology Unit, University of Cambridge School of Clinical Medicine, CB2 0QQ Cambridge, UK; [67]Cardiovascular Epidemiology Unit, Department of Public Health and Primary Care, University of Cambridge, Strangeways Research Laboratory, Wort's Causeway, CB1 8RN Cambridge, UK; [68]Center for Genomic Medicine, Kyoto University Graduate School of Medicine, Kyoto, Japan; [69]Department of Internal Medicine, Erasmus MC, University Medical Center Rotterdam, Rotterdam, the Netherlands; [70]Department of Biological Psychology, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands; [71]Amsterdam Public Health Research Institute, Amsterdam UMC, Amsterdam, the Netherlands; [72]Genetics of Complex Traits, University of Exeter Medical School, University of Exeter, EX2 5DW Exeter, UK; [73]Department of Clinical Acupuncture and Moxibustion, Nanjing University of Chinese Medicine, Nanjing, Jiangsu 210029, China; [74]Research Unit of Molecular Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany; [75]Institute of Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany; [76]Bristol Dental School, University of Bristol, Lower Maudlin Street, BS1 2LY Bristol, UK; [77]Population Health Sciences, Bristol Medical School, University of Bristol, Oakfield Grove, BS8 2BN Bristol, UK; [78]Saw Swee Hock School of Public Health, National University of Singapore and National University Health System, Singapore 117549, Singapore; [79]Center for Clinical Research and Prevention, Bispebjerg and Frederiksberg Hospital, Copenhagen, Denmark; [80]Department of Clinical Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark; [81]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; [82]Center for Public Health Genomics, University of Virginia, Charlottesville, VA 22903, USA; [83]Department of Medical Research, Taichung Veterans General Hospital, Taichung, Taiwan; No. 1650, Sec. 4, Taiwan Boulevard, Taichung City 40705, Taiwan; [84]Institute of Population Health Sciences, National Health Research Institutes, 35 Keyan Road, Zhunan Town, Miaoli County 350, Taiwan, ROC; [85]NIHR Barts Cardiovascular Biomedical Research Centre, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, EC1M 6BQ London, UK; [86]Novo Nordisk Foundation Center for Basic Metabolic Research, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark; [87]Division of Cardiovascular Medicine, Radcliffe Department of Medicine, John Radcliffe Hospital, University of Oxford, OX3 9DU Oxford, UK; [88]Unit of Genomics of Complex Diseases. Sant Pau

(Affiliations continued on next page)

Niek Verweij,[118] Jan W. Benjamins,[118] Jorgen Engmann,[119,120] Tsao L. Noah,[121] Anurag Verma,[1] Roderick C. Slieker,[122,123] Ken Sin Lo,[124] Nuno R. Zilhao,[125] Marcus E. Kleber,[126,127] Graciela E. Delgado,[126] Shaofeng Huo,[128] Daisuke D. Ikeda,[129] Hiroyuki Iha,[129] Jian Yang,[130,131] Jun Liu,[132,133] Ayşe Demirkan,[133,134] Hampton L. Leonard,[135,136] Jonathan Marten,[137] Carina Emmel,[138] Börge Schmidt,[138] Laura J. Smyth,[139] Marisa Cañadas-Garre,[139,140,141,142] Chaolong Wang,[143,144] Masahiro Nakatochi,[145] Andrew Wong,[146] Nina Hutri-Kähönen,[147] Xueling Sim,[78] Rui Xia,[148] Alicia Huerta-Chagoya,[149] Juan Carlos Fernandez-Lopez,[150] Valeriya Lyssenko,[102,151] Suraj S. Nongmaithem,[152] Alagu Sankareswaran,[152,153] Marguerite R. Irvin,[146] Christopher Oldmeadow,[154] Han-Na Kim,[155,156] Seungho Ryu,[157,158] Paul R.H.J. Timmers,[137,159] Liubov Arbeeva,[160] Rajkumar Dorajoo,[144,161] Leslie A. Lange,[162] Gauri Prasad,[153,163] Laura Lorés-Motta,[164] Marc Pauper,[164] Jirong Long,[165] Xiaohui Li,[81] Elizabeth Theusch,[166] Fumihiko Takeuchi,[167] Cassandra N. Spracklen,[168,169] Anu Loukola,[36] Sailalitha Bollepalli,[36] Sophie C. Warner,[170,171] Ya Xing Wang,[172] Wen B. Wei,[173] Teresa Nutile,[174] Daniela Ruggiero,[174,175] Yun Ju Sung,[176] Shufeng Chen,[101] Fangchao Liu,[101] Jingyun Yang,[177,178] Katherine A. Kentistou,[159] Bernhard Banas,[179] Anna Morgan,[180] Karina Meidtner,[181,182] Lawrence F. Bielak,[56] Jennifer A. Smith,[55,56] Prashantha Hebbar,[183] Aliki-Eleni Farmaki,[184,185] Edith Hofer,[186,187] Maoxuan Lin,[188] Maria Pina Concas,[180] Simona Vaccargiu,[189] Peter J. van der Most,[190] Niina Pitkänen,[191,192] Brian E. Cade,[193,194] Sander W. van der Laan,[195] Kumaraswamy Naidu Chitrala,[196,197] Stefan Weiss,[198] Amy R. Bentley,[199] Ayo P. Doumatey,[199] Adebowale A. Adeyemo,[199] Jong Young Lee,[200] Eva R.B. Petersen,[201] Aneta A. Nielsen,[202] Hyeok Sun Choi,[203] Maria Nethander,[204,205] Sandra Freitag-Wolf,[206]

(Author list continued on next page)

Biomedical Research Institute (IIB Sant Pau), Barcelona, Spain; [89]Cardiovascular Medicine Unit, Department of Medicine, Karolinska Institutet, Center for Molecular Medicine, Karolinska University Hospital, Stockholm, Sweden; [90]Department of Internal Medicine, Section of Gerontology and Geriatrics, Leiden University Medical Center, Leiden, the Netherlands; [91]Istituto di Ricerca Genetica e Biomedica, Consiglio Nazionale delle Ricerche, Rome, Italy; [92]Dipartimento di Scienze Biomediche, Università degli Studi di Sassari, Sardinia, Italy; [93]University Center for Primary Care and Public Health, University of Lausanne, Rte de Berne 113, 1010 Lausanne, Switzerland; [94]Swiss Institute of Bioinformatics, 1015 Lausanne, Switzerland; [95]Department of Medicine, Internal Medicine, Lausanne University Hospital and University of Lausanne, Rue du Bugnon 46, 1011 Lausanne, Switzerland; [96]Department of Cardiology, Leiden University Medical Center, Leiden, the Netherlands; [97]BHF Glasgow Cardiovascular Research Centre, Faculty of Medicine, Glasgow, UK; [98]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; [99]Deutsches Herzzentrum München, Klinik für Herz- und Kreislauferkrankungen, Technische Universität München, Munich, Germany; [100]Deutsches Zentrum für Herz-Kreislauf-Forschung (DZHK) e.V., partner site Munich Heart Alliance, Munich, Germany; [101]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; [102]Lund University Diabetes Centre, Lunds University, Malmö, Sweden; [103]Institute of Genetic Epidemiology, Department of Genetics and Pharmacology, Medical University of Innsbruck, Innsbruck, Austria and German Chronic Kidney Disease Study, Austria; [104]Institute for Medical Informatics, Statistics and Epidemiology, University of Leipzig, Haertelstrasse 16-18, 04107 Leipzig, Germany; [105]LIFE Research Centre for Civilization Diseases, University of Leipzig, Philipp-Rosenthal-Straße 27, 04103 Leipzig, Germany; [106]Radboud university medical center, Radboud Institute for Health Sciences, Nijmegen, the Netherlands; [107]Quantinuum Research LLC, Wayne, PA 19087, USA; [108]Division of Statistical Genomics, Department of Genetics, Washington University School of Medicine, St. Louis, MO, USA; [109]Department of Epidemiology, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA 15232, USA; [110]Institute for Biomedicine, Eurac Research, Affiliated Institute of the University of Lübeck, Via Galvani 31, 39100 Bolzano, Italy; [111]Department of Clinical Biochemistry, Lillebaelt Hospital, Vejle, Denmark; [112]Cardiovascular Health Research Unit, Department of Medicine, University of Washington, Seattle, WA 98101, USA; [113]Department of Anthropology, University of Toronto at Mississauga, Mississauga, ON L5L 1C6, Canada; [114]Department of Computer Science, University of Toronto, Toronto, ON M5S 2E4, Canada; [115]The Charles Bronfman Institute for Personalized Medicine, Icahn School of Medicine at Mount Sinai, New York, NY, USA; [116]Department of Twin Research and Genetic Epidemiology, King's College London, SE1 7EH London, UK; [117]NIHR Biomedical Research Centre at Guy's and St Thomas' Foundation Trust, SE1 9RT London, UK; [118]Department of Cardiology, University of Groningen, University Medical Center Groningen, 9700RB Groningen, the Netherlands; [119]Institute of Cardiovascular Sciences, University College London, Gower Street, WC1E 6BT London, UK; [120]Department of Epidemiology and Public Health, University College London, 1-19 Torrington Place, WC1E 6BT London, UK; [121]Department of Surgery, University of Pennsylvania, Philadelphia, PA 19104, USA; [122]Amsterdam UMC, Department of Epidemiology and Data Science, Amsterdam Public Health Research Institute, Amsterdam 1081HV, the Netherlands; [123]Leiden University Medical Center, Department of Cell and Chemical Biology, Leiden 2333ZA, the Netherlands; [124]Montreal Heart Institute, Université de Montréal, 5000 Belanger street, Montreal, QC H1T1C8, Canada; [125]Icelandic Heart Association, 201 Kopavogur, Iceland; [126]Vth Department of Medicine, Medical Faculty Mannheim, Heidelberg University, 68167 Mannheim, Germany; [127]SYNLAB MVZ Humangenetik Mannheim GmbH, 68163 Mannheim, Germany; [128]Shanghai Institute of Nutrition and Health, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Shanghai, China; [129]Biomedical Technology Research Center, Tokushima Research Institute, Otsuka Pharmaceutical Co., Ltd., Tokushima, Japan; [130]School of Life Sciences, Westlake University, Hangzhou, Zhejiang 310024, China; [131]Institute for Molecular Bioscience, The University of Queensland, Brisbane, QLD 4072, Australia; [132]Nuffield Department of Population Health, University of Oxford, Oxford, UK; [133]Department of Epidemiology, Erasmus MC, University Medical Center, Rotterdam, the Netherlands; [134]Section of Statistical Multi-omics, Department of Clinical and Experimental research, University of Surrey, Guildford, Surrey, UK; [135]Laboratory of Neurogenetics, National Institute on Aging, NIH, Bethesda, MD, USA; [136]Data Tecnica International, Glen Echo, MD, USA; [137]MRC Human Genetics Unit, Institute of Genetics and Cancer, University of Edinburgh, Western General Hospital, Crewe Road, Edinburgh EH4 2XU, UK; [138]Institute for Medical Informatics, Biometrie and Epidemiology, University of Duisburg-Essen, Essen, Germany; [139]Centre for Public Health, Queen's University of Belfast, Belfast, Northern Ireland; [140]Genomic Oncology Area, GENYO, Centre for Genomics and Oncological Research: Pfizer-University of Granada-Andalusian Regional Government, Granada, Spain; [141]Hematology Department, Hospital Universitario Virgen de las Nieves, Granada, Spain; [142]Instituto de Investigación Biosanitaria de Granada (ibs.GRANADA), Granada, Spain; [143]Department of Epidemiology and Biostatistics, School of Public Health, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China;

(Affiliations continued on next page)

Lorraine Southam,[45,207] Nigel W. Rayner,[45,60,207,208] Carol A. Wang,[209] Shih-Yi Lin,[210,211,212] Jun-Sing Wang,[213,214] Christian Couture,[215] Leo-Pekka Lyytikäinen,[216,217] Kjell Nikus,[218,219] Gabriel Cuellar-Partida,[220] Henrik Vestergaard,[86,221] Bertha Hidalgo,[222] Olga Giannakopoulou,[4] Qiuyin Cai,[165] Morgan O. Obura,[122] Jessica van Setten,[223] Karen Y. He,[224] Hua Tang,[2] Natalie Terzikhan,[133] Jae Hun Shin,[203] Rebecca D. Jackson,[225] Alexander P. Reiner,[226] Lisa Warsinger Martin,[227] Zhengming Chen,[51,52] Liming Li,[228] Takahisa Kawaguchi,[68] Joachim Thiery,[105,229] Joshua C. Bis,[112] Lenore J. Launer,[230] Huaixing Li,[128] Mike A. Nalls,[135,136] Olli T. Raitakari,[191,192,231] Sahoko Ichihara,[232] Sarah H. Wild,[233] Christopher P. Nelson,[170,171] Harry Campbell,[159] Susanne Jäger,[181,182] Toru Nabika,[234] Fahd Al-Mulla,[183] Harri Niinikoski,[235,236] Peter S. Braund,[170,171] Ivana Kolcic,[237] Peter Kovacs,[238] Tota Giardoglou,[184] Tomohiro Katsuya,[239,240] Dominique de Kleijn,[241] Gert J. de Borst,[241] Eung Kweon Kim,[242] Hieab H.H. Adams,[133,243] M. Arfan Ikram,[133] Xiaofeng Zhu,[224] Folkert W. Asselbergs,[223] Adriaan O. Kraaijeveld,[223] Joline W.J. Beulens,[122,244] Xiao-Ou Shu,[165] Loukianos S. Rallidis,[245] Oluf Pedersen,[86] Torben Hansen,[86] Paul Mitchell,[246] Alex W. Hewitt,[247,248] Mika Kähönen,[249,250] Louis Pérusse,[215,251] Claude Bouchard,[252] Anke Tönjes,[238] Yii-Der Ida Chen,[81] Craig E. Pennell,[209] Trevor A. Mori,[253] Wolfgang Lieb,[254] Andre Franke,[255] Claes Ohlsson,[204,256] Dan Mellström,[204,257] Yoon Shin Cho,[203] Hyejin Lee,[258] Jian-Min Yuan,[109,259] Woon-Puay Koh,[260,261] Sang Youl Rhee,[262] Jeong-Taek Woo,[262] Iris M. Heid,[11] Klaus J. Stark,[11] Martina E. Zimmermann,[11] Henry Völzke,[263] Georg Homuth,[198] Michele K. Evans,[230] Alan B. Zonderman,[230] Ozren Polasek,[237,264] Gerard Pasterkamp,[195] Imo E. Hoefer,[195] Susan Redline,[193,194] Katja Pahkala,[191,192,265] Albertine J. Oldehinkel,[266] Harold Snieder,[190]

(Author list continued on next page)

[144]Genome Institute of Singapore, Agency for Science, Technology and Research, Singapore, Singapore; [145]Public Health Informatics Unit, Department of Integrated Health Sciences, Nagoya University Graduate School of Medicine, Nagoya 461-8673, Japan; [146]University of Alabama at Birmingham, Epidemiology, School of Public Health, Birmingham, AL, USA; [147]Tampere Centre for Skills Training and Simulation, Faculty of Medicine and Health Technology, Tampere University, Tampere, Finland; [148]Brown Foundation Institute of Molecular Medicine, McGovern Medical School, University of Texas Health Science Center at Houston, Houston TX 77030, USA; [149]CONACYT, Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán, Ciudad de Mexico, Mexico; [150]Departamento de Genómica Computacional, Instituto Nacional de Medicina Genómica, Ciudad de Mexico, Mexico; [151]Center for diabetes research, University of Bergen, Bergen, Norway; [152]Genomic Research on Complex diseases (GRC Group), CSIR-Centre for Cellular and Molecular Biology, Hyderabad, Telangana, India; [153]Academy of Scientific and Innovative Research (AcSIR), CSIR-Human Resource Development Centre, New Delhi, India; [154]Hunter Medical Research Institute, Newcastle, NSW, Australia; [155]Medical Research Institute, Kangbuk Samsung Hospital, Sungkyunkwan University School of Medicine, Seoul 03181, Korea; [156]Department of Clinical Research Design & Evaluation, SAIHST, Sungkyunkwan University, Seoul 06355, Korea; [157]Center for Cohort Studies, Total Healthcare Center, Kangbuk Samsung Hospital, Sungkyunkwan University School of Medicine, Seoul 04514, Korea; [158]Department of Occupational and Environmental Medicine, Kangbuk Samsung Hospital, Sungkyunkwan University School of Medicine, Seoul 03181, Korea; [159]Centre for Global Health Research, Usher Institute, University of Edinburgh, Teviot Place, Edinburgh EH8 9AG, UK; [160]Thurston Arthritis Research Center, University of North Carolina, Chapel Hill, NC, USA; [161]Health Services and Systems Research, Duke-NUS Medical School, 169857, Singapore; [162]Division of Biomedical Informatics and Personalized Medicine, Department of Medicine, Anschutz Medical Campus, University of Colorado, Denver, Aurora, CO 80045, USA; [163]Genomics and Molecular Medicine Unit, CSIR-Institute of Genomics and Integrative Biology, New Delhi 110020, India; [164]Departments of Ophthalmology and Human Genetics, Radboud University Nijmegen Medical Center, Philips van Leydenlaan 15, Nijmegen 6525 EX, the Netherlands; [165]Vanderbilt Epidemiology Center, Division of Epidemiology, Vanderbilt University Medical Center, Nashville, TN, USA; [166]Department of Pediatrics, University of California San Francisco, Oakland, CA 94609, USA; [167]National Center for Global Health and Medicine, Tokyo 1628655, Japan; [168]Department of Genetics, University of North Carolina, Chapel Hill, NC 27599, USA; [169]Department of Biostatistics and Epidemiology, University of Massachusetts-Amherst, Amherst, MA 01003, USA; [170]Department of Cardiovascular Sciences, University of Leicester, Leicester, UK; [171]NIHR Leicester Biomedical Research Centre, Glenfield Hospital, Leicester, UK; [172]Beijing Institute of Ophthalmology, Beijing Key Laboratory of Ophthalmology and Visual Sciences, Beijing Tongren Eye Center, Beijing Tongren Hospital, Capital Medical University, Beijing Ophthalmology and Visual Sciences Key Laboratory, 17 Hougou Lane, Chong Wen Men, Beijing 100005, China; [173]Beijing Institute of Ophthalmology, Beijing Tongren Eye Center, Beijing Tongren Hospital, Capital Medical University, 1 Dong Jiao Min Xiang, Dong Cheng District, Beijing 100730, China; [174]Institute of Genetics and Biophysics "Adriano Buzzati-Traverso" - CNR, Naples, Italy; [175]IRCCS Neuromed, Pozzilli, Isernia, Italy; [176]Division of Biostatistics, Washington University School of Medicine, St. Louis, MO 63110, USA; [177]Rush Alzheimer's Disease Center, Rush University Medical Center, Chicago, IL, USA; [178]Department of Neurological Sciences, Rush University Medical Center, Chicago, IL, USA; [179]Department of Nephrology, University Hospital Regensburg, Regensburg, Germany; [180]Institute for Maternal and Child Health—IRCCS, Burlo Garofolo, 34127 Trieste, Italy; [181]Department of Molecular Epidemiology, German Institute of Human Nutrition Potsdam-Rehbruecke, Nuthetal, Germany; [182]German Center for Diabetes Research (DZD), München-Neuherberg, Germany; [183]Department of Genetics and Bioinformatics, Dasman Diabetes Institute, Kuwait City, Kuwait; [184]Department of Nutrition and Dietetics, School of Health Science and Education, Harokopio University of Athens, Eleftheriou Venizelou, Athens, Greece; [185]Department of Population Science and Experimental Medicine, University College London, London, UK; [186]Clinical Division of Neurogeriatrics, Department of Neurology, Medical University of Graz, Graz, Austria; [187]Institute for Medical Informatics, Statistics and Documentation, Medical University of Graz, Graz, Austria; [188]Massachusetts General Hospital Cancer Center, Charlestown, MA 02129, USA; [189]Institute of Genetic and Biomedical Research, National Research Council of Italy, UOS of Sassari, Sassari, Italy; [190]Department of Epidemiology, University of Groningen, University Medical Center Groningen, Groningen 9700 RB, the Netherlands; [191]Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, Turku, Finland; [192]Centre for Population Health Research, University of Turku and Turku University Hospital, Turku, Finland; [193]Sleep Medicine and Circadian Disorders, Brigham and Women's Hospital, Boston, MA 02115, USA; [194]Division of Sleep Medicine, Harvard Medical School, Boston, MA 02115, USA; [195]Central Diagnostics Laboratory, Division Laboratories, Pharmacy, and Biomedical genetics, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands; [196]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; [197]Fels Cancer Institute for Personalized Medicine, Temple University Lewis Katz School of Medicine, Philadelphia, PA, USA; [198]Interfaculty Institute for Genetics and Functional Genomics, Department of Functional Genomics, University of Greifswald and University Medicine Greifswald, Greifswald, Germany; [199]Center for Research on Genomics and Global Health, National Human Genome Research Institute, National Institutes of Health, 12 South Drive, Room 4047, Bethesda, MD

(Affiliations continued on next page)

Ginevra Biino,[267] Reinhold Schmidt,[186] Helena Schmidt,[268] Stefania Bandinelli,[269] George Dedoussis,[184] Thangavel Alphonse Thanaraj,[183] Patricia A. Peyser,[56] Norihiro Kato,[167] Matthias B. Schulze,[181,182,270] Giorgia Girotto,[180,271] Carsten A. Böger,[179,272,273] Bettina Jung,[179,272,273] Peter K. Joshi,[159] David A. Bennett,[177,178] Philip L. De Jager,[10,274] Xiangfeng Lu,[101] Vasiliki Mamakou,[275,276] Morris Brown,[15,85] Mark J. Caulfield,[4,85] Patricia B. Munroe,[4,85] Xiuqing Guo,[81] Marina Ciullo,[174,175] Jost B. Jonas,[172,277,278,279] Nilesh J. Samani,[170,171] Jaakko Kaprio,[36] Päivi Pajukanta,[280] Teresa Tusié-Luna,[281,282] Carlos A. Aguilar-Salinas,[283,311] Linda S. Adair,[284,285] Sonny Augustin Bechayda,[286,287] H. Janaka de Silva,[288] Ananda R. Wickremasinghe,[289] Ronald M. Krauss,[290] Jer-Yuarn Wu,[291] Wei Zheng,[165] Anneke I. den Hollander,[164] Dwaipayan Bharadwaj,[153,292] Adolfo Correa,[293] James G. Wilson,[294] Lars Lind,[295] Chew-Kiat Heng,[296] Amanda E. Nelson,[160,297] Yvonne M. Golightly,[58,160,298,299] James F. Wilson,[137,159] Brenda Penninx,[300,301] Hyung-Lae Kim,[302] John Attia,[154,209] Rodney J. Scott,[154,209] D.C. Rao,[176] Donna K. Arnett,[303] Mark Walker,[304] Laura J. Scott,[13] Heikki A. Koistinen,[37,305,306] Giriraj R. Chandak,[152,153,307] Josep M. Mercader,[308,309,310] Clicerio Gonzalez Villalpando,[312] Lorena Orozco,[313] Myriam Fornage,[148,314] E. Shyong Tai,[78,315] Rob M. van Dam,[78,315] Terho Lehtimäki,[216,217] Nish Chaturvedi,[316] Mitsuhiro Yokota,[317] Jianjun Liu,[144] Dermot F. Reilly,[318] Amy Jayne McKnight,[139] Frank Kee,[139] Karl-Heinz Jöckel,[138] Mark I. McCarthy,[60,208] Colin N.A. Palmer,[319] Veronique Vitart,[137] Caroline Hayward,[137] Eleanor Simonsick,[320] Cornelia M. van Duijn,[132,133] Zi-Bing Jin,[173,321] Fan Lu,[321] Haretsugu Hishigaki,[129] Xu Lin,[128] Winfried März,[126,322,323] Vilmundur Gudnason,[125,324] Jean-Claude Tardif,[124] Guillaume Lettre,[124]

(Author list continued on next page)

20892, USA; [200]Oneomics. co. ltd. 2F, Soonchunhyang Mirai Medical Center 173, Buheuyng-ro, Bucheon-si Gyeonggi-do 14585, Korea; [201]Department of Clinical Biochemistry and Immunology, Hospital of Southern Jutland, Kresten Philipsens Vej 15, 6200 Aabenraa, Denmark; [202]Department of Clinical Biochemistry, Lillebaelt Hospital, Kolding, Denmark; [203]Department of Biomedical Science, Hallym University, Chuncheon, Gangwon 24252, Korea; [204]Centre for Bone and Arthritis Research, Department of Internal Medicine and Clinical Nutrition, Institute of Medicine, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden; [205]Bioinformatics Core Facility, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden; [206]Institute of Medical Informatics and Statistics, Kiel University, Kiel, Germany; [207]Institute of Translational Genomics, Helmholtz Zentrum München – German Research Center for Environmental Health, Neuherberg, Germany; [208]Oxford Centre for Diabetes, Endocrinology, and Metabolism, University of Oxford, Oxford, UK; [209]School of Medicine and Public Health, Faculty of Medicine and Health, University of Newcastle, Newcastle, NSW 2308, Australia; [210]Center for Geriatrics and Gerontology, Division of Endocrinology and Metabolism, Department of Internal Medicine, Taichung Veterans General Hospital, Taichung, Taiwan; [211]School of Medicine, National Yang-Ming University, Taipei, Taiwan; [212]School of Medicine, National Defense Medical Center, Taipei, Taiwan; [213]Division of Endocrinology and Metabolism, Department of Internal Medicine, Taichung Veterans General Hospital, Taichung, Taiwan; [214]Department of Medicine, School of Medicine, National Yang-Ming University, Taipei, Taiwan; [215]Department of Kinesiology, Université Laval, Quebec City, QC, Canada; [216]Department of Clinical Chemistry, Fimlab Laboratories, 33520 Tampere, Finland; [217]Department of Clinical Chemistry, Finnish Cardiovascular Research Center - Tampere, Faculty of Medicine and Health Technology, Tampere University, 33014 Tampere, Finland; [218]Department of Cardiology, Heart Center, Tampere University Hospital, 33521 Tampere, Finland; [219]Department of Cardiology, Finnish Cardiovascular Research Center - Tampere, Faculty of Medicine and Health Technology, Tampere University, 33014 Tampere, Finland; [220]University of Queensland Diamantina Institute, Translational Research Institute, Kent St, Woolloongabba, Brisbane, QLD 4102, Australia; [221]Department of Medicine, Bornholms Hospital, Rønne, Denmark; [222]School of Public Health, University of Alabama at Birmingham, Birmingham, AL, USA; [223]Cardiology, Division Heart & Lungs, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands; [224]Department of Population and Quantitative Health Sciences, Case Western Reserve University, Cleveland, OH 44106, USA; [225]Ohio State University, Division of Endricinology, Columbus, OH 43210, USA; [226]University of Washington, Department of Epidemiology, Seattle, WA 98195, USA; [227]George Washington University, School of Medicine and Health Sciences, Washington, DC 20037, USA; [228]Department of Epidemiology, School of Public Health, Peking University Health Science Center, Beijing, China; [229]Institute for Laboratory Medicine, University Hospital Leipzig, Paul-List-Strasse 13/15, 04103 Leipzig, Germany; [230]Laboratory of Epidemiology and Population Sciences, National Institute on Aging Intramural Research Program, NIH, Baltimore, MD 20892-9205, USA; [231]Department of Clinical Physiology and Nuclear Medicine, Turku University Hospital, Turku, Finland; [232]Department of Environmental and Preventive Medicine, Jichi Medical University School of Medicine, Shimotsuke 329-0498, Japan; [233]Centre for Population Health Sciences, Usher Institute, University of Edinburgh, Teviot Place, Edinburgh EH8 9AG, UK; [234]Department of Functional Pathology, Shimane University School of Medicine, Izumo 6938501, Japan; [235]Department of Pediatrics and Adolescent Medicine, Turku University Hospital and University of Turku, Turku, Finland; [236]Department of Physiology, University of Turku, Turku, Finland; [237]Faculty of Medicine, University of Split, Šoltanska 2, 21000 Split, Croatia; [238]Medical Department III – Endocrinology, Nephrology, Rheumatology, University of Leipzig Medical Center, Liebigstr. 21, 04103 Leipzig, Germany; [239]Department of Clinical Gene Therapy, Osaka University Graduate School of Medicine, Suita 5650871, Japan; [240]Department of Geriatric and General Medicine, Osaka University Graduate School of Medicine, Suita 5650871, Japan; [241]Department of Vascular Surgery, Division of Surgical Specialties, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands; [242]Corneal Dystrophy Research Institute, Department of Ophthalmology, Yonsei University College of Medicine, Seoul 03722, Korea; [243]Department of Radiology and Nuclear Medicine, Erasmus MC - University Medical Center Rotterdam, Rotterdam, the Netherlands; [244]Julius Centre for Health Sciences and Primary Care, University Medical Ceslrntre Utrecht, Utrecht 3584CG, the Netherlands; [245]Second Department of Cardiology, Medical School, National and Kapodistrian University of Athens, Attikon University Hospital, Athens, Greece; [246]Center for Vision Research, Department of Ophthalmology and The Westmead Institute, University of Sydney, Hawkesbury Rd, Sydney, NSW 2145, Australia; [247]Menzies Institute for Medical Research, School of Medicine, University of Tasmania, Liverpool St, Hobart, TAS 7000, Australia; [248]Centre for Eye Research Australia, University of Melbourne, Melbourne, VIC 3002, Australia; [249]Department of Clinical Physiology, Tampere University Hospital, 33521 Tampere, Finland; [250]Department of Clinical Physiology, Finnish Cardiovascular Research Center - Tampere, Faculty of Medicine and Health Technology, Tampere University, 33014 Tampere, Finland; [251]Centre Nutrition, santé et société (NUTRISS), Institute of Nutrition and Functional Foods (INAF), Quebec, QC, Canada; [252]Pennington Biomedical Research Center, Baton Rouge, LA 70808, USA; [253]Discipline of Internal Medicine, Medical School, The University of Western Australia, Perth, WA, Australia; [254]Institute of Epidemiology, Kiel University, Kiel, Germany; [255]Institute of Clinical Molecular Biology, Kiel University, Kiel, Germany; [256]Sahlgrenska University Hospital, Department of Drug Treatment, Gothenburg, Sweden; [257]Geriatric Medicine, Institute of Medicine, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden; [258]Department

(Affiliations continued on next page)

Leen M. t Hart,[122,123,325] Petra J.M. Elders,[326] Daniel J. Rader,[327,407] Scott M. Damrauer,[121,328] Meena Kumari,[329] Mika Kivimaki,[120] Pim van der Harst,[118] Tim D. Spector,[116] Ruth J.F. Loos,[86,115,330] Michael A. Province,[108] Esteban J. Parra,[113] Miguel Cruz,[331] Bruce M. Psaty,[112,226,332] Ivan Brandslund,[111,333] Peter P. Pramstaller,[110] Charles N. Rotimi,[334] Kaare Christensen,[335] Samuli Ripatti,[36,336,337] Elisabeth Widén,[36] Hakon Hakonarson,[338,339] Struan F.A. Grant,[327,339,340] Lambertus Kiemeney,[106] Jacqueline de Graaf,[106] Markus Loeffler,[104,105] Florian Kronenberg,[103] Dongfeng Gu,[101,341] Jeanette Erdmann,[98] Heribert Schunkert,[99,100] Paul W. Franks,[102] Allan Linneberg,[79,80] J. Wouter Jukema,[96,342] Amit V. Khera,[10,343,345,346] Minna Männikkö,[346] Marjo-Riitta Jarvelin,[62,347,348] Zoltan Kutalik,[93,94] Cucca Francesco,[349,350] Dennis O. Mook-Kanamori,[351,352] Ko Willems van Dijk,[353,354,355] Hugh Watkins,[87,88] David P. Strachan,[356] Niels Grarup,[86] Peter Sever,[357] Neil Poulter,[358] Wayne Huey-Herng Sheu,[359,360] Jerome I. Rotter,[81] Thomas M. Dantoft,[79] Fredrik Karpe,[361,362] Matt J. Neville,[361,362] Nicholas J. Timpson,[33,77] Ching-Yu Cheng,[363,364] Tien-Yin Wong,[363,364] Chiea Chuen Khor,[144] Hengtong Li,[365] Charumathi Sabanayagam,[363,364] Annette Peters,[75,100,366] Christian Gieger,[74,75,366] Andrew T. Hattersley,[367] Nancy L. Pedersen,[368] Patrik K.E. Magnusson,[368] Dorret I. Boomsma,[70,301] Eco J.C. de Geus,[70,301] L. Adrienne Cupples,[5,369] Joyce B.J. van Meurs,[69,133]

(Author list continued on next page)

of Internal Medicine, EwhaWomans University School of Medicine, Seoul, Korea; [259]Division of Cancer Control and Population Sciences, UPMC Hillman Cancer Center, University of Pittsburgh, Pittsburgh, PA 15232, USA; [260]Healthy Longevity Translational Research Programme, Yong Loo Lin School of Medicine, National University of Singapore, Singapore 117545, Singapore; [261]Singapore Institute for Clinical Sciences, Agency for Science Technology and Research (A*STAR), Singapore 117609, Singapore; [262]Department of Endocrinology and Metabolism, Kyung Hee University School of Medicine, Seoul 02447, Korea; [263]Institute for Community Medicine, University Medicine Greifswald, Greifswald, Germany; [264]Algebra University College, Ilica 242, Zagreb, Croatia; [265]Paavo Nurmi Centre, Sports and Exercise Medicine Unit, Department of Physical Activity and Health, University of Turku, Turku, Finland; [266]Interdisciplinary Center Psychopathology and Emotion Regulation (ICPE), University of Groningen, University Medical Center Groningen, Groningen 9700 RB, the Netherlands; [267]Institute of Molecular Genetics, National Research Council of Italy, Pavia, Italy; [268]Gottfried Schatz Research Center for Cell Signaling, Metabolism and Aging, Medical University of Graz, Graz, Austria; [269]Local Health Unit Toscana Centro, Firenze, Italy; [270]Institute of Nutritional Science, University of Potsdam, Nuthetal, Germany; [271]Department of Medicine, Surgery and Health Sciences, University of Trieste, Strada di Fiume 447, 34149 Trieste, Italy; [272]Department of Nephrology, Diabetology, Rheumatology, Traunstein Hospital, Traunstein, Germany; [273]KfH Kidney Center Traunstein, Traunstein, Germany; [274]Center for Translational and Systems Neuroimmunology, Department of Neurology, Columbia University Medical Center, New York, NY, USA; [275]Medical School, National and Kapodistrian University Athens, 75 M. Assias Street, 115 27 Athens, Greece; [276]Dromokaiteio Psychiatric Hospital, 124 61 Athens, Greece; [277]Department of Ophthalmology, Medical Faculty Mannheim, Heidelberg University, Kutzerufer 1, Mannheim 68167, Germany; [278]Institute of Molecular and Clinical Ophthalmology, Basel, Switzerland; [279]Privatpraxis Prof Jonas und Dr Panda-Jonas, Heidelberg, Germany; [280]Department of Human Genetics, David Geffen School of Medicine at UCLA, University of California, Los Angeles, Los Angeles, CA, USA; [281]Unidad de Biología Molecular y Medicina Genómica, Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán, Mexico City 14080, Mexico; [282]Instituto de Investigaciones Biomédicas, UNAM, Mexico City, Mexico; [283]Departamento de Endocrinología y Metabolismo, Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán, Mexico City 14080, Mexico; [284]Department of Nutrition, Gillings School of Global Public Health, University of North Carolina, Chapel Hill, NC 27599, USA; [285]Carolina Population Center, University of North Carolina, Chapel Hill, NC 27516, USA; [286]USC–Office of Population Studies Foundation, University of San Carlos, Cebu City 6000, Philippines; [287]Department of Anthropology, Sociology, and History, University of San Carlos, Cebu City 6000, Philippines; [288]Department of Medicine, Faculty of Medicine, University of Kelaniya, Ragama 11010, Sri Lanka; [289]Department of Public Health, Faculty of Medicine, University of Kelaniya, Ragama 11010, Sri Lanka; [290]Children's Hospital Oakland Research Institute, Oakland, CA 94609, USA; [291]Institute of Biomedical Sciences, Academia Sinica, Taipei, Taiwan; [292]Systems Genomics Laboratory, School of Biotechnology, Jawaharlal Nehru University, New Delhi 110067, India; [293]Department of Medicine, University of Mississippi Medical Center, Jackson, MS 39216, USA; [294]Department of Physiology and Biophysics, University of Mississippi Medical Center, Jackson, MS 39216, USA; [295]Department of Medical Sciences, Uppsala University, Uppsala, Sweden; [296]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, Singapore; [297]Department of Medicine, University of North Carolina, Chapel Hill, NC, USA; [298]Injury Prevention Research Center, University of North Carolina, Chapel Hill, NC, USA; [299]Division of Physical Therapy, University of North Carolina, Chapel Hill, NC, USA; [300]Department of Psychiatry, Amsterdam UMC, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands; [301]Amsterdam Public Health Research Institute, VU Medical Center Amsterdam, Amsterdam, the Netherlands; [302]Department of Biochemistry, College of Medicine, Ewha Womans University, Seoul 07804, Korea; [303]University of Kentucky, College of Public Health, Lexington, KY, USA; [304]Institute of Cellular Medicine (Diabetes), The Medical School, Newcastle University, Framlington Place, Newcastle upon Tyne NE2 4HH, UK; [305]University of Helsinki and Department of Medicine, Helsinki University Hospital, P.O.Box 340, Haartmaninkatu 4, 00029 Helsinki, Finland; [306]Minerva Foundation Institute for Medical Research, Biomedicum 2U, Tukholmankatu 8, 00290 Helsinki, Finland; [307]JSS Academy of Higher Education and Research, Mysuru, India; [308]Programs in Metabolism and Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, MA, USA; [309]Diabetes Unit and Center for Genomic Medicine, Massachusetts General Hospital, Boston, MA, USA; [310]Harvard Medical School, Boston, MA 02115, USA; [311]Dirección de Nutrición and Unidad de Estudios de Enfermedades Metabólicas, Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán, Mexico City, Mexico; [312]Instituto Nacional de Salud Publica y Centro de Estudios en Diabetes, Cuernavaca, Morelos, Mexico; [313]Instituto Nacional de Medicina Genómica, Mexico City, Mexico; [314]Human Genetics Center, School of Public Health, University of Texas Health Science Center at Houston, Houston, TX 77030, USA; [315]Yong Loo Lin School of Medicine, National University of Singapore and National University Health System, Singapore 119228, Singapore; [316]MRC Unit for Lifelong Health and Ageing at UCL, 1-19 Torrington Place, WC1E 7HB London, UK; [317]Kurume University School of Medicine, Kurume 830-0011, Japan; [318]Genetics, Merck Sharp & Dohme Corp, Kenilworth, NJ 07033, USA; [319]Division of Population Health and Genomics, University of Dundee, Ninwells Hospital and Medical School, Dundee DD1 9SY, UK; [320]Intramural Research Program, National Institute on Aging, 3001 S. Hanover St., Baltimore, MD 21225, USA; [321]The Eye Hospital, School of Ophthalmology & Optometry, Wenzhou Medical University, Wenzhou, Zhejiang 325027, China; [322]Synlab Academy, SYNLAB Holding Deutschland GmbH, Mannheim and Augsburg, Germany; [323]Clinical Institute of Medical and Chemical Laboratory Diagnostics, Medical University of Graz, Graz, Austria; [324]Faculty of Medicine, University of Iceland, 101 Reykjavik, Iceland; [325]Leiden University Medical Center, Department of Biomedical Data Sciences, Section Molecular Epidemiology, 2333ZA Leiden, the Netherlands; [326]Amsterdam UMC, Department of General Practice and

(Affiliations continued on next page)

Arfan Ikram,[133] Mohsen Ghanbari,[133,370] Penny Gordon-Larsen,[284,285] Wei Huang,[371] Young Jin Kim,[23] Yasuharu Tabara,[68] Nicholas J. Wareham,[66] Claudia Langenberg,[66] Eleftheria Zeggini,[45,207,372] Jaakko Tuomilehto,[37,336,373] Johanna Kuusisto,[374] Markku Laakso,[374] Erik Ingelsson,[7,65,375,376] Goncalo Abecasis,[13,377] John C. Chambers,[62,63,64,378] Jaspal S. Kooner,[63,64,357,379] Paul S. de Vries,[61] Alanna C. Morrison,[61] Scott Hazelhurst,[59,380] Michèle Ramsay,[59] Kari E. North,[58] Martha Daviglus,[381] Peter Kraft,[57,382] Nicholas G. Martin,[383] John B. Whitfield,[383] Shahid Abbas,[53] Danish Saleheen,[53,384,385] Robin G. Walters,[51,52,386] Michael V. Holmes,[51,52,362] Corri Black,[387] Blair H. Smith,[319] Aris Baras,[377] Anne E. Justice,[388] Julie E. Buring,[48,310] Paul M. Ridker,[48,310] Daniel I. Chasman,[48,310] Charles Kooperberg,[47] Gen Tamiya,[44] Masayuki Yamamoto,[44] David A. van Heel,[46] Richard C. Trembath,[389] Wei-Qi Wei,[390] Gail P. Jarvik,[391] Bahram Namjou,[392] M. Geoffrey Hayes,[39,393,394] Marylyn D. Ritchie,[38] Pekka Jousilahti,[37] Veikko Salomaa,[37] Kristian Hveem,[32,395,396] Bjørn Olav Åsvold,[32,395,397] Michiaki Kubo,[398] Yoichiro Kamatani,[24,399] Yukinori Okada,[24,26,400,401] Yoshinori Murakami,[402] Bong-Jo Kim,[403] Unnur Thorsteinsdottir,[20,324] Kari Stefansson,[20,324] Jifeng Zhang,[3] Y. Eugene Chen,[3] Yuk-Lam Ho,[404] Julie A. Lynch,[405,406] Philip S. Tsao,[6,7,408] Kyong-Mi Chang,[328,407] Kelly Cho,[404,409] Christopher J. O'Donnell,[404,409] John M. Gaziano,[404,409] Peter Wilson,[17,410] Karen L. Mohlke,[168]

(Author list continued on next page)

Elderly Care, Amsterdam Public Health Research Institute, 1081HV Amsterdam, the Netherlands; [327]Department of Genetics, University of Pennsylvania, Philadelphia, PA 19104, USA; [328]Corporal Michael Crescenz VA Medical Center, Philadelphia, PA 19104, USA; [329]Institute of Social and Economic Research, University of Essex, Wivenhoe Park CO4 3SQ, UK; [330]Department of Environmental Medicine and Public Health, Icahn School of Medicine at Mount Sinai, New York, NY, USA; [331]Unidad de Investigacion Medica en Bioquimica, Hospital de Especialidades, Centro Medico Nacional Siglo XXI, Instituto Mexicano del Seguro Social, Mexico City, Mexico; [332]Department of Health Services, University of Washington, Seattle, WA, USA; [333]Institute of Regional Health Research, University of Southern Denmark, Odense, Denmark; [334]16Center for Research on Genomics and Global Health, National Human Genome Research Institute, National Institutes of Health, 12 South Drive, Room 4047, Bethesda, MD 20892, USA; [335]Danish Aging Research Center, University of Southern Denmark, Odense C, Denmark; [336]Department of Public Health, Faculty of Medicine, University of Helsinki, Helsinki, Finland; [337]Broad Institute of MIT and Harvard, Cambridge, MA, USA; [338]Center for Applied Genomics, Children's Hospital of Philadelphia, Philadelphia, PA 19104, USA; [339]Department of Pediatrics, The University of Pennsylvania Perelman School of Medicine, Philadelphia, PA 19104, USA; [340]Division of Human Genetics, Children's Hospital of Philadelphia, Philadelphia, PA 19104, USA; [341]School of Medicine, Southern University of Science and Technology, Shenzhen, China; [342]Netherlands Heart Institute, Utrecht, the Netherlands; [343]Division of Cardiology, Department of Medicine, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA; [344]Center for Genomic Medicine, Department of Medicine, Massachusetts General Hospital, Boston, MA, USA; [345]Department of Medicine, Harvard Medical School, Boston, MA, USA; [346]Northern Finland Birth Cohorts, Infrastructure for population studies, Faculty of Medicine, University of Oulu, Oulu, Finland; [347]Center for Life Course Health Research, Faculty of Medicine, University of Oulu, Oulu, Finland; [348]Biocenter of Oulu, University of Oulu, Oulu, Finland; [349]Institute for Genetic and Biomedical Research, Italian National Council of Research (IRGB CNR), Cagliari, Italy; [350]University of Sassari, Sassari, Italy; [351]Department of Clinical Epidemiology, Leiden University Medical Center, Leiden, the Netherlands; [352]Department of Public Health and Primary Care, Leiden University Medical Center, Leiden, the Netherlands; [353]Department of Internal Medicine, Division of Endocrinology, Leiden University Medical Center, Leiden, the Netherlands; [354]Einthoven Laboratory for Experimental Vascular Medicine, Leiden University Medical Center, Leiden, the Netherlands; [355]Department of Human Genetics, Leiden University Medical Center, Leiden, the Netherlands; [356]Population Health Research Institute, St George's, University of London, SW17 0RE London, UK; [357]National Heart and Lung Institute, Imperial College London, W2 1PG London, UK; [358]School of Public Health, Imperial College London, W12 7RH London, UK; [359]Taichung Veterans General Hospital, Taichung, Taiwan; No. 1650, Sec. 4, Taiwan Boulevard, Xitun District Taichung City 40705, Taiwan; [360]Division of Endocrinology and Metabolism, Department of Medicine, Taipei Veterans General Hospital, Taipei, Taiwan; No. 201, Sec. 2, Shipai Road, Beitou District, Taipei City 112201, Taiwan; [361]OCDEM, University of Oxford, Churchill Hospital, OX3 7LE Oxford, UK; [362]NIHR Oxford Biomedical Research Centre, Churchill Hospital, Oxford, UK; [363]Ocular Epidemiology, Singapore Eye Research Institute, Singapore National Eye Centre, Singapore 168751, Singapore; [364]Ophthalmology & Visual Sciences Academic Clinical Program (Eye ACP), Duke-NUS Medical School, Singapore 169857, Singapore; [365]Data Science, Singapore Eye Research Institute, Singapore National Eye Centre, Singapore 168751, Singapore; [366]German Center for Diabetes Research (DZD), Neuherberg, Germany; [367]University of Exeter Medical School, University of Exeter, EX2 5DW Exeter, UK; [368]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden; [369]Framingham Heart Study, National Heart, Lung, and Blood Institute, US National Institutes of Health, Bethesda, MD, USA; [370]Department of Genetics, School of Medicine, Mashhad University of Medical Sciences, Mashhad, Iran; [371]Department of Genetics, Shanghai-MOST Key Laboratory of Health and Disease Genomics, Chinese National Human Genome Center at Shanghai, Shanghai 201203, China; [372]Technical University of Munich (TUM) and Klinikum Rechts der Isar, TUM School of Medicine, Munich, Germany; [373]Diabetes Research Group, King Abdulaziz University, Jeddah, Saudi Arabia; [374]Institute of Clinical Medicine, Internal Medicine, University of Eastern Finland and Kuopio University Hospital, Kuopio, Finland; [375]Stanford Cardiovascular Institute, Stanford University, Stanford, CA 94305, USA; [376]Stanford Diabetes Research Center, Stanford University, Stanford, CA 94305, USA; [377]Regeneron Pharmaceuticals, Tarrytown, NY, USA; [378]Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore 308232, Singapore; [379]MRC-PHE Centre for Environment and Health, Imperial College London, London W2 1PG, UK; [380]School of Electrical & Information Engineering, University of the Witwatersrand, Witwatersrand, South Africa; [381]Institute for Minority Health Research, University of Illinois College of Medicine, Chicago, IL, USA; [382]Department of Biostatistics, Harvard T.H. Chan School of Public Health, 677 Huntington Avenue, Boston, MA 02115, USA; [383]QIMR Berghofer Medical Research Institute, 300 Herston Road, Brisbane, QLD 4006, Australia; [384]Department of Medicine, Columbia University Irving Medical Center, New York, NY, USA; [385]Department of Cardiology, Columbia University Irving Medical Center, New York, NY, USA; [386]Big Data Instutute, University of Oxford, OX3 7LF Oxford, UK; [387]Aberdeen Centre for Health Data Science,1:042 Polwarth Building School of Medicine, Medical Science and Nutrition University of Aberdeen, Foresterhill, Aberdeen AB25 2ZD, UK; [388]Biomedical and Translational Informatics, Geisinger Health, Danville, PA 17822, USA; [389]School of Basic and Medical Biosciences, Faculty of Life Sciences and Medicine, King's College London, London, UK; [390]Department of Biomedical Informatics, Vanderbilt University Medical Center, Nashville, TN, USA; [391]Departments of Medicine (Medical Genetics) and Genome Sciences, University of Washington, Seattle, WA, USA; [392]Center for Autoimmune Genomics and Etiology, Cincinnati Children's Hospital Medical Center (CCHMC), Cincinnati, OH, USA; [393]Division of Endocrinology, Metabolism, and Molecular Medicine, Department of Medicine, Northwestern University, Feinberg School of Medicine, Chicago, IL 60611, USA; [394]Department of Anthropology, Northwestern University, Evanston, IL 60208, USA; [395]HUNT Research Centre, Department of Public Health and Nursing, NTNU, Norwegian University of Science and Technology, Levanger 7600, Norway; [396]Department of Medicine, Levanger

(Affiliations continued on next page)

Timothy M. Frayling,[72] Joel N. Hirschhorn,[9,10,411] Sekar Kathiresan,[10,344,345] Michael Boehnke,[13] Million Veterans Program, Global Lipids Genetics Consortium, Struan Grant,[1,340] Pradeep Natarajan,[10,343,412,413] Yan V. Sun,[16,17] Andrew P. Morris,[414] Panos Deloukas,[4,415] Gina Peloso,[5] Themistocles L. Assimes,[6,7,408] Cristen J. Willer,[3,14,416] Xiang Zhu,[6,417,418,419,*] and Christopher D. Brown[1,*]

Linking trait-associated variants to genome function has emerged as a promising model for mechanistic interpretation of noncoding findings in GWAS. This "variant-to-function" model is premised on recent observations that noncoding 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 GWASs has become more feasible as large-scale functional genomic resources, such as epigenomic[7] and transcriptomic[8] catalogs, have been systematically generated across a wide range of human cell types and tissues. The integration of functional genomics with GWASs 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 GWASs. 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 been confidently assigned at only a small fraction of these loci,[2] and the regulatory mechanism connecting variant to phenotype has been conclusively characterized for only a handful of genes.[5] Furthermore, systematic mapping of lipid-associated variants to their biological functions has been missing in the literature at the time of this study.

Here we conduct a genome-scale integrative analysis on the largest published GWAS to date of five lipid phenotypes (LDL, or low-density lipoprotein; HDL, or high-density lipoprotein; TC, or total cholesterol; nonHDL, or non-high density lipoprotein; and TG, or triglycerides) involving 1.65 million individuals from five ancestries.[12] Combining the lipid GWAS with a wide array of functional genomic resources in diverse human tissues and cell types, we identify regulatory mechanisms of noncoding genetic variation in lipids with a full suite of computational approaches. Further, we develop a generalizable framework to understand how tissue-specific gene regulation can explain GWAS findings and we demonstrate its real-world value on lipid-associated loci.

Material and methods

GWAS

We used the recently published GWAS data from the Global Lipids Genetics Consortium (GLGC) for five blood lipid traits (LDL, HDL, TC, TG, and nonHDL) in 1.65 million individuals from five ancestry groups[12] (African and African-admixed, East Asian, European, Hispanic, South Asian) at 91 million variants imputed primarily from the Haplotype Reference Consortium[13] or 1,000 Genomes Phase 3.[14] GWASs of individual cohorts were based on the hg19 version of the human reference genome. MR-MEGA[15] was used for meta-analysis across cohorts.

We defined "sentinel variants" as the most significant variant at independent trait-associated loci in the genome. The windows are the greater of 500 kb or 0.25 cM 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 of (or dependent on) a neighboring locus to enforce independence of associated loci.

For each sentinel variant, we defined credible sets of potentially causal variants within ± 500 kb region around the sentinel variant representing the set of variants harboring the causal variant with

Hospital, Nord-Trøndelag Hospital Trust, Levanger 7600, Norway; [397]Department of Endocrinology, St. Olavs Hospital, Trondheim University Hospital, Trondheim, Norway; [398]RIKEN Center for Integrative Medical Sciences, Yokohama, Japan; [399]Laboratory of Complex Trait Genomics, Department of Computational Biology and Medical Sciences, Graduate School of Frontier Sciences, The University of Tokyo, Tokyo, Japan; [400]Laboratory of Statistical Immunology, WPI Immunology Frontier Research Center, Osaka University, Osaka, Japan; [401]Integrated Frontier Research for Medical Science Division, Institute for Open and Transdisciplinary Research Initiatives, Osaka University, Osaka, Japan; [402]Division of Molecular Pathology, Institute of Medical Science, The University of Tokyo, Tokyo, Japan; [403]Division of Genome Research, Center for Genome Science, National Institute of Health, Chungcheongbuk-do, South Korea; [404]VA Boston Healthcare System, Boston, MA, USA; [405]VA Informatics and Computing Infrastructure, VA Salt Lake City Health Care System, Salt Lake City, UT, USA; [406]University of Massachusetts, Boston, MA, USA; [407]Department of Medicine, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, USA; [408]Cardiovascular Institute, Stanford University School of Medicine, Stanford, CA, USA; [409]Department of Medicine, Brigham Women's Hospital, Boston, MA, USA; [410]Division of Cardiology, Emory University School of Medicine, Atlanta, GA, USA; [411]Departments of Pediatrics and Genetics, Harvard Medical School, Boston, MA, USA; [412]Department of Medicine, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA; [413]Cardiovascular Research Center and Center for Genomic Medicine, Massachusetts General Hospital, Boston, MA, USA; [414]Centre for Genetics and Genomics Versus Arthritis, Centre for Musculoskeletal Research, Division of Musculoskeletal and Dermatological Sciences, The University of Manchester, Manchester, UK; [415]Princess Al-Jawhara Al-Brahim Centre of Excellence in Research of Hereditary Disorders (PACER-HD), King Abdulaziz University, Jeddah, Saudi Arabia; [416]Department of Human Genetics, University of Michigan, Ann Arbor, MI 48109, USA; [417]Department of Statistics, The Pennsylvania State University, University Park, PA, USA; [418]Huck Institutes of the Life Sciences, The Pennsylvania State University, University Park, PA, USA; [419]Department of Statistics, Stanford University, Stanford, CA, USA
[420]These authors contributed equally
*Correspondence: xiangzhu@psu.edu (X.Z.), chrbro@upenn.edu (C.D.B.)
https://doi.org/10.1016/j.ajhg.2022.06.012.

a 95% posterior probability. Full details of the credible set construction are reported in our recent GWAS publication.[12] The credible sets are freely available (web resources).

Colocalization of GWAS associations with eQTLs

We performed statistical colocalization of lipid GWASs 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 1 Mb 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 (R v.3.4.3, coloc v.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 500 kb or 0.25 cM on either side of the lead variant) that are present in both datasets. Because eQTL summary statistics were in GRCh38, we lifted over the GWAS summary statistics from 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 types (web resources) to capture physical interactions between gene promoters and their regulatory elements. The four Capture-C datasets are (1) three biological replicates of HepG2 liver carcinoma cells (HepG2.1),[20] (2) another HepG2 dataset described in Selvarajan et al. (HepG2.2),[21] (3) 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 (4) an adipose dataset obtained from Pan et al.[23] that was produced using primary human white adipocytes. Across the four datasets, the number of significant interactions on the same chromosome ranges from 67,819 (adipose) to 126,565 (HLC). The bait end has a median size of 2,141 (HepG2.1) to 6,567 (HepG2.2) bases. The interacting end has a median size of 2,100 (HepG2.1) to 3,243 base pairs (HepG2.2) for all datasets. The median distance between the bait and interacting ends for all interactions on the same chromosome ranges from 71,722 (HLC) to 285,140 base pairs (adipose).

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, totalling 36,691 RNA baited fragments. Each library was then sequenced on an Illumina HiSeq 4,000 (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 [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 v.19 [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 v.19 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 to 2.96 for bait distances from 0 to 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 v.1.30.3[27] to find overlap between credible sets for each trait's GWAS and the previously annotated promoter Capture-C data. Given the bait end of a gene, we defined a GWAS locus as interacting with this gene if a variant in the credible set for this GWAS locus fell inside the interacting end.

Presence of gene-variant pairs in same topologically associated domains

To assess the frequency of colocalized gene-sentinel variant pairs in the same topologically associated domain (TAD), we used a list of 2,499 publicly available TADs from human liver[28] (web resources). We computed as a fraction 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 whether this fraction was statistically significant, we generated random TAD boundaries (using bedtools shuffle) 1,000 times and calculated the same fraction 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 enriched pathways in the latest version of the KEGG database.[30] We first re-mapped Gencode IDs to gene symbols using the Gencode v.24 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 that were 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 four known sets of lipid-associated genes using the Fisher exact test (R function fisher.test). First, we identified 33 Mendelian genes from ClinVar[32] with lipidemia-associated ICD10 codes (E78). Second, we used 35 genes with rare-coding variants associated with lipid levels.[33] Third, we extracted 1,115 genes associated with "cholesterol" or "lipidemia" phenotypes in mouse knockouts from the Mouse Genome Informatics (MGI) database.[34] Fourth, we identified 4,008 genes from a transcriptome-wide association study (TWAS) on the same GWAS and GTEx v8 summary

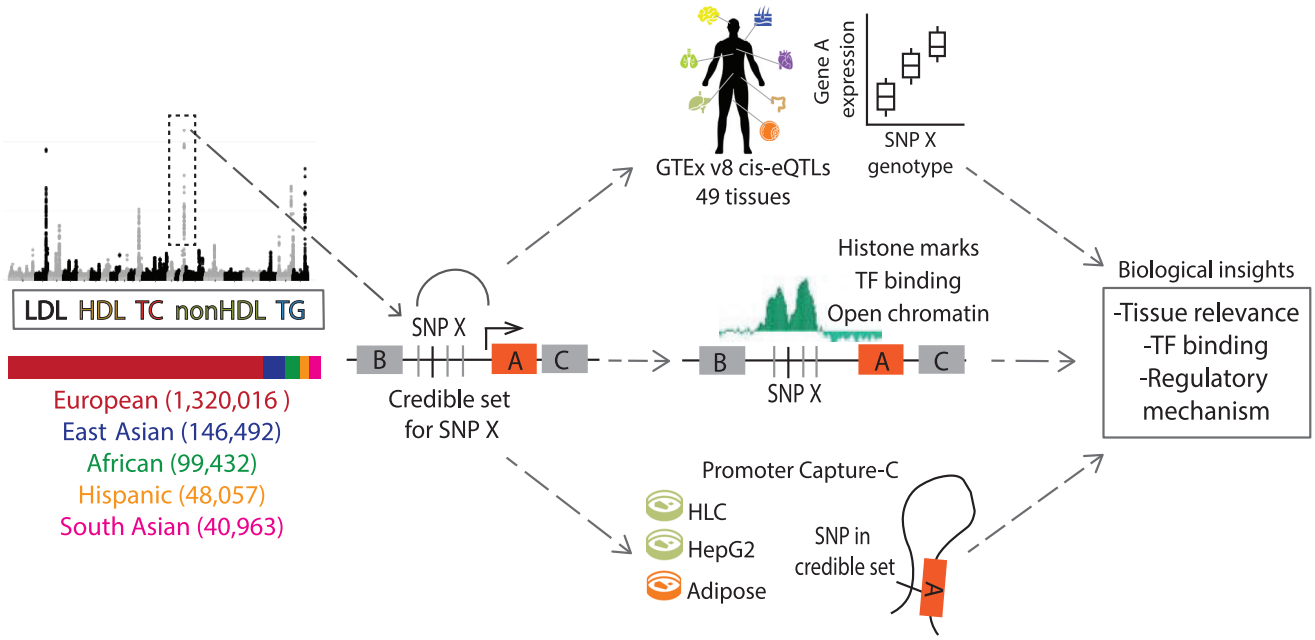

Figure 1. Schematic overview of the multi-layer functional genomic analysis
We integrate GWAS summary statistics for five lipid phenotypes with eQTL and chromatin interaction data to identify potential genes mediating the GWAS loci, and use epigenomic annotations to identify regulatory mechanisms at these loci. For a GWAS locus indexed by a lead variant X, A, B, and C represent nearby eGenes across tissues, and SNPs around SNP X represent variants in the credible set for this locus.

statistics using the S-PrediXcan software[35] default setup. The TWAS method accounts for allelic heterogeneity and thus complements the eQTL colocalization approach that assumes one causal variant per locus.

TF binding sites

We extracted TF binding sites from ChIP-seq data of 161 TFs in 91 cell types from the ENCODE project[7] (web resources). We included all cell types in our primary analysis because TFs were not comprehensively assayed in most cell lines. We also performed a secondary analysis using TF binding sites from HepG2 only. All TF binding sites were aligned to the hg19 version of human reference genome (https://www.encodeproject.org/chip-seq/transcription_factor/).

Stratified LD score (S-LDSC) regression analysis

We used LDSC version 1.0.1[36] to estimate the enrichment of heritability explained using GWAS summary statistics in different epigenetic and transcriptomic annotations, including gene expression, chromatin marks, and TF binding sites. The gene expression and chromatin mark annotations across 205 datasets from more than 170 tissues and cell types and the corresponding LD scores were provided as Multitissuegeneexpr1000Gv3 and Multitissuechromatin1000Gv3 databases in LDSC software (web resources). The LD scores for binding sites of each TF were estimated from 1,000 Genomes Phase 3 European samples using ldsc.py –l2. We first converted the summary statistics for each phenotype to LDSC-formatted summary statistics using munge_sumstats.py. Second, we ran ldsc.py using the baseline_v1.2 model on each annotation to estimate enrichment of heritability. For primary analyses, we used multi-ancestry GWAS summary statistics and LD scores estimated from 1,000 Genomes Phase 3 European samples. For secondary analyses on East Asian (EAS) GWAS alone, we obtained EAS-specific LD scores for the same functional annotations.[37]

Genomic regulatory elements and GWAS overlap algorithm (GREGOR) analysis

We used GREGOR[38] to estimate enrichment of sentinel variants for each lipid phenotype in TF binding sites for 161 TFs from ENCODE compared to a null distribution of variants matched for allele frequency. We ran GREGOR with default parameters, specifying 0.8 as the R^2 threshold, window size of 1 Mb, and 'EUR' as the population. Annotations with enrichment >2 and FDR-adjusted p value < 0.05 were considered significant.

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[39] 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 quantified the overrepresentation of our gene list in ranked genes for this cluster via an enrichment p value computed by the fgsea[40] R package v.1.4.1 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

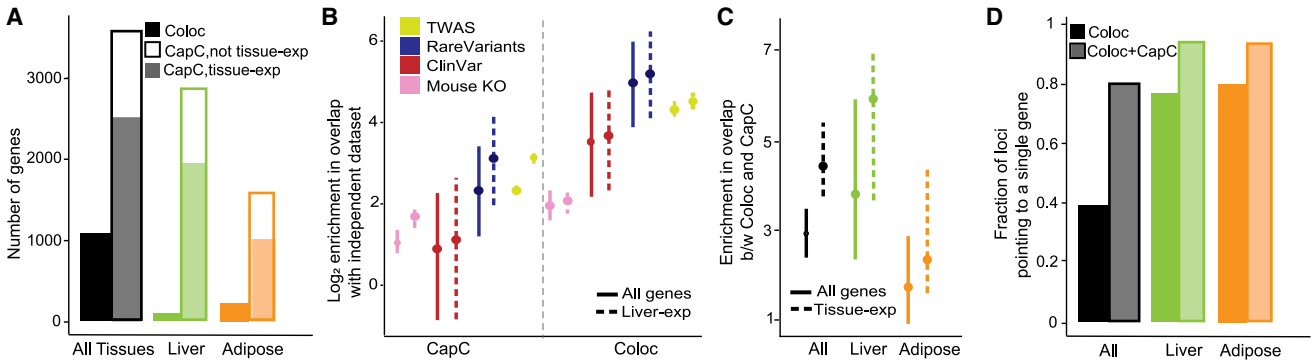

Figure 2. Overlap between eQTL colocalized genes and Capture-C prioritized genes, and their enrichments in known lipid-associated genes

(A) Numbers of genes identified by two approaches: eQTL colocalization (Coloc) and promoter Capture-C interaction (CapC). Capture-C interactions restricted to genes expressed in the tissue of interest (or in the union of adipose and liver for "all tissues") are shaded.

(B) Overlap between two list of prioritized genes (left: Capture-C prioritized genes; right: eQTL colocalized genes) with four external sets of genes previously associated with lipid biology (MGI knockout genes, ClinVar lipidemia-associated genes, genes implicated in rare burden of lipids, and genes from a lipid TWAS). Dashed lines represent enrichments using only genes expressed in the liver.

(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.

Dashed lines represent genes expressed in the tissue of interest (or in the union of adipose or liver for "all"). Enrichment estimates and 95% confidence intervals shown in (B) and (C) are based on the Fisher exact test.

(D) Fraction of colocalized loci that point to a single candidate gene when using eQTL data alone or using both eQTL and Capture-C data.

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.

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.[41] To do so, we tested for colocalization of each significant lipid GWAS signal with significant *cis*-eQTL data across 49 human tissues from the GTEx consortium.[8] The significant GWAS signals were 1,750 loci reaching genome-wide significance and corrected for shadow signals in our multi-ancestry meta-analysis for at least one of five lipid traits. 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.

Second, we restricted our analysis to loci most likely mediated through regulatory mechanisms as opposed to coding variation. Specifically, 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]) 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 and 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,[42] we performed all downstream analyses with all 1,076 colocalized genes, to further prioritize functional genes at loci with multiple eGenes.

Since lipid-associated genetic variants are often enriched in the liver and adipose,[43,44] we repeated the colocalization analysis on eQTLs only from liver or adipose. Compared to the 1,076 colocalized eGenes identified from all 49 tissues, the liver- and adipose-only analysis identified 119 and 225, respectively (Figure 2A). The reduced discovery of colocalized eGenes in the liver- and adipose-only analysis is likely due to the small sample sizes of liver (n = 208) and adipose (n = 581) in GTEx v8 (Figure S1). Leveraging the large degree of tissue sharing in eQTLs,[19,45] our cross-tissue colocalization analysis enhanced the discovery power through the collectively large sample size across all 49 tissues (n = 15,201). For example, several well-documented lipid-relevant genes such as *PPARA*[46] and *LPL*[47] were not identified in the liver- or adipose-only analysis but were identified as significant in our cross-tissue analysis.

To acquire additional functional insights into the 1,076 colocalized genes, we assessed their enrichments across existing biological and clinical gene sets (Figure 2B, Tables S2 and S3). Colocalized genes showed enrichments in (1) 20 KEGG pathways[30] at FDR 5%, including known lipid-related processes such as cholesterol metabolism, PPAR signaling, and bile secretion; (2) 33 Mendelian genes from ClinVar[32] associated with lipid-related ICD10 codes (11.61-fold enrichment, $p = 2.08 \times 10^{-6}$, including *APOB, LPL*, and *APOE*), suggesting the shared genetic basis of Mendelian and complex lipid phenotypes;[48] (3) 35

genes with rare-variant burden for lipid phenotypes in a recent multi-ancestry analysis[33] (30.82-fold enrichment, $p = 1.77 \times 10^{-16}$, including *APOB, LPL, LIPG*, and *ANGPTL4*), confirming shared mechanisms of rare and common variation underlying lipid traits;[49] (4) genes implicated by cholesterol or lipidemia phenotypes in mouse knockouts (3.92-fold enrichment, p = 2.18e−20), suggesting the shared genetic basis of lipid traits between human and mouse.[50] Colocalized genes also showed enrichment with genes implicated in TWAS (Table S4) run on the same GWAS and eQTL summary statistics (20.14-fold enrichment, p < 2.22e−308). These enrichment results demonstrate the biological relevance of candidate functional genes prioritized by our approach.

Chromatin-chromatin interactions shortlist 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.[51] To add this layer to our analysis, we generated Capture-C data from HepG2 liver carcinoma cells (HepG2.1) and hepatocyte-like cells (HLC) derived from differentiating iPSCs,[22] as well as publicly available Capture-C datasets from HepG2[21] (HepG2.2) and white adipocytes.[23] Based on the Capture-C data, we defined an interaction between a GWAS locus and a gene as a significant interaction between the bait end (promoter) for this gene and the interacting end that contains a variant in the credible set for this GWAS locus. In total, 1,079 of 1,750 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 and Table S5).

Unlike eQTL-colocalized genes, genes interacting with GWAS credible sets were not significantly enriched in lipid-relevant KEGG pathways (Table S2) and lipid-related genes from ClinVar (Figure 2B and Table S3). These genes were significantly enriched in genes with rare-variant lipid associations (5.36-fold enrichment, $p = 2.8 \times 10^{-5}$), genes with lipid-related mouse knockouts (1.43-fold enrichment, $p = 2.8 \times 10^{-4}$), and TWAS-prioritized genes (5.05-fold enrichment, $p = 2.5 \times 10^{-288}$), but their enrichments were consistently lower than enrichments of eQTL-colocalized genes nonetheless (Figure 2B and Table S3).

Since genes expressed in the liver are most likely to harbor genuine lipid-relevant variant-gene interactions, we repeated the enrichment analyses above restricting both eQTL colocalization and Capture-C interactions to genes expressed in the liver (>0.1 TPM and ≥6 reads in at least 20% of GTEx liver samples). Reassuringly, we observed higher enrichments for each combination of two methods (eQTL, Capture-C) and four databases (ClinVar, Rare Variant, Mouse Knockout, TWAS) when we restricted our analyses to genes expressed in the liver (Figure 2B and Table S3). For the same database, we observed higher enrichments in eQTL colocalized genes than Capture-C prioritized genes, consistent with the results based on all genes.

Genes physically interacting with GWAS loci significantly overlapped with eQTL colocalized genes despite their reduced enrichments in lipid-related gene sets. Of 1,079 credible sets with promoter interactions, 224 also colocalized with eQTLs for the same gene. Across 49 eQTL tissues and four Capture-C cell lines, 233 genes were implicated in both eQTL colocalizations and Capture-C interactions (Table S6), representing an enrichment of 3-fold compared to random chance (Figure 2C, $p = 3.11 \times 10^{-38}$). Because our Capture-C data came from liver and adipose only, we observed a stronger enrichment in overlap when restricting genes expressed in the liver or adipose (4.5-fold enrichment, $p = 2.89 \times 10^{-65}$). We observed similar enrichment patterns when analyzing liver and adipose Capture-C data separately (Figure 2C). Together, the enrichments in overlap suggest that, despite a large number of genes identified by Capture-C (Figure 2A), many of them are likely to harbor functional interactions with GWAS loci.

Chromatin-chromatin interactions helped shortlist functional genes from eQTL colocalization. Among 224 loci with concordant eQTL colocalizations and Capture-C interactions across all tissues, only 39% (88) mapped to a single gene using eQTL data alone, whereas adding Capture-C information increased this fraction to 80% (180). We observed the same trend in the adipose-only and liver-only analysis: 80% (12/15) and 79% (26/33) of loci mapped to a single gene using adipose and liver eQTLs alone, compared to 93% (14/15) and 97% (32/33) after the integration of adipose-only and liver-only Capture-C data, respectively (Figure 2D). 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 (TAD) as the genes they regulate,[52] we examined TADs from an independent human liver dataset[28] at lipid GWAS loci with eQTL colocalizations to confirm GWAS variant-target gene colocalization within the same TAD. Of eQTL-GWAS colocalizations in which the sentinel variant resided within a TAD, 84.8% (1,040 out of 1,235) had the colocalized gene residing in the same TAD ($p < 0.001$ with 1,000 permutations). When we restricted to all colocalizations concordant with Capture-C data in any cell type, 96.9% (252 out of 260) of gene-variant pairs fell in the same TAD. This fraction further increased to 100% (33 out of 33) when we repeated the analysis using liver eQTLs and liver Capture-C interactions only. These results add to the existing evidence for TAD boundaries being regulatory insulators in the cell[53] 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

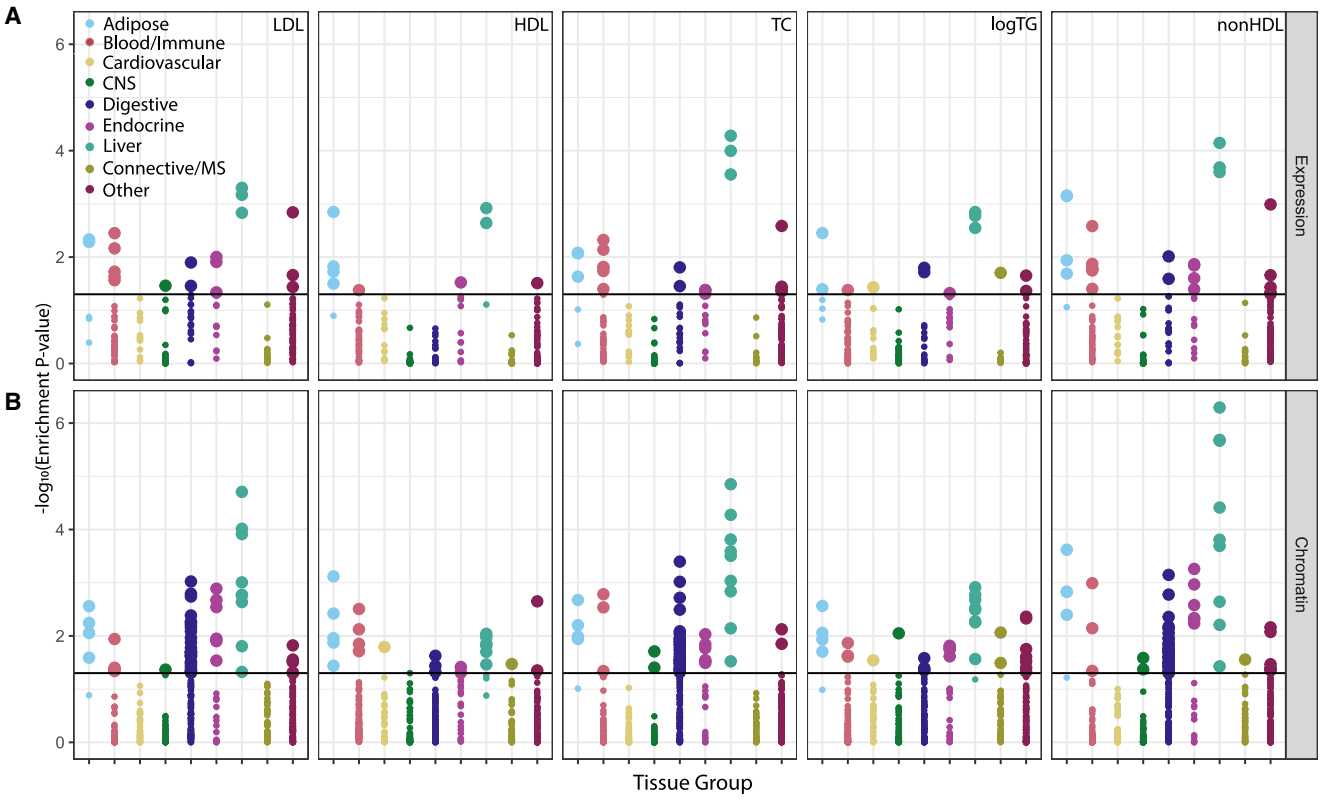

Figure 3. Tissue relevance of lipid-associated loci
Partitioning heritability of lipid GWAS summary statistics on gene expression (A) and active chromatin marks (B) across tissues. Each plotted point represents a tested dataset for enrichment of heritability, with larger dots representing datasets with enrichment p value < 0.05. Each color represents a tissue group (Table S7), and the y-axis represents $-\log10$ p value for enrichment of heritability.

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),[36] a polygenic approach not restricted to genome-wide significant variants, on tissue-specific transcriptomic and epigenomic annotations across 205 datasets from more than 170 tissues and cell types, to identify relevant tissues for each lipid trait. Consistent with previous studies[43,44] and our eQTL-based analysis, liver-related tissues (Tables S7 and S8) showed strong enrichments across all lipid traits (S-LDSC enrichment p values ranging from 0.001 in TG to 0.0001 in TC), for both expression (Figure 3A) and chromatin annotations (Figure 3B). This result was confirmed by analysis using two other approaches: DEPICT[54] (Figure S2 and Table S9) and RSS-NET[55] (Table S10). To assess the robustness of our S-LDSC results based on multi-ancestry GWASs, we applied S-LDSC to population-specific GWASs in European and East Asian ancestry participants together with population-specific LD scores and obtained similar results (Table S11, Figures S3 and S4).

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 ($p = 7.6 \times 10^{-4}$); for TG, enrichment in liver-related tissues ($p = 1.2 \times 10^{-3}$) is similar to enrichment in adipose ($p = 2.7 \times 10^{-3}$). For LDL, TC, and nonHDL, 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,[56] and the role of adipose as TG deposits.[57] These results were corroborated by eQTL colocalizations stratified by phenotype (Figure S1) and DEPICT analysis on gene expression[54] (Figure S2 and Table S9). Together, these results confirm the liver as a tissue of action for all five lipid traits and highlight the additional role of adipose primarily in HDL and TG.

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,[39] we performed gene set enrichment analysis[58] to identify cell-type clusters enriched for genes with eQTL colocalizations for any lipid trait. Out

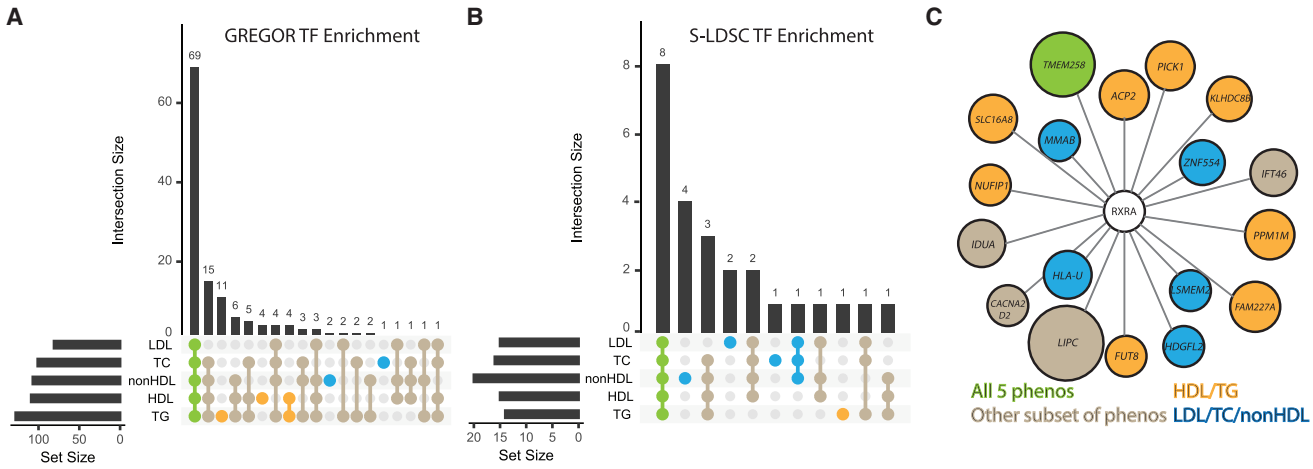

Figure 4. TF enrichment identified by GREGOR and S-LDSC
(A) Number of TFs enriched in the GREGOR analysis on GWAS loci for each of the five lipid traits.
(B) Number of TFs enriched in S-LDSC analysis on each of the five lipid traits.
(C) TF RXRA binds to the promoters of 26 colocalized genes (18 protein-coding). Colors represent the subsets of lipid phenotypes with colocalization. Larger node sizes represent smaller GWAS p values of colocalized loci.

of 11 identified cell types in 20 clusters in the liver, only hepatocytes were enriched at FDR-adjusted $p < 0.05$ (Figure S5 and Table S12), 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 S6 and Table S12). Of note, the enrichment in adipocytes was significant when we restricted this analysis to genes that were colocalized 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.[59] To understand lipid GWASs in the context of TF activity, we assessed enrichment of genome-wide significant variants at TF binding sites using GREGOR[38] and performed polygenic enrichment analysis of TF binding sites using S-LDSC. Because TFs were not comprehensively assayed in most cell lines (Figure S7), we used all cell types in our primary analysis presented below.

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,[38] 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 4A and Table S13). We obtained similar results when repeating the GREGOR analysis on TF binding sites derived from HepG2 only (Table S14). To assess the impact of GWAS power on TF enrichments, we repeated the GREGOR analysis on the same TF binding sites using a previous version of lipid GWAS, and we identified 54 enriched TFs (Table S15). Between the two versions of lipid GWASs, the total sample size and number of GWAS loci increased 8.7-fold (from 188,577 to 1,650,000) and 11-fold (from 156 to 1,750), respectively, but the number of enriched TFs only increased 2.5-fold (from 54 to 137), suggesting that the large number of enriched TFs is unlikely driven by the GWAS power alone.

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 4A and Table S13). The TF with the strongest enrichment in all phenotypes was ESRRA (estrogen-related receptor alpha), a nuclear receptor active in metabolic tissues;[60] ESRRA has been implicated in adipogenesis and lipid metabolism, and ESRRA-null mice display an increase in fat mass and obesity.[60]

The GREGOR analysis also highlighted 68 TFs significantly enriched in specific subsets of (but not all five) lipid phenotypes (Figure 4A and Table S13). 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[61] and fat cell differentiation has been demonstrated.[62] 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.

We also performed polygenic enrichment analysis of TF binding sites using S-LDSC (Figure 4B and Table S16), which differed from GREGOR analysis by looking at not only the genome-wide significant associations but also

the polygenic signal without GWAS p value cutoff. On the same 161 ENCODE TFs, this polygenic analysis identified 25 TFs whose binding sites were significantly enriched in heritability explained (nominal $p < 0.05$) for at least one lipid phenotype; reassuringly, 24 of 25 TFs were also significant in the GREGOR analysis. As a sensitivity check, we repeated the S-LDSC analysis on TF binding sites derived from HepG2 only, and we obtained similar results (Table S17).

Among 24 enriched TFs identified by both GREGOR and S-LDSC, eight were significantly enriched in all five lipid traits (CEBPB, CEBPD, FOXA2, HDAC2, HNF4G, NFYA, RXRA, SP1). RXRA (retinoid X receptor alpha) is 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.[63] Moreover, 145 lipid GWAS loci overlap RXRA binding peaks, and RXRA binds to the promoters of 26 colocalized genes (18 of which are protein-coding) (Figure 4C and Table S18), suggesting that the GWAS variants might affect lipids (partially) through affecting the binding activity of RXRA. While *RXRA* has been previously implicated,[64] 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 shortlisted 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 and Capture-C interactions in the liver or adipose. We next annotated each variant in the 95% credible set with various indicators of regulatory function, including its open chromatin status in liver[20] or adipose-related cell types,[65] its proximity to a promoter or an enhancer,[66] and its RegulomeDB regulation probability;[67] see Table S19 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 5A). 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 the 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 GWASs we prioritized 28 candidate loci with the strongest multi-layer evidence, 13 of which point to a single functional variant (Table 1). We have also made the full results of variant prioritization freely available (web resources). 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 5B). The *RRBP1* eQTL signals in the liver colocalize with LDL, TC, and nonHDL GWAS signals. The credible set at this locus contains a single lead variant (chr20:17,844,684). The "T" allele of this lead variant decreases *RRBP1* expression levels and increases LDL, TC, and nonHDL levels. This lead variant is in open chromatin in HLC and adipose and physically interacts with the *RRBP1* promoter (250 kb away) in adipose. 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 very low-density lipoprotein levels.[68] Rare variants in *RRBP1* are associated with LDL in humans[69] and silencing *RRBP1* in liver affects lipid homeostasis in mice.[68]

CREBRF (CREB3 regulatory factor) further demonstrates the power of our multi-layer integration framework in prioritizing functional variants (Figure 5C). 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 and is 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.[70]

Finally, to compare the power of functional fine-mapping with multi-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 ($p < 1 \times 10^{-4}$ based on 10,000 permutations randomly sampling variants from the European-only credible sets). This convergence of complementary approaches to the same smaller set of fine-mapped variants highlights the power of multi-ancestry datasets as an approach to narrow in on functional variants.

Discussion

Here we integrate the largest multi-ancestry 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

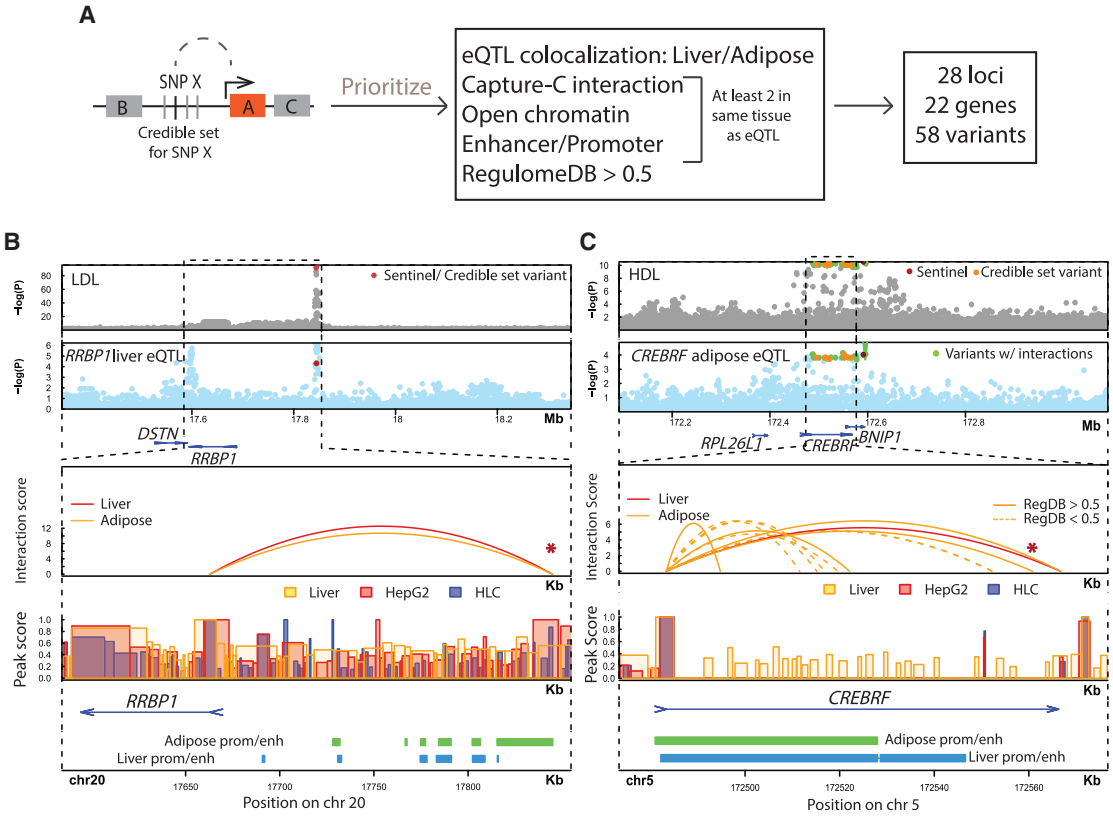

Figure 5. Multi-layer functional integration to prioritize variants at GWAS loci
(A) Variant annotation and prioritization scheme at each GWAS credible set.
(B) Evidence for *RRBP1* from functional genomics data. The LDL GWAS locus at this region (first row) is an eQTL for *RRBP1* in the liver (second row). Variants in the credible set of this locus interact with the gene promoter in both adipose and HepG2 Capture-C data (third row). The interacting variant is also in an open chromatin peak in three liver-related cell types (fourth row).
(C) Multiple sources of functional genomics data support *CREBRF* as a gene contributing to HDL levels. The HDL GWAS locus at this region (first row) is an eQTL for *CREBRF* in adipose (second row). Variants in the credible set at this locus interact with the CREBRF promoter in adipose (third row). The interacting variant is also in open chromatin in liver-related cell types (fourth row).

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 GWASs 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,[21,71,72] the size of our multi-ancestry 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 accessibility 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 GWASs, current functional datasets[73] 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

Table 1. Thirteen prioritized loci with highest confidence of a single functional variant in the credible set

Gene Name	Tissue	Sentinel	Prioritized Var	Open	CapC	Enhancer	Prom-oter	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	none	both	0.705
MICAL2	liver	11:12071855	12221016	liver	liver	none	liver	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	none	both	0.60906
FAM210A	liver	18:13725674	13725674	liver	liver	none	both	0.7571
RRBP1	liver	20:17844684	17844684	both	ad	both	none	0.6091

The "sentinel" column represents the lead variant at the locus. The "prioritized var" column 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 tissues. The "RegDB" column represents the RegulomeDB score of the prioritized variant.

are based on eQTLs from bulk tissue RNA-seq,[8,74] which may miss detailed cell types and biological processes in which lipid-associated SNPs regulate gene expression. Third, some functional datasets are not available across the full spectrum of human tissues and cell types. One example is that our chromatin-chromatin interaction analysis examines only a few cell types in two known lipid-related tissues (liver and adipose), producing results that may be biased toward known lipid biology. Another example is that ENCODE TF ChIP-seq data are not available in adipose-related cell lines. Fourth, our results are validated computationally but not experimentally. That said, our results provide a high-confidence list of regulatory mechanisms at lipid GWAS loci, forming a useful basis for future experiments. 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[75] 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 modeling multiple correlated SNPs through LD matrix,[76] which is computationally intensive in large-scale analyses derived from many cohorts with diverse ancestries, like the multi-ancestry GWASs examined here. Second, due to restricted access to individual genotypes of 201 cohorts, we cannot produce multi-ancestry 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.[77] 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[55]), but our approach simplifies computation and allows for scalability of the underlying 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.

Data and code availability

The accession number for the HLC Capture-C data reported in this paper is https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE189026.

Supplemental information

Supplemental information can be found online at https://doi.org/10.1016/j.ajhg.2022.06.012

Acknowledgments

Xiang Zhu is supported by the Stein Fellowship from Stanford University and Institute for Computational and Data Sciences Seed Grant from The Pennsylvania State University. C.D.B. is supported by the NIH (R01-HL133218). Funding for the Global Lipids Genetics Consortium was provided by the NIH (R01-HL127564). This research was conducted using the UK Biobank Resource under application number 24460. This research is based on data from the Million Veteran Program, Office of Research and Development, Veterans Health Administration, and was supported by awards 2I01BX003362-03A1 and 1I01BX004821-01A1. This publication does not represent the views of the Department of Veteran Affairs or the United States Government. We thank Bethany Klunder for administrative support. Study-specific acknowledgments are provided in the supplemental information.

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. Psaty serves on the steering committee of the Yale Open Data Access Project funded by Johnson & Johnson. G. Thorleifsson, A.H., D.F.G., H. Holm, 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. McCarthy has served on advisory panels for Pfizer, NovoNordisk, and Zoe Global and 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. McCarthy and A. Mahajan are employees of Genentech and 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, and 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 and 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. Kathiresan 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; and 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. Saleheen has received support from the British Heart Foundation, Pfizer, Regeneron, Genentech, and Eli Lilly pharmaceuticals. P.N. reports investigator-initated grants from Amgen, Apple, AstraZeneca, Boston Scientific, and Novartis, personal fees from Apple, AstraZeneca, Blackstone Life Sciences, Foresite Labs, Novartis, Roche / Genentech, is a co-founder of TenSixteen Bio, is a scientific advisory board member of Esperion Therapeutics, geneXwell, and TenSixteen Bio, and spousal employment at Vertex, all unrelated to the present work. The spouse of C.J.W. is employed by Regeneron.

Received: December 2, 2021
Accepted: June 23, 2022
Published: August 4, 2022

Web resources

Adipose single-cell data, https://singlecell.broadinstitute.org/single_cell/study/SCP133/human-adipose-svf-single-cell
bedtools, https://bedtools.readthedocs.io/en/latest/
biomaRt, https://bioconductor.org/packages/release/bioc/html/biomaRt.html
Browser of noncoding variant prioritization, http://csg.sph.umich.edu/willer/public/glgc-lipids2021/variant_prioritization.html
CHiCAGO, https://www.bioconductor.org/packages/release/bioc/html/Chicago.html
ClinVar, https://www.ncbi.nlm.nih.gov/clinvar/
ClusterProfiler, https://guangchuangyu.github.io/clusterProfiler
coloc, https://cran.r-project.org/web/packages/coloc
DEPICT, https://data.broadinstitute.org/mpg/depict
East Asian LD scores and related annotations, http://jenger.riken.jp/en/data
ENCODE ChIP-Seq data, https://hgdownload.cse.ucsc.edu/goldenpath/hg19/encodeDCC/wgEncodeRegTfbsClustered/wgEncodeRegTfbsClusteredWithCellsV3.bed.gz
European LD scores and related annotations, https://data.broadinstitute.org/alkesgroup/LDSCORE/
fgsea, http://bioconductor.org/packages/release/bioc/html/fgsea.html
GenomicRanges, https://bioconductor.org/packages/release/bioc/html/GenomicRanges.html
GLGC GWAS summary statistics and credible sets, http://csg.sph.umich.edu/willer/public/glgc-lipids2021/
GREGOR, https://genome.sph.umich.edu/wiki/GREGOR
GTEx v8 summary statistics, https://www.gtexportal.org/home/datasets
HepG2 Capture-C data (Chesi et al.[20]), https://www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-7144/
HepG2 Capture-C data (Selvarajan et al.[21]), https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE157306
HiCUP, https://www.bioinformatics.babraham.ac.uk/projects/hicup/
Human liver Hi-C data, https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE58752

Human white adipocyte Capture-C data, https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE110619
LDSC software, https://github.com/bulik/ldsc
liftOver, https://genome.ucsc.edu/cgi-bin/hgLiftOver
Liver single-cell data, http://shiny.baderlab.org/HumanLiverAtlas/
MGI, http://www.informatics.jax.org/downloads/reports/index.html#pheno
Open chromatin data from HepG2, https://www.omicsdi.org/dataset/arrayexpress-repository/E-MTAB-7543
Open chromatin data from adipose, https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE110734
Roadmap epigenomic data (promoters and enhancer annotation), https://egg2.wustl.edu/roadmap/data/byFileType/chromhmmSegmentations/ChmmModels/coreMarks/jointModel/final/
RegulomeDB, https://regulomedb.org/regulome-search/
RSS-NET, https://github.com/SUwonglab/rss-net
S-PrediXcan, https://github.com/hakyimlab/MetaXcan

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. https://doi.org/10.1016/j.ajhg.2018.04.002.
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. https://doi.org/10.3389/fgene.2020.00424.
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. https://doi.org/10.1038/s41576-018-0016-z.
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. https://doi.org/10.1038/nature13138.
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. https://doi.org/10.1038/nature09266.
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. https://doi.org/10.1038/s41576-019-0200-9.
7. ENCODE Project Consortium (2012). An integrated encyclopedia of DNA elements in the human genome. Nature *489*, 57–74. https://doi.org/10.1038/nature11247.
8. The GTEx Consortium (2020). The GTEx Consortium atlas of genetic regulatory effects across human tissues. Science *369*, 1318–1330. https://doi.org/10.1126/science.aaz1776.
9. Loos, R.J.F., and Yeo, G.S.H. (2021). The genetics of obesity: from discovery to biology. Nat. Rev. Genet. *23*, 120–133.
10. Huo, Y., Li, S., Liu, J., Li, X., and Luo, X.-J. (2019). Functional genomics reveal gene regulatory mechanisms underlying schizophrenia risk. Nat. Commun. *10*, 670. https://doi.org/10.1038/s41467-019-08666-4.
11. Willer, C.J., Sanna, S., Jackson, A.U., Scuteri, A., Bonnycastle, L.L., Clarke, R., Heath, S.C., Timpson, N.J., Najjar, S.S., Stringham, H.M., et al. (2008). Newly identified loci that influence lipid concentrations and risk of coronary artery disease. Nat. Genet. *40*, 161–169. https://doi.org/10.1038/ng.76.
12. Graham, S.E., Shoa, L., Clarke, K.-H.H.W., Kanoni, S., Zajac, G.J.M., Greg, J.M.Z., Ramdas, S., Surakka, I., Ntalla, I., Vedantam, S., et al. (2021). The power of genetic diversity in genome-wide association studies of lipids. Nature *600*, 675–679. https://doi.org/10.1038/s41586-021-04064-3.
13. McCarthy, S., Das, S., Kretzschmar, W., Delaneau, O., Wood, A.R., Teumer, A., Kang, H.M., Fuchsberger, C., Danecek, P., Sharp, K., et al. (2016). A reference panel of 64, 976 haplotypes for genotype imputation. Nat. Genet. *48*, 1279–1283. https://doi.org/10.1038/ng.3643.
14. 1000 G. Project Consortium, Auton, A., Brooks, L.D., Durbin, R.M., Garrison, E.P., Kang, H.M., Korbel, J.O., Marchini, J.L., McCarthy, S., McVean, G.A., and Abecasis, G.R. (2015). A global reference for human genetic variation. Nature *526*, 68–74. https://doi.org/10.1038/nature15393.
15. Mägi, R., Horikoshi, M., Sofer, T., Mahajan, A., Kitajima, H., Franceschini, N., McCarthy, M.I., Morris, A.P., T2D-GENES Consortium, and Morris, A.P. (2017). Trans-ethnic meta-regression of genome-wide association studies accounting for ancestry increases power for discovery and improves fine-mapping resolution. Hum. Mol. Genet. *26*, 3639–3650. https://doi.org/10.1093/hmg/ddx280.
16. The International HapMap 3 Consortium, Altshuler, D.M., Gibbs, R.A., Peltonen, L., Altshuler, D.M., Gibbs, R.A., Peltonen, L., Dermitzakis, E., Schaffner, S.F., Yu, F., Peltonen, L., et al. (2010). Integrating common and rare genetic variation in diverse human populations. Nature *467*, 52–58. https://doi.org/10.1038/nature09298.
17. Giambartolomei, C., Vukcevic, D., Schadt, E.E., Franke, L., Hingorani, A.D., Wallace, C., and Plagnol, V. (2014). Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. PLoS Genet. *10*, e1004383. https://doi.org/10.1371/journal.pgen.1004383.
18. Kuhn, R.M., Haussler, D., and Kent, W.J. (2013). The UCSC genome browser and associated tools. Briefings Bioinf. *14*, 144–161. https://doi.org/10.1093/bib/bbs038.
19. Çalışkan, M., Manduchi, E., Rao, H.S., Segert, J.A., Beltrame, M.H., Trizzino, M., Park, Y., Baker, S.W., Chesi, A., Johnson, M.E., et al. (2019). Genetic and epigenetic fine mapping of complex trait associated loci in the human liver. Am. J. Hum. Genet. *105*, 89–107. https://doi.org/10.1016/j.ajhg.2019.05.010.
20. Chesi, A., Wagley, Y., Johnson, M.E., Manduchi, E., Su, C., Lu, S., Leonard, M.E., Hodge, K.M., Pippin, J.A., Hankenson, K.D., et al. (2019). Genome-scale Capture C promoter interactions implicate effector genes at GWAS loci for bone mineral density. Nat. Commun. *10*, 1260. https://doi.org/10.1038/s41467-019-09302-x.
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. https://doi.org/10.1016/j.ajhg.2021.02.006.
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. https://doi.org/10.1016/j.stem.2017.03.017.

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. https://doi.org/10.1038/s41467-018-03554-9.
24. Wingett, S.W., 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. https://doi.org/10.12688/f1000research.7334.1.
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. https://doi.org/10.1186/s13059-016-0992-2.
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. https://doi.org/10.1101/gr.135350.111.
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. https://doi.org/10.1371/journal.pcbi.1003118.
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. https://doi.org/10.1038/nature14217.
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 J. Integr. Biol. *16*, 284–287. https://doi.org/10.1089/omi.2011.0118.
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. https://doi.org/10.1093/nar/gkv1070.
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. https://doi.org/10.1038/nprot.2009.97.
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. https://doi.org/10.1093/nar/gkz972.
33. Hindy, G., Dornbos, P., Chaffin, M.D., Liu, D.J., Wang, M., Selvaraj, M.S., Zhang, D., Park, J., Aguilar-Salinas, C.A., Antonacci-Fulton, L., et al. (2022). Rare coding variants in 35 genes associate with circulating lipid levels-A multi-ancestry analysis of 170, 000 exomes. Am. J. Hum. Genet. *109*, 81–96. https://doi.org/10.1016/j.ajhg.2021.11.021.
34. Ringwald, M., Richardson, J.E., Baldarelli, R.M., Blake, J.A., Kadin, J.A., Smith, C., and Bult, C.J. (2022). Mouse genome Informatics (MGI): latest news from MGD and GXD. Mamm. Genome *33*, 4–18. https://doi.org/10.1007/s00335-021-09921-0.
35. 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 phenotypic consequences of tissue specific gene expression variation inferred from GWAS summary statistics. Nat. Commun. *9*, 1825. https://doi.org/10.1038/s41467-018-03621-1.
36. Finucane, H.K., Bulik-Sullivan, B., Gusev, A., Trynka, G., Reshef, Y., Loh, P.-R., Anttila, V., Xu, H., Zang, C., Farh, K., et al. (2015). Partitioning heritability by functional annotation using genome-wide association summary statistics. Nat. Genet. *47*, 1228–1235. https://doi.org/10.1038/ng.3404.
37. Kanai, M., Akiyama, M., Takahashi, A., Matoba, N., Momozawa, Y., Ikeda, M., Iwata, N., Ikegawa, S., Hirata, M., Matsuda, K., et al. (2018). Genetic analysis of quantitative traits in the Japanese population links cell types to complex human diseases. Nat. Genet. *50*, 390–400. https://doi.org/10.1038/s41588-018-0047-6.
38. Schmidt, E.M., Zhang, J., Zhou, W., Chen, J., Mohlke, K.L., Chen, Y.E., and Willer, C.J. (2015). GREGOR: evaluating global enrichment of trait-associated variants in epigenomic features using a systematic, data-driven approach. Bioinformatics *31*, 2601–2606. https://doi.org/10.1093/bioinformatics/btv201.
39. MacParland, S.A., Liu, J.C., Ma, X.-Z., Innes, B.T., Bartczak, A.M., Gage, B.K., Manuel, J., Khuu, N., Echeverri, J., Linares, I., et al. (2018). Single cell RNA sequencing of human liver reveals distinct intrahepatic macrophage populations. Nat. Commun. *9*, 4383. https://doi.org/10.1038/s41467-018-06318-7.
40. Korotkevich, G., Sukhov, V., Budin, N., Shpak, B., Artyomov, M.N., and Sergushichev, A. (2021). Fast gene set enrichment analysis. Preprint at bioRxiv. https://doi.org/10.1101/060012.
41. Neumeyer, S., Hemani, G., and Zeggini, E. (2020). Strengthening causal inference for complex disease using molecular quantitative trait loci. Trends Mol. Med. *26*, 232–241. https://doi.org/10.1016/j.molmed.2019.10.004.
42. Tung, Y., Yeo, G.S.H., O'Rahilly, S., and Coll, A.P. (2014). Obesity and FTO: changing focus at a complex locus. Cell Metabol. *20*, 710–718. https://doi.org/10.1016/j.cmet.2014.09.010.
43. Hauberg, M.E., Zhang, W., Giambartolomei, C., Franzén, O., Morris, D.L., Vyse, T.J., Ruusalepp, A., Consortium, C.M., Sklar, P., Björkegren, J.L., et al. (2017). Large-scale identification of common trait and disease variants affecting gene expression. Am. J. Hum. Genet. *100*, 885–894. https://doi.org/10.1016/j.ajhg.2017.04.016.
44. Boix, C.A., James, B.T., Park, Y.P., Meuleman, W., and Kellis, M. (2021). Regulatory genomic circuitry of human disease loci by integrative epigenomics. Nature *590*, 300–307. https://doi.org/10.1038/s41586-020-03145-z.
45. GTE. Consortium, Laboratory, Data Analysis &Coordinating Center (LDACC)-Analysis Working Group, Statistical Methods groups-Analysis Working Group, Enhancing GTEx (eGTEx) groups, NIH Common Fund, NIH/NCI, NIH/NHGRI, NIH/NIMH, NIH/NIDA, Biospecimen Collection Source Site-NDRI (2017). Genetic effects on gene expression across human tissues. Nature *550*, 204–213.
46. Yoon, M. (2009). The role of PPARα in lipid metabolism and obesity: focusing on the effects of estrogen on PPARα actions. Pharmacol. Res. *60*, 151–159. https://doi.org/10.1016/j.phrs.2009.02.004.
47. Wang, H., and Eckel, R.H. (2009). Lipoprotein lipase: from gene to obesity. Am. J. Physiol. Endocrinol. Metab. *297*, E271–E288. https://doi.org/10.1152/ajpendo.90920.2008.

48. Blair, D.R., Lyttle, C.S., Mortensen, J.M., Bearden, C.F., Jensen, A.B., Khiabanian, H., Melamed, R., Rabadan, R., Bernstam, E.V., Brunak, S., et al. (2013). A nondegenerate code of deleterious variants in Mendelian loci contributes to complex disease risk. Cell *155*, 70–80. https://doi.org/10.1016/j.cell.2013.08.030.
49. Teslovich, T.M., Musunuru, K., Smith, A.V., Edmondson, A.C., Stylianou, I.M., Koseki, M., Pirruccello, J.P., Ripatti, S., Chasman, D.I., Willer, C.J., et al. (2010). Biological, clinical and population relevance of 95 loci for blood lipids. Nature *466*, 707–713. https://doi.org/10.1038/nature09270.
50. Brown, S.D.M., Holmes, C.C., Mallon, A.-M., Meehan, T.F., Smedley, D., and Wells, S. (2018). High-throughput mouse phenomics for characterizing mammalian gene function. Nat. Rev. Genet. *19*, 357–370. https://doi.org/10.1038/s41576-018-0005-2.
51. Gorkin, D.U., Leung, D., and Ren, B. (2014). The 3D genome in transcriptional regulation and pluripotency. Cell Stem Cell *14*, 762–775. https://doi.org/10.1016/j.stem.2014.05.017.
52. Yu, J., Hu, M., and Li, C. (2019). Joint analyses of multi-tissue Hi-C and eQTL data demonstrate close spatial proximity between eQTLs and their target genes. BMC Genet. *20*, 43. https://doi.org/10.1186/s12863-019-0744-x.
53. Matharu, N.K., and Ahanger, S.H. (2015). Chromatin insulators and topological domains: adding new dimensions to 3D genome architecture. Genes *6*, 790–811. https://doi.org/10.3390/genes6030790.
54. 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. https://doi.org/10.1038/ncomms6890.
55. 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. https://doi.org/10.1038/s41467-021-22588-0.
56. 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. https://doi.org/10.1186/s12944-019-1170-9.
57. Sniderman, A.D., Cianflone, K., Arner, P., Summers, L.K.M., and Frayn, K.N. (1998). The adipocyte, fatty acid trapping, and atherogenesis. Arterioscler. Thromb. Vasc. Biol. *18*, 147–151. https://doi.org/10.1161/01.atv.18.2.147.
58. 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., and Mesirov, J.P. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. USA *102*, 15545–15550. https://doi.org/10.1073/pnas.0506580102.
59. 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*, 6454. https://doi.org/10.3390/ijms22126454.
60. 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*, 1645. https://doi.org/10.3390/ijms21051645.
61. 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. https://doi.org/10.1371/journal.pone.0008460.
62. 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. https://doi.org/10.7554/elife.03346.
63. Neuschwander-Tetri, B.A. (2015). Retinoid X receptor: the forgotten partner in regulating lipid metabolism? Am. J. Clin. Nutr. *102*, 5–6. https://doi.org/10.3945/ajcn.115.114330.
64. 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. https://doi.org/10.1194/jlr.p008268.
65. Cannon, M.E., Currin, K.W., Young, K.L., Perrin, H.J., Vadlamudi, S., Safi, A., Song, L., Wu, Y., Wabitsch, M., Laakso, M., et al. (2019). Open chromatin profiling in adipose tissue marks genomic regions with functional roles in cardiometabolic traits. G3 (Bethesda) *9*, 2521–2533. https://doi.org/10.1534/g3.119.400294.
66. Epigenomics Consortium, R., Kundaje, A., Meuleman, W., Ernst, J., Bilenky, M., Yen, A., Heravi-Moussavi, A., Kheradpour, P., Zhang, Z., Ziller, M.J., et al. (2015). Integrative analysis of 111 reference human epigenomes. Nature *518*, 317–330. https://doi.org/10.1038/nature14248.
67. Boyle, A.P., Hong, E.L., Hariharan, M., Cheng, Y., Schaub, M.A., Kasowski, M., Karczewski, K.J., Park, J., Hitz, B.C., Weng, S., et al. (2012). Annotation of functional variation in personal genomes using RegulomeDB. Genome Res. *22*, 1790–1797. https://doi.org/10.1101/gr.137323.112.
68. 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. https://doi.org/10.1016/j.celrep.2021.108873.
69. 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. (2022). Analysis of rare genetic variation underlying cardiometabolic diseases and traits among 200, 000 individuals in the UK Biobank. Nat. Genet. *54*, 240–250. https://doi.org/10.1038/s41588-021-01011-w.
70. 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., et al. (2016). A thrifty variant in CREBRF strongly influences body mass index in Samoans. Nat. Genet. *48*, 1049–1054. https://doi.org/10.1038/ng.3620.
71. 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. https://doi.org/10.1038/s41588-018-0222-9.
72. 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 Metabol. *31*, 741–754.e5. https://doi.org/10.1016/j.cmet.2020.02.015.

73. 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. https://doi.org/10.1534/genetics.118.301525.
74. 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 *9*, e52155.
75. Arvanitis, M., Tayeb, K., Strober, B.J., and Battle, A. (2022). Redefining tissue specificity of genetic regulation of gene expression in the presence of allelic heterogeneity. Am. J. Hum. Genet. *109*, 223–239. https://doi.org/10.1016/j.ajhg.2022.01.002.
76. 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. https://doi.org/10.1214/17-aoas1046.
77. 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. https://doi.org/10.1038/s41586-019-1310-4.