OPEN ACCESS Freely available online

PLOS | GENETICS

Meta-Analysis of Genome-Wide Association Studies in African Americans Provides Insights into the Genetic Architecture of Type 2 Diabetes

Maggie C. Y. Ng[1,2☯], Daniel Shriner[3☯], Brian H. Chen[4,5], Jiang Li[2], Wei-Min Chen[6,7], Xiuqing Guo[8], Jiankang Liu[9], Suzette J. Bielinski[10], Lisa R. Yanek[11], Michael A. Nalls[12], Mary E. Comeau[13,14], Laura J. Rasmussen-Torvik[15], Richard A. Jensen[16,17], Daniel S. Evans[18], Yan V. Sun[19], Ping An[20], Sanjay R. Patel[21], Yingchang Lu[22,23], Jirong Long[24], Loren L. Armstrong[25], Lynne Wagenknecht[26], Lingyao Yang[14], Beverly M. Snively[14], Nicholette D. Palmer[1,2,27], Poorva Mudgal[2], Carl D. Langefeld[13,14], Keith L. Keene[28], Barry I. Freedman[29], Josyf C. Mychaleckyj[6,7], Uma Nayak[6,7], Leslie J. Raffel[30], Mark O. Goodarzi[30], Y-D Ida Chen[8], Herman A. Taylor Jr.[31,32], Adolfo Correa[31], Mario Sims[31], David Couper[33], James S. Pankow[34], Eric Boerwinkle[35], Adebowale Adeyemo[3], Ayo Doumatey[3], Guanjie Chen[3], Rasika A. Mathias[11,36], Dhananjay Vaidya[11,37], Andrew B. Singleton[12], Alan B. Zonderman[38], Robert P. Igo Jr.[39], John R. Sedor[40,41], the FIND Consortium‡, Edmond K. Kabagambe[42], David S. Siscovick[16,17,43], Barbara McKnight[16,44], Kenneth Rice[16,44], Yongmei Liu[45], Wen-Chi Hsueh[46], Wei Zhao[47], Lawrence F. Bielak[47], Aldi Kraja[20], Michael A. Province[20], Erwin P. Bottinger[22], Omri Gottesman[22], Qiuyin Cai[24], Wei Zheng[24], William J. Blot[48], William L. Lowe[25], Jennifer A. Pacheco[49], Dana C. Crawford[50], the eMERGE Consortium‡, the DIAGRAM Consortium‡, Elin Grundberg[51], the MuTHER Consortium‡, Stephen S. Rich[6], M. Geoffrey Hayes[25], Xiao-Ou Shu[24], Ruth J. F. Loos[22,23,52], Ingrid B. Borecki[20], Patricia A. Peyser[47], Steven R. Cummings[18], Bruce M. Psaty[16,17,43,53], Myriam Fornage[35], Sudha K. Iyengar[39], Michele K. Evans[54], Diane M. Becker[11,55], W. H. Linda Kao[37], James G. Wilson[56], Jerome I. Rotter[8], Michèle M. Sale[6,57,58], Simin Liu[4,59,60¶*], Charles N. Rotimi[3¶*], Donald W. Bowden[1,2,27¶*], for the MEta-analysis of type 2 DIabetes in African Americans (MEDIA) Consortium

1 Center for Genomics and Personalized Medicine Research, Wake Forest School of Medicine, Winston-Salem, North Carolina, United States of America, **2** Center for Diabetes Research, Wake Forest School of Medicine, Winston-Salem, North Carolina, United States of America, **3** Center for Research on Genomics and Global Health, National Human Genome Research Institute, Bethesda, Maryland, United States of America, **4** Program on Genomics and Nutrition, School of Public Health, University of California Los Angeles, Los Angeles, California, United States of America, **5** Center for Metabolic Disease Prevention, School of Public Health, University of California Los Angeles, Los Angeles, California, United States of America, **6** Center for Public Health Genomics, University of Virginia, Charlottesville, Virginia, United States of America, **7** Department of Public Health Sciences, University of Virginia, Charlottesville, Virginia, United States of America, **8** Institute for Translational Genomics and Population Sciences, Los Angeles BioMedical Research Institute at Harbor-UCLA Medical Center, Torrance, California, United States of America, **9** Department of Medicine, University of Mississippi Medical Center, Jackson, Mississippi, United States of America, **10** Division of Epidemiology, Department of Health Sciences Research, Mayo Clinic, Rochester, Minnesota, United States of America, **11** The GeneSTAR Research Program, Division of General Internal Medicine, Department of Medicine, The Johns Hopkins University School of Medicine, Baltimore, Maryland, United States of America, **12** Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, Maryland, United States of America, **13** Center for Public Health Genomics, Division of Public Health Sciences, Wake Forest School of Medicine, Winston-Salem, North Carolina, United States of America, **14** Department of Biostatistical Sciences, Division of Public Health Sciences, Wake Forest School of Medicine, Winston-Salem, North Carolina, United States of America, **15** Department of Preventive Medicine, Northwestern University Feinberg School of Medicine, Chicago, Illinois, United States of America, **16** Cardiovascular Health Research Unit, University of Washington, Seattle, Washington, United States of America, **17** Department of Medicine, University of Washington, Seattle, Washington, United States of America, **18** San Francisco Coordinating Center, California Pacific Medical Center Research Institute, San Francisco, California, United States of America, **19** Department of Epidemiology and Biomedical Informatics, Emory University, Atlanta, Georgia, United States of America, **20** Division of Statistical Genomics, Washington University School of Medicine, St. Louis, Missouri, United States of America, **21** Division of Sleep Medicine, Brigham and Women's Hospital, Boston, Massachusetts, United States of America, **22** The Charles Bronfman Institute for Personalized Medicine, Icahn School of Medicine at Mount Sinai, New York, New York, United States of America, **23** The Genetics of Obesity and Related Metabolic Traits Program, Icahn School of Medicine at Mount Sinai, New York, New York, United States of America, **24** Division of Epidemiology, Department of Medicine, Vanderbilt Epidemiology Center, Vanderbilt-Ingram Cancer Center, Vanderbilt University School of Medicine, Nashville, Tennessee, United States of America, **25** Division of Endocrinology, Metabolism and Molecular Medicine, Northwestern University Feinberg School of Medicine, Chicago, Illinois, United States of America, **26** Division of Public Health Sciences, Wake Forest School of Medicine, Winston-Salem, North Carolina, United States of America, **27** Department of Biochemistry, Wake Forest School of Medicine, Winston-Salem, North Carolina, United States of America, **28** Department of Biology, Center for Health Disparities, East Carolina University, Greenville, North Carolina, United States of America, **29** Department of Internal Medicine, Wake Forest School of Medicine, Winston-Salem, North Carolina, United States of America, **30** Medical Genetics Research Institute, Cedars-Sinai Medical Center, Los Angeles, California, United States of America, **31** Department of Medicine, University of Mississippi Medical Center, Jackson, Mississippi, United States of America, **32** Jackson State University, Tougaloo College, Jackson, Mississippi, United States of America, **33** Collaborative Studies Coordinating Center, Department of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, United States of America, **34** Division of Epidemiology and Community Health, University of Minnesota, Minneapolis, Minnesota, United States of America, **35** Human Genetics Center, University of Texas Health Science Center at Houston, Houston, Texas, United States of America, **36** Division of Allergy and Clinical Immunology, Department of Medicine, The Johns Hopkins University School of Medicine, Baltimore, Maryland, United States of America, **37** Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, United States of America, **38** Laboratory of Personality and Cognition, National Institute on Aging, National Institutes of Health, Baltimore, Maryland, United States of America, **39** Department of Epidemiology and Biostatistics, Case

Western Reserve University, Cleveland, Ohio, United States of America, **40** Department of Medicine, Case Western Reserve University, MetroHealth System campus, Cleveland, Ohio, United States of America, **41** Department of Physiology and Biophysics, Case Western Reserve University, Cleveland, Ohio, United States of America, **42** Division of Epidemiology, Department of Medicine, Vanderbilt University Medical Center, Nashville, Tennessee, United States of America, **43** Department of Epidemiology, University of Washington, Seattle, Washington, United States of America, **44** Department of Biostatistics, University of Washington, Seattle, Washington, United States of America, **45** Department of Epidemiology and Prevention, Division of Public Health Sciences, Wake Forest School of Medicine, Winston-Salem, North Carolina, United States of America, **46** Department of Medicine, University of California, San Francisco, California, United States of America, **47** Department of Epidemiology, School of Public Health, University of Michigan, Ann Arbor, Michigan, United States of America, **48** Division of Epidemiology, Department of Medicine, Vanderbilt Epidemiology Center, Vanderbilt-Ingram Cancer Center, Vanderbilt University School of Medicine, Nashville, Tennessee; International Epidemiology Institute, Rockville, Maryland, United States of America, **49** Center for Genetic Medicine, Northwestern University Feinberg School of Medicine, Chicago, Illinois, United States of America, **50** Center for Human Genetics Research and Department of Molecular Physiology and Biophysics, Vanderbilt University, Nashville, Tennessee, United States of America, **51** Department of Twin Research and Genetic Epidemiology, King's College London, London, United Kingdom, **52** Child Health and Development Institute, Icahn School of Medicine at Mount Sinai, New York, New York, United States of America, **53** Department of Health Services, University of Washington, Seattle, Washington, United States of America, **54** Health Disparities Unit, National Institute on Aging, National Institutes of Health, Baltimore Maryland, United States of America, **55** Department of Health Policy and Management, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, United States of America, **56** Department of Physiology and Biophysics, University of Mississippi Medical Center, Jackson, Mississippi, United States of America, **57** Department of Medicine, University of Virginia, Charlottesville, Virginia, United States of America, **58** Department of Biochemistry and Molecular Genetics, University of Virginia, Charlottesville, Virginia, United States of America, **59** Department of Epidemiology, University of California Los Angeles, Los Angeles, California, United States of America, **60** Departments of Epidemiology and Medicine, Brown University, Providence, Rhode Island, United States of America

Abstract

Type 2 diabetes (T2D) is more prevalent in African Americans than in Europeans. However, little is known about the genetic risk in African Americans despite the recent identification of more than 70 T2D loci primarily by genome-wide association studies (GWAS) in individuals of European ancestry. In order to investigate the genetic architecture of T2D in African Americans, the MEta-analysis of type 2 DIabetes in African Americans (MEDIA) Consortium examined 17 GWAS on T2D comprising 8,284 cases and 15,543 controls in African Americans in stage 1 analysis. Single nucleotide polymorphisms (SNPs) association analysis was conducted in each study under the additive model after adjustment for age, sex, study site, and principal components. Meta-analysis of approximately 2.6 million genotyped and imputed SNPs in all studies was conducted using an inverse variance-weighted fixed effect model. Replications were performed to follow up 21 loci in up to 6,061 cases and 5,483 controls in African Americans, and 8,130 cases and 38,987 controls of European ancestry. We identified three known loci (*TCF7L2*, *HMGA2* and *KCNQ1*) and two novel loci (*HLA-B* and *INS-IGF2*) at genome-wide significance ($4.15\times10^{-94}<P<5\times10^{-8}$, odds ratio (OR) = 1.09 to 1.36). Fine-mapping revealed that 88 of 158 previously identified T2D or glucose homeostasis loci demonstrated nominal to highly significant association (2.2×10^{-23} < locus-wide $P<0.05$). These novel and previously identified loci yielded a sibling relative risk of 1.19, explaining 17.5% of the phenotypic variance of T2D on the liability scale in African Americans. Overall, this study identified two novel susceptibility loci for T2D in African Americans. A substantial number of previously reported loci are transferable to African Americans after accounting for linkage disequilibrium, enabling fine mapping of causal variants in trans-ethnic meta-analysis studies.

Citation: Ng MCY, Shriner D, Chen BH, Li J, Chen W-M, et al. (2014) Meta-Analysis of Genome-Wide Association Studies in African Americans Provides Insights into the Genetic Architecture of Type 2 Diabetes. PLoS Genet 10(8): e1004517. doi:10.1371/journal.pgen.1004517

Editor: Eleftheria Zeggini, Wellcome Trust Sanger Institute, United Kingdom

Received January 21, 2014; **Accepted** June 5, 2014; **Published** August 7, 2014

Funding: Atherosclerosis Risk in Communities Study (ARIC) is carried out as a collaborative study supported by National Heart, Lung, and Blood Institute contracts (HHSN268201100005C, HHSN268201100006C, HHSN268201100007C, HHSN268201100008C, HHSN268201100009C, HHSN268201100010C, HHSN268201100011C, and HHSN268201100012C), R01HL087641, R01HL59367 and R01HL086694; National Human Genome Research Institute contract U01HG004402; and National Institutes of Health contract HHSN268200625226C. Infrastructure was partly supported by Grant Number UL1RR025005, a component of the National Institutes of Health and NIH Roadmap for Medical Research. Funding for Coronary Artery Risk Development in Young Adults (CARDIA) include support to University of Alabama at Birmingham (N01-HC-48047), University of Minnesota (N01-HC-48048), Northwestern University (N01-HC-48049), Kaiser Foundation Research Institute (N01-HC-48050), University of Alabama at Birmingham (N01-HC-95095), Tufts-New England Medical Center (N01-HC-45204), Wake Forest School of Medicine (N01-HC-45205), Harbor-UCLA Research and Education Institute (N01-HC-05187), University of California Irvine (N01-HC-45134, N01-HC-95100). Funding to Candidate-gene Association Resource (CARe) (http://public.nhlbi.nih.gov/GeneticsGenomics/home/care.aspx) include support to Massachusetts Institute of Technology - Broad Institute (N01-HC-65226). Cleveland Family Study (CFS) was supported by a grant to Case Western Reserve University (NIH HL 46380, M01RR00080). Cardiovascular Health Study (CHS) was supported by NHLBI contracts HL085251, HHSN268201200036C, HHSN268200800007C, N01HC55222, N01HC85079, N01HC85080, N01HC85081, N01HC85082, N01HC85083, N01HC85086; and NHLBI grants HL080295, HL087652, HL105756, HL103612, and HL120393 with additional contribution from the National Institute of Neurological Disorders and Stroke (NINDS). Additional support was provided through AG023629 rom the National Institute on Aging (NIA). A full list of principal CHS investigators and institutions can be found at CHS-NHLBI.org/. The provision of genotyping data was supported in part by the National Center for Advancing Translational Sciences, CTSI grant UL1TR000124, and the National Institute of Diabetes and Digestive and Kidney Disease Diabetes Research Center (DRC) grant DK063491 to the Southern California Diabetes Endocrinology Research Center. DNA handling and genotyping was supported in part by National Center of Advancing Translational Technologies CTSI grant UL1TR000124, the National Institute of Diabetes and Digestive and Kidney Diseases grant DK063491 to the Southern California Diabetes Endocrinology Research Center. The Electronic Medical Records and Genomics Network (eMERGE) Network was initiated and funded by NHGRI, with additional funding from NIGMS through the following grants: U01-HG-004610 (Group Health Cooperative); U01-HG-004608 (Marshfield Clinic); U01-HG-04599 (Mayo Clinic); U01HG004609 (Northwestern University); U01-HG-04603 (Vanderbilt University, also serving as the Coordinating Center), and the State of Washington Life Sciences Discovery Fund award to the Northwest Institute of Medical Genetics. The Northwestern University Enterprise Data Warehouse was funded in part by a grant from the National Center for Research Resources, UL1RR025741. The dataset(s) used for the analyses described were obtained from Vanderbilt University Medical Center's BioVU which is supported by institutional funding and by the Vanderbilt CTSA grant UL1 TR000445 from NCATS/NIH. Family Heart Study (FamHS) was supported by NIH grants R01-HL-087700 and R01-HL-088215 (MAP, Principal Investigator) from NHLBI; and R01-DK-8925601 and R01-DK-075681 (IBB, Principal Investigator) from NIDDK. Family Investigation of Nephropathy in Diabetes (FIND) was supported by FIND grants U01DK57292, U01DK57329, U01DK057300, U01DK057298, U01DK057249, U01DK57295, U01DK070657, U01DK057303, and U01DK070657, U01DK57304. This project has been funded in whole or in part with federal funds

from the NIH National Cancer Institute (NCI) under contract N01-CO-12400 and the Intramural Research Program of the NIH-NCI Center for Cancer Research. This work also was supported by the National Center for Research Resources for the General Clinical Research Center grants: Case Western Reserve University, M01-RR-000080; Wake Forest University, M01-RR-07122; Harbor-University of California, Los Angeles Medical Center, M01-RR-00425; College of Medicine, University of California, Irvine, M01-RR-00827-29; University of New Mexico, HSC M01-RR-00997; and Frederic C. Bartter, M01-RR-01346. The CHOICE Study was supported in part by HS08365 from the Agency for Healthcare Research and Quality, Rockville, MD, and HL62985 from the National Heart, Lung, and Blood Institute, Bethesda, MD. Genetic Study of Atherosclerosis Risk (GeneSTAR) was supported by NIH grants through the National Heart, Lung, and Blood Institute (HL58625-01A1, HL59684, HL071025-01A1, U01HL72518, and HL087698) and the National Institute of Nursing Research (NR0224103) and by M01-RR000052 to the Johns Hopkins General Clinical Research Center. Genetic Epidemiology Network of Arteriopathy (GENOA) study is supported by the National Institutes of Health grant numbers HL087660, HL100245 and HL100185 from the National Heart, Lung, and Blood Institute. Healthy Aging in Neighborhoods of Diversity across the Life Span Study (HANDLS) was supported by the Intramural Research Program of the NIH, National Institute on Aging and the National Center on Minority Health and Health Disparities (project # Z01-AG000513 and human subjects protocol # 2009-149). Health, Aging, and Body Composition Study (Health ABC Study) was supported by NIA contracts N01AG62101, N01AG62103, and N01AG62106. The genome-wide association study was funded by NIA grant 1R01AG032098-01A1 to Wake Forest University Health Sciences and genotyping services were provided by the Center for Inherited Disease Research (CIDR). CIDR is fully funded through a federal contract from the National Institutes of Health to The Johns Hopkins University, contract number HHSN268200782096C. This research was supported in part by the Intramural Research Program of the NIH, National Institute on Aging. Howard University Family Study (HUFS) was supported by National Institutes of Health grants S06GM008016-320107 to CNR and S06GM008016-380111 to AA. Participant enrollment was carried out at the Howard University General Clinical Research Center, supported by National Institutes of Health grant 2M01RR010284. Genotyping support was provided by the Coriell Institute for Medical Research. This research was supported by the Intramural Research Program of the Center for Research on Genomics and Global Health (CRGGH). The CRGGH is supported by the National Human Genome Research Institute, the National Institute of Diabetes and Digestive and Kidney Diseases, the Center for Information Technology, and the Office of the Director at the National Institutes of Health (Z01HG200362). The Charles Bronfman Institute for Personalized Medicine (IPM) BioBank Program is supported by The Andrea and Charles Bronfman Philanthropies. Insulin Resistance Atherosclerosis Study (IRAS) was supported by the National Heart, Lung, and Blood Institute (HL047887, HL047889, HL047890, HL47902). IRAS Family Study was supported by the National Heart, Lung, and Blood Institute (HL060944, HL060894, HL061210). The Jackson Heart Study (JHS) is supported by contracts HHSN268201300046C, HHSN268201300047C, HHSN268201300048C, HHSN268201300049C, HHSN268201300050C from the National Heart, Lung, and Blood Institute and the National Institute on Minority Health and Health Disparities. Multi-Ethnic Study of Atherosclerosis (MESA), MESA Family, and the MESA SHARe project are conducted and supported by the National Heart, Lung, and Blood Institute (NHLBI) in collaboration with MESA investigators. Support for MESA is provided by contracts N01-HC-95159 through N01-HC-95169 and UL1-RR-024156. Funding for MESA Family is provided by grants R01-HL-071051, R01-HL-071205, R01-HL-071250, R01-HL-071251, R01-HL-071252, R01-HL-071258, R01-HL-071259, and UL1-RR-025005. Funding for genotyping was provided by NHLBI Contract N02-HL-6-4278 and N01-HC-65226. MESA Air is funded by the US EPA - Science to Achieve Results (STAR)Program Grant #RD831697. The project described was supported by the National Center for Research Resources, Grant UL1RR033176, and is now at the National Center for Advancing Translational Sciences, Grant UL1TR000124. In Southern Community Cohort Study (SCCS), sample preparation was conducted at the Survey and Biospecimen Shared Resources, which is supported in part by Vanderbilt-Ingram Cancer Center (P30 CA68485). The SCCS dataset used for the present analyses was supported by U.S. NIH grant R01CA92447. In Sea Islands Genetic Network (SIGNET) - Reasons for Geographic And Racial Differences in Stroke (SIGNET-REGARDS), the REGARDS Study is supported by a cooperative agreement U01 NS041588 (PI George Howard) and SIGNET was supported by R01 DK084350 (MMS) from the National Institutes of Health. In Wake Forest School of Medicine (WFSM), genotyping services were provided by the Center for Inherited Disease Research (CIDR). CIDR is fully funded through a federal contract from the National Institutes of Health to The Johns Hopkins University, contract number HHSC268200782096C. The work at Wake Forest was supported by NIH grants K99 DK081350 (NDP), R01 DK066358 (DWB), R01 DK053591 (DWB), R01 HL56266 (BIF), R01 DK070941 (BIF) and in part by the General Clinical Research Center of the Wake Forest School of Medicine grant M01 RR07122. Women's Health Initiative (WHI) program is funded by the National Heart, Lung, and Blood Institute, National Institutes of Health, U.S. Department of Health and Human Services through contracts N01-WH-22110, 24152, 32100-2, 32105-6, 32108-9, 32111-13, 32115, 32118-32119, 32122, 42107-26, 42129-32, and 44221. Funding for WHI SHARe genotyping was provided by NHLBI Contract N02-HL-64278. BHC was funded by the Burroughs Wellcome Fund Inter-school Training Program in Metabolic Diseases and UCLA Genomic Analysis Training Program (NHGRI T32-HG002536). The contents of this publication are solely the responsibility of the authors and do not necessarily represent the official view of the National Institutes of Health. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* Email: simin_liu@brown.edu (SL); rotimic@mail.nih.gov (CNR); dbowden@wakehealth.edu (DWB)

❍ These authors contributed equally to this work.

¶ SL, CNR and DWB are joint senior authors on this work.

‡Membership of the FIND Consortium, the eMERGE Consortium, the DIAGRAM Consortium, and the MuTHER Consortium is provided in the Acknowledgments.

Introduction

The prevalence of type 2 diabetes (T2D) among adults in the USA is currently 11.3%, with substantially higher prevalence in African Americans (18.7%) than in European Americans (10.2%) [1]. To date, genome-wide association studies (GWAS) have identified >70 susceptibility loci for T2D [2–8]. While it is known that T2D is heritable in African Americans [9], it is unclear how much heritability is explained by the known genetic associations discovered primarily from European ancestry populations and whether there are risk loci specific to African Americans. Given that individuals of African ancestry tend to harbor more genetic diversity than individuals of other ancestries [10], we hypothesized that large-scale association analyses in African Americans could shed light on the genetic architecture of T2D and the risk attributable to cosmopolitan *vs*. population-specific variants.

Results

Study overview

We conducted a meta-analysis of 17 African American GWAS on T2D comprising 8,284 cases and 15,543 controls (Tables S1 and S2). Missing genotypes in individual studies were imputed to one of the HapMap reference panels (Phase II release 21–24 CEU+YRI, Phase II release 22 all populations, Phase II+III release 27 CEU+YRI, Phase II+III release 27 CEU+YRI+ASW or Phase II+III release 27 all populations) using MACH, IMPUTE2 or BEAGLE (Table S3). Genomic control corrections [11] were applied to each study ($\lambda = 1.01$–1.08) and after meta-analysis ($\lambda = 1.06$) due to modest inflated association results (Table S3) [12]. Association results for ~2.6M SNPs were subsequently examined.

From stage 1 meta-analysis, 49 SNPs moderately associated with T2D ($P<1\times10^{-5}$) and two candidate SNPs near the p value threshold (rs231356 at *KCNQ1*, $P = 2.84\times10^{-5}$ and rs2244020 at *HLA-B*, $P = 1.02\times10^{-5}$) totaling 51 SNPs in 21 loci were followed up for replication. rs231356 is 14 kb downstream of the reported T2D index SNP, rs231362, in Europeans [3]. Moderate associations have also been observed across the *HLA* region in Europeans [3]. The stage 2 replication included *in silico* and *de novo* replication in up to 11,544 African American T2D cases and controls, as well as *in silico* replication in 47,117 individuals of European ancestry from DIAGRAMv2 [3] (Table S4). Meta-analyses were performed to combine results from African Americans (stage 1+2a, $n\leq35{,}371$, Table S4) and both African Americans and Europeans (stage 1+2a+2b, $n\leq82{,}488$, Table S4).

T2D loci reaching genome-wide significance

Five independent loci reached genome-wide significance ($P<5\times10^{-8}$). Stage 1 meta-analysis identified the established *TCF7L2* locus. Stage 1+2a meta-analysis identified the established *KCNQ1* and *HMGA2* loci. Stage 1+2a+2b meta-analysis identified a second signal at *KCNQ1* and a novel *HLA-B* locus. Secondary analysis including body mass index (BMI) adjustment in stage 1+2a meta-analysis identified the second novel locus at *INS-IGF2* (Table 1 and Figure 1). None of the most strongly associated SNPs at these loci demonstrated significant heterogeneity of effect sizes among studies within each stage, between African Americans in

Author Summary

Despite the higher prevalence of type 2 diabetes (T2D) in African Americans than in Europeans, recent genome-wide association studies (GWAS) were examined primarily in individuals of European ancestry. In this study, we performed meta-analysis of 17 GWAS in 8,284 cases and 15,543 controls to explore the genetic architecture of T2D in African Americans. Following replication in additional 6,061 cases and 5,483 controls in African Americans, and 8,130 cases and 38,987 controls of European ancestry, we identified two novel and three previous reported T2D loci reaching genome-wide significance. We also examined 158 loci previously reported to be associated with T2D or regulating glucose homeostasis. While 56% of these loci were shared between African Americans and the other populations, the strongest associations in African Americans are often found in nearby single nucleotide polymorphisms (SNPs) instead of the original SNPs reported in other populations due to differential genetic architecture across populations. Our results highlight the importance of performing genetic studies in non-European populations to fine map the causal genetic variants.

stages 1 and 2a, or between African Americans in stage 1+2a and Europeans in stage 2b after Bonferroni correction of multiple comparisons ($P_{het}>0.001$) (Figure S1).

At the *TCF7L2* locus, the most strongly associated SNP in stage 1+2a African Americans samples was rs7903146 (OR = 1.33, $P=4.78\times10^{-44}$, Table 1 and Figure 2). rs7903146 is also the index SNP (most significantly associated with T2D in prior studies) in Europeans (OR = 1.40, $P=2.21\times10^{-51}$) [3], South Asians (OR = 1.25, $P=3.4\times10^{-19}$) [4] and East Asians (OR = 1.48, $P=2.44\times10^{-15}$) [13].

Two association signals were observed at *KCNQ1* (Table 1 and Figure 2). The first association signal was represented by rs2283228 located at the 3′ end of *KCNQ1* (stage 1+2a OR = 1.20, $P=9.90\times10^{-11}$; stage 1+2a+2b OR = 1.19, $P=4.87\times10^{-13}$). Using data from individuals of African ancestry in Southwest USA (ASW) from the 1000 Genomes Project (1KGP) [14], rs2283228 mapped to the same linkage disequilibrium (LD)-based interval as index SNPs from other populations (rs2283228 [15] and rs2237892 [16–17] in Japanese, rs2237892 in Hispanics [18], rs163182 [19] and rs2237895 [20] in Han Chinese). The second association signal was represented by rs231356 ($r^2=0$ with rs2283228 in both ASW and CEU) (stage 1+2a OR = 1.11, $P=1.94\times10^{-5}$; stage 1+2a+2b OR = 1.09, $P=3.93\times10^{-8}$), located 144 kb upstream of the first signal. rs231356 is located at the same LD interval as the index SNPs rs231362 in Europeans [3] and rs231359 in Chinese [20].

At the *HMGA2* locus, the most strongly associated SNP was rs343092 (stage 1+2a OR = 1.16, $P=8.79\times10^{-9}$; stage 1+2a+2b OR = 1.14, $P=2.75\times10^{-12}$; Table 1 and Figure 2). rs343092 is located 76 kb downstream and at the same LD interval as of the index SNP rs1531343 reported in Europeans [3].

Two novel T2D loci were identified. The effect sizes of rs2244020 located near *HLA-B* were similar in African Americans and Europeans (OR = 1.11 *vs*. 1.07, $P_{het}=0.26$; stage 1+2a+2b $P=6.57\times10^{-9}$) (Table 1 and Figure 2). *HLA-B* encodes the class I major histocompatibility complex involved in antigen presentation in immune responses.

The most strongly associated SNP near *INS-IGF2* was rs3842770 in African Americans (OR = 1.14, $P=2.78\times10^{-8}$, stage 1+2a BMI adjusted, Table 1 and Figure 2) but the risk A allele was absent in the CEU population. Insulin plays a key role in glucose homeostasis. Mutations at *INS* lead to neonatal diabetes, type 1 diabetes, and hyperinsulinemia [21]. Insulin-like growth factor 2 (IGF2) is involved in growth and development. IGF2 overexpression in transgenic mice leads to islet hyperplasia [22] and IGF2 deficiency in the Goto–Kakizaki rat leads to beta cell mass anomaly [23].

Associations at previously reported T2D and glucose homeostasis loci

We investigated index SNPs from 158 independent loci associated with T2D and/or glucose homeostasis from prior genome-wide and candidate gene studies in individuals of European, East Asian, South Asian, or African American ancestry (Table S5). Among the 104 T2D-associated index SNPs, 19 were associated with T2D in stage 1 African American samples ($P<0.05$). Most of the 17 T2D-associated SNPs that showed consistent direction of effects had similar effect sizes between this study and prior reports, despite that rs10440833 at *CDKAL1* had substantially stronger effect size in Europeans (OR = 1.25) than in African Americans (OR = 1.06, $P_{het}=5.86\times10^{-6}$). Additionally, 3 out of 54 trait-increasing alleles from glucose homeostasis-associated index SNPs were associated with increased T2D risk in African Americans ($P<0.05$).

We also performed a locus-wide analysis to test for associations of all SNPs within the LD region at $r^2\geq0.3$ with the previously reported index SNPs and results were corrected for the effective number of SNPs [24]. Since the causal variant(s) at each locus may be different or reside on different haplotypes across populations with different LD structures, this approach allows the identification of the most strongly associated SNPs in African Americans that may or may not be in LD with the index SNPs reported in other populations. A total of 55 T2D- and 29 glucose-associated loci were associated with T2D in African Americans ($P_{locus}<0.05$, corrected for LD in ASW for SNPs within a locus; Table S6). We compared the genetic architecture between the previously reported index SNPs and our fine-mapped SNPs for these 84 loci. The respective average risk allele frequencies were 0.51 and 0.46, and the distributions or pairwise differences of risk allele frequencies were not significantly different ($P=0.255$, Wilcoxon rank sum test; and $P=0.295$, Wilcoxon signed-rank test, respectively, Figure S2). In contrast, the average odds ratios for the risk alleles were higher for the fine-mapped SNPs as compared to the index SNPs (1.14 *vs*. 1.05). The distributions and pairwise differences of risk allele odds ratios were significantly different ($P=1.18\times10^{-19}$ and 5.55×10^{-14}, respectively, Figure S2). Thus, the locus-wide analysis identified variants with larger effect sizes and similar allele frequencies.

We leveraged differences in LD between African Americans and Europeans to fine-map and re-annotate several established loci. The association signal spanning ~100 kb at *INTS8* in African Americans overlapped the ~200 kb *TP53INP1* T2D locus in Europeans [3]. The most strongly associated SNP in MEDIA tended to have larger effect size in African Americans than in Europeans (rs17359493, OR = 1.13 vs. 1.06, $P=1.39\times10^{-7}$ vs. 3.20×10^{-2}, respectively, $P_{het}=0.06$) (Table S4). However, rs17359493 at intron 10 of *INTS8* was only in weak LD with the reported index SNP rs896854 in Europeans ($r^2=0.21$ in CEU, 0.10 in ASW). Neither the reported index SNP rs896854 nor its proxies from the CEU data demonstrated significant association to T2D in African Americans (Table S6 and Figure S3a,b), suggesting that rs17359493 may be an independent novel signal. *INTS8* encodes a subunit of the integrator complex which is involved in the cleavage of small nuclear RNAs. At *KCNQ1*, the most strongly associated SNP rs231356 was in weak LD with the

Table 1. Novel and previously identified loci associated with T2D at $P<5\times10^{-8}$.

						Stage 1 GWAS meta-analysis in in African Americans: up to 8,284 cases and 15,543 controls			Stage 2a replication in African Americans: up to 6,061 cases and 5,483 controls			Stage 1+2a meta-analysis in African Americans: up to 14,345 cases and 21,026 controls			Stage 2b replication in Europeans (DIAGRAMv2): up to 8,130 cases and 38,987 controls		Stage 1+2a+2b meta-analysis of all African Americans and Europeans: up to 22,475 cases and 60,013 controls		
Loci	**Chr**	**Position (Build 36)**	**SNP**	**Alleles[a]**	**RAF[b]**	**OR (95% CI)**	***P***	P_{het}	**OR (95% CI)**	***P***	P_{het}	**OR (95% CI)**	***P***	P_{het}	**OR (95% CI)**	***P***	**OR (95% CI)**	***P***	P_{het}
Previously identified T2D loci																			
TCF7L2[c]	10	114748339	rs7903146	T/C	0.30	1.32 (1.25–1.4)	**6.62E-24**	1.81E-01	1.34 (1.26–1.43)	**8.38E-20**	6.01E-03	1.33 (1.28–1.39)	**4.78E-44**	7.34E-01	1.4 (1.34–1.46)	**2.21E-51**	1.36 (1.32–1.4)	**4.15E-94**	1.16E-01
KCNQ1[c]	11	2661919	rs231356	T/A	0.27	1.14 (1.07–1.21)	2.84E-05	9.11E-01	1.05 (0.98–1.14)	1.68E-01	3.26E-01	1.11 (1.06–1.16)	1.94E-05	1.08E-01	1.08 (1.04–1.13)	4.37E-04	1.09 (1.06–1.13)	**3.93E-08**	5.27E-01
KCNQ1[c]	11	2806106	rs2283228	A/C	0.89	1.22 (1.14–1.31)	6.10E-08	9.48E-02	1.17 (1.06–1.28)	1.04E-03	7.10E-01	1.2 (1.14–1.27)	**9.90E-11**	4.34E-01	1.16 (1.06–1.26)	9.73E-04	1.19 (1.13–1.24)	**4.87E-13**	4.90E-01
HMGA2[c]	12	64537207	rs343092	T/G	0.81	1.16 (1.09–1.24)	1.91E-06	9.48E-01	1.15 (1.04–1.26)	3.99E-03	3.37E-01	1.16 (1.1–1.22)	**8.79E-09**	7.93E-01	1.12 (1.06–1.19)	5.43E-05	1.14 (1.1–1.19)	**2.75E-12**	4.41E-01
Newly identified T2D loci																			
HLA-B[c]	6	31455430	rs2244020	G/A	0.69	1.12 (1.06–1.17)	1.02E-05	2.11E-02	1.1 (0.98–1.22)	1.01E-01	1	1.11 (1.07–1.16)	1.14E-06	7.57E-01	1.07 (1.03–1.12)	7.67E-04	1.09 (1.06–1.13)	**6.57E-09**	2.55E-01
INS-IGF2[d]	11	2135246	rs3842770	A/G	0.23	1.18 (1.11–1.25)	8.18E-08	7.16E-01	1.07 (0.99–1.16)	8.09E-02	7.16E-01	1.14 (1.09–1.19)	**2.78E-08**	7.37E-02	-	-	-	-	-

Abbreviations: Chr, chromosome; SNP, single nucleotide polymorphism; RAF, risk allele frequency; OR, odds ratio for risk allele; CI, confidence interval; P_{het}, heterogeneity *P* value.

[a]Alleles are ordered as risk allele/other allele aligned to the forward strand of NCBI Build 36.

[b]Risk allele frequency in Stage 1 samples.

[c]Associations were performed with adjustment for age, sex, study sites, and study-specific principal components.

[d]Associations were performed with adjustment for age, sex, study sites, study-specific principal components and body mass index.

$P<5\times10^{-8}$ are in bold.

doi:10.1371/journal.pgen.1004517.t001

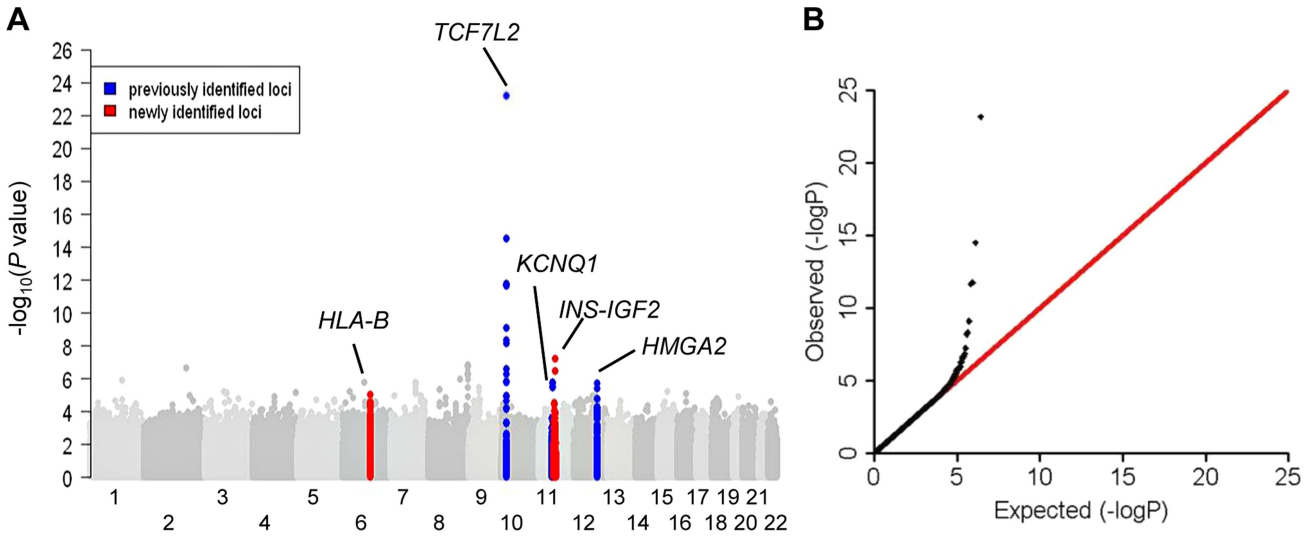

Figure 1. Association results of stage 1 meta-analysis in African Americans in a model adjusted for age, sex, study sites and study-specific principle components. (**A**) Manhattan plot. Previously identified loci are denoted in red. Novel loci identified in this study are denoted in blue. (**B**) Quantile-quantile plot. The observed *P* values (*y* axis) were compared with the expected *P* values under the null distribution (*x* axis).
doi:10.1371/journal.pgen.1004517.g001

index SNP rs231362 reported in Europeans [3] ($r^2 = 0.24$ in CEU and 0.17 in ASW). Given rs231362 was modestly associated with T2D in African American ($P = 0.04$) and was in weak LD ($r^2 = 0.21$ to 0.46 in CEU) with other associated SNPs in this region (Table S6 and Figure S3c,d), the results suggest a refinement of the localization of causal variant(s) to variants in strong LD with rs231356. At *HMGA2*, the most strongly associated SNP rs343092 was in moderate LD with the index SNP rs1531343 ($r^2 = 0.60$ in CEU and 0.32 in ASW). Despite rs1531343 and its proxies in high LD were not associated with T2D in African Americans ($P > 0.05$), several SNPs in moderate LD, including rs343092, showed nominal to strong associations (Table S6 and Figure S3e,f). Trans-ethic fine mapping will be particularly useful to dissect the causal variant(s) at this locus.

Effect of obesity on T2D susceptibility loci

We investigated the influence of obesity by comparing the stage 1 meta-analysis results with or without adjustment for BMI at the 51 most significantly associated SNPs from the GWAS for follow up (Tables S4 and S7) and 158 established T2D or glucose homeostasis index SNPs (Table S5). Association results were highly similar with and without BMI adjustment (correlation coefficients were 0.99 for both effect sizes and $-\log P$ values). Of particular note, *FTO* is suggested to influence T2D primarily through modulation of adiposity in Europeans [3,25], but evidence is contradictory across multiple ethnic groups [26–28]. The index SNP rs11642841 was not significantly associated with T2D in African Americans without and with BMI adjustment ($P = 0.06$ and 0.23, respectively) (Table S5). The frequency of the risk A allele was 0.13 in this study. It had 100% power to detect association at the reported OR of 1.13 at type 1 error rate of 0.05, suggesting that *FTO* is unlikely a key T2D susceptibility gene in African Americans.

Gene expression and bioinformatics analyses

Among the six genome-wide significant loci (Table 1), we found no coding variants in the most significantly associated SNPs or their proxies. These SNPs demonstrated only weak associations with expression quantitative trait loci (eQTLs) ($P > 0.001$, Table S8). Examination of the ENCODE data [29] revealed that several SNPs at *TCF7L2*, *KCNQ1*, and *HMGA2* were located at protein binding sites or were predicted to alter motif affinity for transcription factors implicated in energy homeostasis (Table S9). The most strongly associated SNP rs7903146 in *TCF7L2* is predicted to alter the binding affinity for a POU3F2 regulatory motif [30]. POU3F2 is a neural transcription factor that enhances the activation of genes regulated by corticotropin-releasing hormone which stimulates adrenocorticotropic hormone (ACTH). ACTH is synthesized from pre-pro-opiomelanocortin (pre-POMC) which regulates energy homeostasis. For the 3′ signal at *KCNQ1*, several tag SNPs are predicted to alter the binding affinity for regulatory motifs, including SREBP, CTCF and HNF4A. SREBP is a transcription factor involved in sterol biosynthesis. CTCF regulates the expression of IGF2 [31]. HNF4A is a master regulator of hepatocyte and islet transcription. The tag SNP rs2257883 at *HMGA2* is predicted to alter the binding affinity of MEF2, which regulates GLUT4 transcription in insulin responsive tissues [32].

Discussion

We have performed the largest genetic association analysis to date for T2D in African Americans. Our data support the hypothesis that risk for T2D is partly attributable to a large number of common variants with small effects [7]. We identified *HLA-B* and *INS-IGF2* as novel T2D loci, the latter specific to African Americans. We found evidence supporting association for 88 previously identified T2D and glucose homeostasis loci. Taken together, these 90 loci yielded a sibling relative risk of 1.19. The phenotypic variance measured on the liability scale is substantially larger in African Americans than in European Americans (17.5% *vs*. 5.7%) [7] due to larger effect sizes upon fine-mapping as well as higher disease prevalence in African Americans.

The two novel T2D loci, *HLA-B* and *INS-IGF2*, have been implicated in type 1 diabetes (T1D) risk in Europeans [33–35]. One limitation of our study is the lack of autoantibody measurement. However, our results are unlikely to be confounded by the presence of misclassified patients. Among diabetic youth

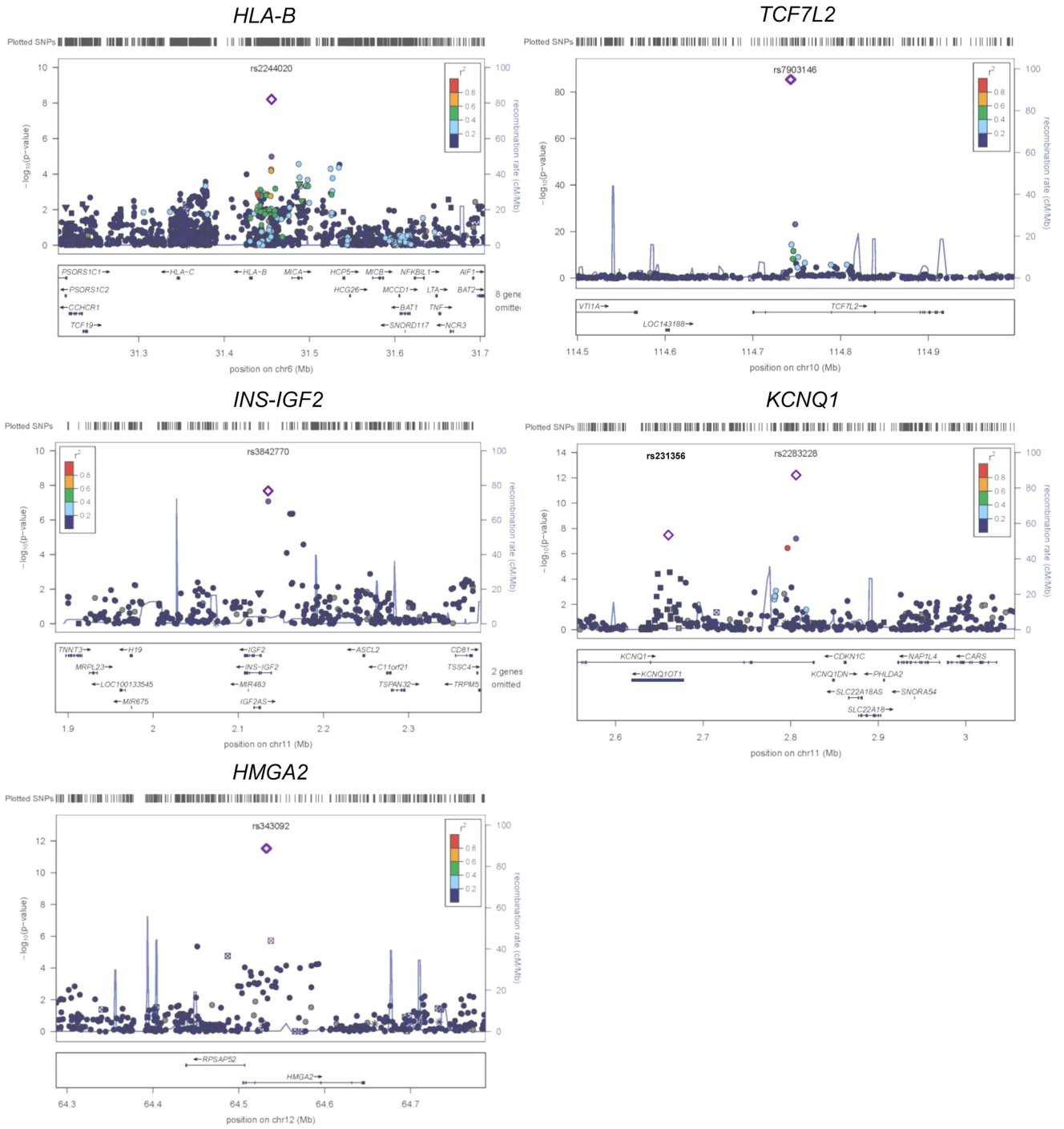

Figure 2. Regional plots of five previously and newly identified T2D loci in African Americans. Association *P* values (on a $-\log_{10}$ scale) of genotyped and imputed SNPs from stage 1 meta-analysis are plotted as a function of genomic position (NCBI Build 36). Plots for *HLA-B*, *TCF7L2*, *KCNQ1*, and *HMGA2* used the model without BMI adjustment whereas plots for *INS-IGF2* used the model with BMI adjustment. In each panel, the most strongly associated SNP from stage 1 and stage 1+2a+2b meta-analysis is denoted by a purple circle and a purple diamond, respectively. The color of all other SNPs indicates LD with the most strongly associated SNP based on the HapMap 2 YRI data. At *KCNQ1*, two independent signals are shown.
doi:10.1371/journal.pgen.1004517.g002

aged <20 years, T2D characterized by insulin resistance without autoimmunity is more prevalent in African Americans (40.1%) than in European Americans (6.2%), while African Americans less often present with autoimmunity and insulin deficiency resembling T1D compared to European Americans (32.5% vs. 62.9%, respectively) [36]. Autoimmunity is also uncommon in African American diabetic adults [37]. Furthermore, associations for T1D are stronger at *HLA* class II (*HLA-DRB1*, *-DQA1*, and *-DQB1*) than *HLA* class I regions in Europeans [33–34,38–41] (http://www.t1dbase.org). In African Americans, T1D individuals showed both shared and unique risk and protective *HLA* class II haplotypes as compared to European T1D individuals [42–43].

More importantly, these individuals also showed substantially stronger associations at *HLA* class II ($P<1\times10^{-25}$) than class I regions ($P<1\times10^{-5}$) [42], which is in contradiction with our finding of stronger associations at *HLA* class I than class II regions in T2D individuals (*HLA-B*, Figure S4). The observed *HLA-B* association may be due to LD with nearby causal gene(s) since there is long range LD in this region. Recently, rs3130501 near *POU5F1* and *TCF19* was reported for association with T2D in a trans-ancestry meta-analysis [8]. rs3130501 was located 211 kb upstream of rs2244020 and mapped to the same LD interval. However, the two SNPs were not correlated in both CEU ($D' = 0.57$, $r^2 = 0.05$) and ASW ($D' = 0.68$, $r^2 = 0.16$) from 1KGP nor strongly associated with T2D in the stage 1 meta-analysis ($P = 0.04$). Other potential non-*HLA* candidate genes may include *TNFA* which regulates immune and inflammatory response. It has been hypothesized that activated innate and adaptive immune cells stimulate release of cytokines such as TNFα and IL-1β, which promote both systemic insulin resistance and β-cell damage [44]. On the other hand, evidence has implicated T1D loci *HLA-DQ/DR*, *GLIS3* and *INS* in the susceptibility of latent autoimmune diabetes in adults (LADA) and/or T2D [7,34,45–46], while T2D loci such as *PPARG* and *TCF7L2* was associated with T1D [47] and LADA [46,48], respectively. More comprehensive studies are needed to understand the shared and distinct genetic risks in different forms of diabetes which will facilitate diagnosis and personalized treatment.

Our results have several implications regarding the genetic architecture of T2D. First, fine-mapping suggests that currently known loci explain more of the risk than previously estimated. Second, the loci conferring the largest risk for T2D appear to act through regulatory rather than protein-coding changes. Third, many, but not all, of the previously identified T2D loci are shared across ancestries. The differential LD structure of African-ancestry populations at shared loci provides an opportunity for fine mapping in trans-ethnic meta-analysis. Fourth, the ~2.6M MEDIA SNPs achieved only 43.3% coverage of the 1KGP ASW common SNPs, suggesting that risk loci that are specific to African-ancestry individuals are difficult to discover with the genotyping arrays being used. Large-scale sequencing studies, such as those focusing on whole genomes, exomes, and targeted resequencing for associated non-coding regions, will be necessary to further delineate the causal variants for T2D risk in African Americans.

Materials and Methods

Samples and clinical characterization

Stage 1 discovery samples included 17 T2D GWAS studies (ARIC, CARDIA, CFS, CHS, FamHS, GeneSTAR, GENOA, HANDLS, Health ABC, HUFS, JHS, MESA, MESA Family, SIGNET-REGARDS, WFSM, FIND, and WHI) with up to 23,827 African American subjects (8,284 cases and 15,543 controls). Stage 2 replication samples included up to 11,544 African American subjects (6,061 cases and 5,483 controls), using *in silico* replication of GWAS data from eMERGE and IPM Biobank and *de novo* genotyping in IRAS, IRASFS, SCCS, and WFSM. In general, T2D cases were defined as having at least one of the following: fasting plasma glucose ≥126 mg/dl, 2 hour glucose during oral glucose tolerance test (OGTT) ≥200 mg/dl, random glucose ≥200 mg/dl, oral hypoglycemic agent or insulin treatment, or physician-diagnosed diabetes. All cases were diagnosed at ≥25 years (or age at study ≥25 years if age at diagnosis was not available). For cohort studies, individuals who met the criteria at any of the visits were defined as cases. Controls with normal glucose tolerance (NGT) were defined by satisfying all the following criteria: fasting plasma glucose <100 mg/dl, 2 hour OGTT<140 mg/dl (if available), no treatment of diabetes, and age ≥25 years. For cohort studies, individuals who met the criteria at all visits were defined as controls. All study participants provided written informed consent, except for eMERGE that use an opt out program, and approval was obtained from the institutional review board (IRB) from the respective local institutions. Detailed descriptions of the participating studies are provided in Text S1.

Genotyping, imputation and quality control

For stage 1 and 2 GWAS studies, genotyping was performed with Affymetrix or Illumina genome-wide SNP arrays. Imputation of missing genotypes was performed using MACH [49], IMPUTE2 [50] or BEAGLE [51] using HapMap reference haplotypes. For each study, samples reflecting duplicates, low call rate, gender mismatch, or population outliers were excluded. In general, SNPs were excluded by the following criteria: call rate < 0.95, minor allele frequency (MAF)<0.01, minor allele count < 10, Hardy-Weinberg P-value $<1\times10^{-4}$, or imputation quality score <0.5 (Table S3). For *de novo* replication studies, genotyping was performed using the Sequenom MassArray platform (Sequenom; San Diego, CA). Sample and SNP quality controls were performed as with GWAS data.

Statistical analysis

Single SNP association was performed for each study by regressing T2D case/control status on genotypes. To account for uncertainty of genotype calls during imputation, genotype probabilities or dosage were used for association tests in imputed SNPs. The association tests assumed an additive genetic model and adjusted for age, sex, study centers, and principal components. Principal components were included to control for confounding effects of admixture proportion and population structure. Secondary analysis with additional adjustment for BMI was performed for SNPs with $P<1\times10^{-5}$ in stage 1 meta-analysis and index SNPs previously reported to be associated with T2D or glucose homeostasis traits. BMI adjustment allows increasing power to detect T2D loci independent of BMI effect and diminish associations at T2D loci with effects modulated through BMI. Logistic regression was used for samples of unrelated individuals. Generalized estimating equations [52] or SOLAR [53] were used for samples of related individuals. Association results with extreme values (absolute beta coefficient or standard error >10), primarily due to low cell counts resulting from small sample sizes and/or low minor allele frequencies, were excluded (Table S3).

Meta-analysis

In stage 1, association results were combined by a fixed effect model with inverse variance weighted method using the METAL software [12]. Genomic control correction [11] was applied to each study before meta-analysis, and to the overall results after meta-analysis. Results from SNPs genotyped in <10,000 samples and those with allele frequency difference >0.3 among studies were excluded. A total of 2,579,389 SNPs were analyzed in the meta-analysis (Table S3). In stage 2a, association results from African American replication studies were also combined using a fixed effect inverse variance weighted method. To assess the overall effects in African Americans (stage 1+2a) and both African Americans and Europeans (stage 1+2a+2b), association results from studies in the respective stages were combined using a fixed effect inverse variance weighted method. Genome-wide significance is declared at $P<5\times10^{-8}$ from the meta-analysis result of all stages, which has better power than the replication-based strategy [54].

Among the 51 SNPs carried forward for replication, heterogeneity of effect sizes across studies within each stage was assessed using Cochran's Q statistic implemented in METAL. Meta-analysis results from stages 1 and 2a, stage 1+2a and 2b were used to assess heterogeneity of effect sizes between discovery and replication stages in African Americans, and between African Americans and Europeans, respectively. For SNPs with significant heterogeneous effect size after multiple comparison corrections ($P_{het}<0.001$), meta-analysis results including studies of all stages assessed by the random effect model implemented in GWAMA [55] were reported. Heterogeneous associations may partly due to differences in ascertainment scheme across studies. For index SNPs reported in prior studies, assessment of heterogeneity using Cochran's Q statistic between prior studies and this study were also reported.

Transferability analysis

Index SNPs associated with T2D or glucose homeostasis traits from prior GWAS and candidate gene studies were examined for association with T2D in African Americans (Table S5). For the index SNP association tests, a per-SNP P value <0.05 was defined as significant. In the locus-wide analysis, the boundaries of a locus were defined by the most distant markers (within ± 500 kb) using the 1KGP CEU data with $r^2 \geq 0.3$ with the index SNP. All MEDIA SNPs within these bounds were examined for association analysis. All pairwise LD values within each locus were estimated using the 1KGP CEU and ASW data. To estimate the effective number of SNPs at a locus, we retrieved genotypes from the 1KGP ASW data for markers present in MEDIA, estimated the sample covariance matrix from those genotypes, and spectrally decomposed the covariance matrix [24]. The effective number of SNPs was estimated using the relationship $N_{eff} = \left(\sum_{k=1}^{K} \lambda_k\right)^2 \Big/ \left(\sum_{k=1}^{K} \lambda_k^2\right)$, in which λ_k is the k^{th} eigenvalue of the $K \times K$ covariance matrix for the K SNPs in the locus [24]. The per-locus significance level was defined as 0.05/effective number of SNPs (Table S6). By accounting for all SNPs within the bounds of LD, the per-locus significance level is corrected to account for markers in LD with the index SNP as well as markers not in LD with the index SNP, thereby potentially allowing for discovery of new associations at markers not tagged by the index SNP.

Liability-scale variance explained

For each independent locus, we estimated the sibling relative risk using the most strongly associated SNP within that locus. Let p_i and ψ_i be the risk allele frequency and the corresponding odds ratio at the i^{th} SNP, respectively. Assuming the additive genetic model and independence between SNPs, the contribution to the sibling relative risk λ_s for a set of N SNPs is given by $\lambda_s = \prod_{i=1}^{N} \left[\left(1 + \frac{p_i(1-p_i)(\psi_i - 1)^2}{2((1-p_i)+p_i\psi_i)^2} \right)^2 \right]$ [56]. Let K be the disease prevalence. The liability-scale variance h_L^2 explained by the set of N SNPs is given by $h_L^2 = \frac{2\left(T - T_1\sqrt{1-\left(T^2 - T_1^2\right)\left(1-\frac{T}{\omega}\right)}\right)}{\omega + T_1^2(\omega - T)}$, in which $T = \phi^{-1}(1-K)$, $T_1 = \phi^{-1}(1-\lambda_s K)$, and $\omega = \frac{z}{K}$, with ϕ^{-1} representing the standard normal quantile function and z representing the standard normal density at T [57].

Coverage

The coverage of MEDIA SNPs to the human genome was estimated using HaploView [58] via pairwise tagging at the $r^2 = 0.8$ threshold. We used all SNPs with minor allele frequencies $\geq 1\%$ in both MEDIA and the 1KGP ASW sequence data. Coverage was estimated using non-overlapping bins of 1,000 SNPs.

Power analysis

Study power was calculated using the genetic power calculator [59]. For SNPs with MAF$\geq$0.3, our study had $>80\%$ power to detect odds ratios for T2D at OR$\geq$1.06 and $\geq$1.13 at $P<0.05$ and $P<5\times10^{-8}$, respectively, in stage 1 samples under an additive model. The observed odds ratios among our stage 1 most significantly associated SNPs with $P<1\times10^{-5}$ ranged from 1.11 to 1.56 (Table S4). Given our African American sample size in stage 1+2a, our study had $>80\%$ power to detect OR$\geq$1.1 at $P<5\times10^{-8}$ at MAF$\geq$0.3, thus provided good power to detect genome-wide significance among the most significantly associated SNPs using all African American samples. For T2D SNPs reported from the literature, power was also calculated from the reported effect size using the risk allele frequency from this study for stage 1 samples at $P<0.05$ and $P<5\times10^{-8}$, respectively (Table S5).

Gene expression analysis

The MuTHER resource (www.muther.ac.uk) includes lymphoblastoid cell lines (LCLs), skin, and adipose tissue derived simultaneously from a subset of well-phenotyped healthy female twins from the TwinsUK adult registry [60]. Whole-genome expression profiling of the samples, each with either two or three technical replicates, was performed using the Illumina Human HT-12 V3 BeadChips (Illumina Inc.) according to the protocol supplied by the manufacturer. Log_2-transformed expression signals were normalized separately per tissue as follows: quantile normalization was performed across technical replicates of each individual followed by quantile normalization across all individuals. Genotyping was performed with a combination of Illumina arrays (HumanHap300, HumanHap610Q, 1M-Duo, and 1.2MDuo 1M). Untyped HapMap2 SNPs were imputed using the IMPUTE2 software package. In total, 776 adipose and 777 LCL samples had both expression profiles and imputed genotypes. Association between all SNPs (MAF$>5\%$, IMPUTE info >0.8) within a gene or within 1 Mb of the gene transcription start or end site and normalized expression values were performed with the GenABEL/ProbABEL packages [61–62] using the polygenic linear model incorporating a kinship matrix in GenABEL followed by the ProbABEL mmscore score test with imputed genotypes. Age and experimental batch were included as cofactors.

Genotype and gene expression in LCL in HapMap samples were also available [63]. Association of genotypes and gene expression of transcripts within 1 MB of tested SNPs were analyzed separately for CEU and YRI populations. The variance components model implemented in SOLAR was used for association analysis which accounts for correlation among related individuals [53].

In this study, we examined the association of the most significantly associated SNPs from the six genome-wide significant loci and their proxies ($r^2 \geq 0.8$ in ASW) within 1 Mb of the associated SNPs with *cis*-expression quantitative trait loci (eQTLs) in peripheral blood leukocytes (LCL) and adipose tissue (Table S8).

ENCODE data analysis

We examined putative function of non-coding genome-wide significant SNPs and their proxies within 1 Mb ($r^2 \geq 0.8$ in 1KGP ASW) using HaploReg [30] and RegulomeDB [64]. These databases interrogated multiple chromatin features from the

Encyclopedia of DNA Elements (ENCODE) project [29]. High priority was given to variants annotated as protein-binding via ChIP-seq, and motif-changing via position weight matrices, with the respective transcription factors implicated in diabetes pathogenesis and related biological processes.

Supporting Information

Figure S1 Forest plots of the most strongly associated SNPs at five previously and newly identified T2D loci in African Americans. Odds ratio and 95% CIs are presented for individual studies (black circle and line) and meta-analysis results (red diamond and line). At *KCNQ1*, two independent associated SNPs are shown.
(PDF)

Figure S2 (A) Distributions of risk allele frequencies for the previously reported index SNPs (in black) vs. the MEDIA most strongly associated SNPs (in red) in African Americans from stage 1 meta-analysis. (B) Distributions of odds ratios for risk alleles of the index SNPs (in black) vs. the most strongly associated MEDIA SNPs (in red) in African Americans from stage 1 meta-analysis.
(PDF)

Figure S3 Regional plots of stage 1 meta-analysis association results in African Americans for the most strongly associated SNPs from this study and the index SNPs from previous studies. (A–B) *INTS8-TP53INP1* region; (C–D) *KCNQ1* region; (E–F) *HMGA2* region. (A, C, E) The most strongly associated SNP in MEDIA is denoted by a purple circle and a red arrow with LD colored based on the HapMap 2 YRI data. (B, D, F) The index SNP is denoted by a purple circle and a blue arrow with LD colored based on the HapMap 2 CEU data.
(PDF)

Figure S4 Regional plots of *HLA-B* and *HLA-DQ/DR* regions for (A, C) stage 1 meta-analysis association results in African Americans and HapMap 2 YRI LD data, and (B, D) stage 3 DIAGRAMv2 results in Europeans using HapMap 2 CEU LD data. (A, B) The most strongly associated SNP rs2244020 at *HLA-B* region from this study is denoted by a purple circle and a red arrow. (C, D) The index SNP rs9272346 from Burton PR *et al* (2007) [65] is denoted by a purple circle and a blue arrow.
(PDF)

Table S1 Design of studies in stage 1 GWAS and stage 2a replication in African Americans.
(PDF)

Table S2 Clinical characteristics of study samples in stage 1 GWAS and stage 2a replication studies in African Americans.
(PDF)

Table S3 Genotyping methods, quality controls, imputation and statistical analysis in stage 1 GWAS and stage 2a replication studies in African Americans.
(PDF)

Table S4 SNPs with P value$\leq 1\times10^{-5}$ from stage 1 GWAS meta-analysis (BMI unadjusted) selected for stage 2 *in silico* and *de novo* replication in African Americans and *in silico* replication in individuals of European ancestry from DIAGRAMv2.
(PDF)

Table S5 Stage 1 GWAS meta-analysis results for index SNPs at established T2D or glucose homeostasis loci in African Americans.
(PDF)

Table S6 Locus-wide association at established T2D or glucose homeostasis loci in stage 1 GWAS meta-analysis in African Americans.
(PDF)

Table S7 BMI-adjusted association for SNPs from stage 1 GWAS meta-analysis selected for replication.
(PDF)

Table S8 Expression Quantitative Trait Loci (eQTL) analysis for the genome-wide significant SNPs for T2D. Results are shown for suggestive evidence of *cis*-association ($P<0.05$) between the genome-wide significant SNPs and their proxies with the genes within 1 Mb of the associated SNPs.
(PDF)

Table S9 Putative regulatory SNPs predicted from the ENCODE project for the genome-wide significant SNPs and their proxies at *TCF7L2*, *INS-IGF2*, *KCNQ1* and *HMGA2*.
(PDF)

Text S1 Description of GWAS and replication studies.
(PDF)

Acknowledgments

We thank all the study participants for their valuable contributions to the parent studies (ARIC, CARDIA, CFS, CHS, eMERGE, FamHS, FIND, GeneSTAR, GENOA, HANDLS, Health ABC, HUFS, IPM Biobank, IRAS, IRASFS, JHS, MESA, MESA Family, SCCS, SIGNET-REGARDS, WFSM and WHI) of the MEDIA consortium. We thank the contributions of investigators and staff of the parent studies for data collection, genotyping, data analysis and sharing of association results to the MEDIA consortium at the discovery and replication stages. We also thank the DIAGRAM Consortium for sharing the European T2D GWAS meta-analysis results for replication. We thank the MuTHER Consortium to share the expression quantitative trait loci data. The Candidate-gene Association Resource (CARe) acknowledge the support of the National Heart, Lung, and Blood Institute and the contributions of the involved research institutions, study investigators, field staff, and study participants of ARIC, CARDIA, CFS, JHS, and MESA in creating the Candidate-gene Association Resource for biomedical research (http://public.nhlbi.nih.gov/GeneticsGenomics/home/care.aspx). The parent studies have contributed parent study data, ancillary study data, and DNA samples through the Massachusetts Institute of Technology - Broad Institute to create this genotype/phenotype database for wide dissemination to the biomedical research community. Women's Health Initiative (WHI) investigators have reviewed and approved this manuscript. WHI investigators are listed at http://www.whiscience.org/publications/WHI_investigators_shortlist_2010-2015.pdf. The datasets used for the analyses described in this manuscript were obtained from dbGaP at http://www.ncbi.nlm.nih.gov/sites/entrez?db=gap through dbGaP accession phs000200.v1.p1.

Members of the FIND Consortium are:

Genetic Analysis and Data Coordinating Center, Case Western Reserve University, Cleveland, OH: RC Elston, S Iyengar, K Goddard, J Olson, S Ialacci, S Edwards, C Fondran, A Horvath, G Jun, K Kramp, M Slaughter, E Zaletel.

Clinical centers:

Case Western Reserve University, Cleveland, OH: JR Sedor, J Schelling, A Sehgal, A Pickens, L Humbert, L Getz-Fradley.

Harbor—University of California Los Angeles Medical Center: S Adler, HE Collins-Schramm§, E Ipp, H Li§, M. Pahl†, MF Seldin§, J LaPage, B Walker, C Garcia, J Gonzalez, L Ingram-Drake.

§University of California, Davis, CA.

†University of California, Irvine, CA.

Johns Hopkins University, Baltimore, MD: M. Klag, J Coresh, L Kao, L Mead, R Parekh, N Fink, P Bayton, Y Long, L Wei, T Whitehead.

National Institute of Diabetes and Digestive and Kidney Diseases, Phoenix, AZ: WC Knowler, RL Hanson, RG Nelson, L Jones, R Juan, R Lovelace, C Luethe, LM Phillips, J Sewemaenewa, I Sili, B Waseta.

University of California, Los Angeles, CA: MF Saad, X Guo, J Rotter, K Taylor, M Budgett.

University of New Mexico, Albuquerque, NM: P Zager, V Shah, M Scavini, A Bobelu.

University of Texas Health Science Center at San Antonio, San Antonio, TX: H Abboud, N Arar, R Duggirala, BS Kasinath, R Plaetke, M Stem, C Goyes, V Sartorio.

Wake-Forest University, Winston-Salem, NC: BI Freedman, DW Bowden, SC Satko, SS Rich, S Warren, S Viverette, G Brooks, R Young.

Laboratory of Genomic Diversity, National Cancer Institute, Frederick, MD: C Winkler, MW Smith, M Thompson, R Hanson.

National Institute of Diabetes and Digestive and Kidney Diseases program office: JP Briggs, PL Kimmel, R Rasooly.

External Advisory Committee: D Warnock (chair), R Chakraborty, GM Dunston, RP Lifton, SJ O'Brien (ad hoc), R Spielman.

Members of the eMERGE Consortium are:

Marshfield Clinic, Marshfield, WI: Catherine A. McCarty, Justin Starren, Peggy Peissig, Richard Berg, Luke Rasmussen, James Linneman, Aaron Miller, Vidhu Choudary, Lin Chen, Carol Waudby, Terrie Kitchner, Jonathan Reeser, Norman Fost, Marylyn Ritchie, Russell A. Wilke.

Northwestern University, Chicago, IL: Rex L. Chisholm, Pedro C. Avila, Philip Greenland, M. Geoffrey Hayes, Abel N. Kho, Warren A. Kibbe, Amy A. Lemke, William L. Lowe, Maureen E. Smith, Wendy A. Wolf, Jennifer A. Pacheco, William K. Thompson, Joel Humowiecki, May Law, Laura Rasmussen-Torvik.

Mayo Clinic, Rochester, MN: Christopher Chute, Iftikar Kullo, Barbara Koenig, Mariza de Andrade, Suzette Bielinski, Jyotishman Pathak, Guergana Savova, Joel Wu, Joan Henriksen, Keyue Ding, Lacey Hart, Jeremy Palbicki.

Group Health/University of Washington/Fred Hutchinson Cancer Research Center, Seattle, WA: Eric B. Larson, Katherine Newton, Evette Ludman, Leslie Spangler, Gene Hart, David Carrell, Gail Jarvik, Paul Crane, Wylie Burke, Stephanie Malia Fullerton, Susan Brown Trinidad, Chris Carlson, Andrew McDavid.

Vanderbilt University, Nashville, TN: Dan M. Roden, Ellen Clayton, Jonathan L. Haines, Daniel R. Masys, Larry R. Churchill, Daniel Cornfield, Dana Crawford, Dawood Darbar, Joshua C. Denny, Bradley A Malin, Marylyn D. Ritchie, Jonathan S. Schildcrout, Hua Xu, Andrea Havens Ramirez, Melissa Basford, Jill Pulley.

Members of the DIAGRAM Consortium are:

Benjamin F. Voight[1–3,100], Laura J. Scott[4,100], Valgerdur Steinthorsdottir[5,100], Andrew P. Morris[6,100], Christian Dina[7,8,100], Ryan P. Welch[9], Eleftheria Zeggini[6,10], Cornelia Huth[11,12], Yurii S. Aulchenko[13], Gudmar Thorleifsson[5], Laura J. Mcculloch[14], Teresa Ferreira[6], Harald Grallert[11,12], Najaf Amin[13], Guanming Wu[15], Cristen J. Willer[4], Soumya Raychaudhuri[1,2,16], Steve A. Mccarroll[1,17], Claudia Langenberg[18], Oliver M. Hofmann[19], Josée Dupuis[20,21], Lu Qi[22–24], Ayellet V. Segrè[1,2,17], Mandy van Hoek[25], Pau Navarro[26], Kristin Ardlie[1], Beverley Balkau[27,28], Rafn Benediktsson[29,30], Amanda J. Bennett[14], Roza Blagieva[31], Eric Boerwinkle[32], Lori L. Bonnycastle[33], Kristina Bengtsson Boström[34], Bert Bravenboer[35], Suzannah Bumpstead[10], Noisël P. Burtt[1], Guillaume Charpentier[36], Peter S. Chines[33], Marilyn Cornelis[24], David J. Couper[37], Gabe Crawford[1,] Alex S. F. Doney[38,39], Katherine S. Elliott[6], Amanda L. Elliott[1,17,40], Michael R. Erdos[33], Caroline S. Fox[21,41], Christopher S. Franklin[42], Martha Ganser[4], Christian Gieger[11], Niels Grarup[43], Todd Green[1,2], Simon Griffin[18], Christopher J. Groves[14], Candace Guiducci[1], Samy Hadjadj[44], Neelam Hassanali[14], Christian Herder[45], Bo Isomaa[46,47], Anne U. Jackson[4], Paul R. V. Johnson[48], Torben Jørgensen[49,50], Wen H. L. Kao[51,52], Norman Klopp[11], Augustine Kong[5], Peter Kraft[22,23], Johanna Kuusisto[53], Torsten Lauritzen[54], Man Li[51], Aloysius Lieverse[55], Cecilia M. Lindgren[6,] Valeriya Lyssenko[56], Michel Marre[57,58], Thomas Meitinger[59,60], Kristian Midthjell[61], Mario A. Morken[33], Narisu Narisu[33], Peter Nilsson[56], Katharine R. Owen[14], Felicity Payne[10], John R. B. Perry[62,63], Ann-Kristin Petersen[11], Carl Platou[61], Christine Proença[7,] Inga Prokopenko[6,14], Wolfgang Rathmann[64], N. William Rayner[6,14], Neil R. Robertson[6,14], Ghislain Rocheleau[65–67], Michael Roden[45,68], Michael J. Sampson[69], Richa Saxena[1,2,40], Beverley M. Shields[62,63], Peter Shrader[3,70], Gunnar Sigurdsson[29,30], Thomas Sparsø[43], Klaus Strassburger[64], Heather M. Stringham[4], Qi Sun[22,23], Amy J. Swift[33], Barbara Thorand[11], Jean Tichet[71], Tiinamaija Tuomi[46,72], Rob M. van Dam[24], Timon W. van Haeften[73], Thijs van Herpt[25,55], Jana V. van Vliet-Ostaptchouk[74], G. Bragi Walters[5], Michael N Weedon[62,63], Cisca Wijmenga[75], Jacqueline Witteman[13], the MAGIC investigators[99], the GIANT consortium[99], Richard N. Bergman[76], Stephane Cauchi[7], Francis S. Collins[77], Anna L. Gloyn[14], Ulf Gyllensten[78], Torben Hansen[43,79], Winston A. Hide[19], Graham A. Hitman[80], Albert Hofman[13], David J. Hunter[22,23], Kristian Hveem[61,81], Markku laakso[53], Karen L. Mohlke[82], Andrew D. Morris[38,39], Colin N. A. Palmer[38,39], Peter P. Pramstaller[83], Igor Rudan[42,84,85], Eric Sijbrands[25], Lincoln D. Stein[15], Jaakko Tuomilehto[86–88], Andre Uitterlinden[25], Mark Walker[89], Nicholas J. Wareham[18], Richard M. Watanabe[76,90,] Gonçalo R. Abecasis[4], Bernhard O. Boehm[31], Harry Campbell[42], Mark J. Daly[1,2], Andrew T. Hattersley[62,63], Frank B. Hu[22–24], James B. Meigs[3,70], James S. Pankow[91], Oluf Pedersen[43,92,93], H-Erich Wichmann[11,12,94], Inês Barroso[10], Jose C. Florez[1–3,95], Timothy M. Frayling[62,63], Leif Groop[56,72], Rob Sladek[65–67], Unnur Thorsteinsdottir[5,96], James F Wilson[42], Thomas Illig[11], Philippe Froguel[17,97], Cornelia M. van duijn[13], Kari Stefansson[5,96], David Altshuler 1–3,17,40,95, Michael Boehnke[4] & Mark I. Mccarthy[6,14,98]

1 Broad Institute of Harvard and Massachusetts Institute of Technology (MIT), Cambridge, Massachusetts, USA.

2 Center for Human Genetic Research, Massachusetts General Hospital, Boston, Massachusetts, USA.

3 Department of Medicine, Harvard Medical School, Boston, Massachusetts, USA.

4 Department of Biostatistics, University of Michigan, Ann Arbor, Michigan, USA.

5 deCODE Genetics, Reykjavik, Iceland.

6 Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK.

7 CNRS-UMR-8090, Institute of Biology and Lille 2 University, Pasteur Institute, Lille, France.

8 INSERM UMR915 CNRS ERL3147, Nantes, France.

9 Bioinformatics Program, University of Michigan, Ann Arbor, Michigan, USA.

10 Wellcome Trust Sanger Institute, Hinxton, UK.

11 Institute of Epidemiology, Helmholtz Zentrum Muenchen, Neuherberg, Germany.

12 Institute of Medical Informatics, Biometry and Epidemiology, Ludwig-Maximilians-Universität, Munich, Germany.

13 Department of Epidemiology, Erasmus University Medical Center, Rotterdam, The Netherlands.

14 Oxford Centre for Diabetes, Endocrinology and Metabolism, University of Oxford, Oxford, UK.

15 Ontario Institute for Cancer Research, Toronto, Ontario, Canada.

16 Division of Rheumatology, Immunology and Allergy, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts, USA.

17 Department of Molecular Biology, Harvard Medical School, Boston, Massachusetts, USA.

18 Medical Research Council (MRC) Epidemiology Unit, Institute of Metabolic Science, Addenbrooke's Hospital, Cambridge, UK.

19 Department of Biostatistics, Harvard School of Public Health, Boston, Massachusetts, USA.

20 Department of Biostatistics, Boston University School of Public Health, Boston, Massachusetts, USA.

21 National Heart, Lung, and Blood Institute's Framingham Heart Study, Framingham, Massachusetts, USA.

22 Department of Nutrition, Harvard School of Public Health, Boston, Massachusetts, USA.

23 Department of Epidemiology, Harvard School of Public Health, Boston, Massachusetts, USA.

24 Channing Laboratory, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts, USA.

25 Department of Internal Medicine, Erasmus University Medical Centre, Rotterdam, The Netherlands.

26 MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine, Western General Hospital, Edinburgh, UK.

27 INSERM, CESP Centre for Research in Epidemiology and Population Health, U1018, Epidemiology of Diabetes, Obesity and Chronic Kidney Disease over the Lifecourse, Villejuif, France.

28 University Paris-Sud 11, UMRS 1018, Villejuif, France.

29 Landspitali University Hospital, Reykjavik, Iceland.

30 Icelandic Heart Association, Kopavogur, Iceland.

31 Division of Endocrinology, Diabetes and Metabolism, Ulm University, Ulm, Germany.

32 The Human Genetics Center and Institute of Molecular Medicine, University of Texas Health Science Center, Houston, Texas, USA.

33 National Human Genome Research Institute, National Institute of Health, Bethesda, Maryland, USA.

34 Research and Development Centre, Skaraborg Primary Care, Skövde, Sweden.

35 Department of Internal Medicine, Catharina Hospital, Eindhoven, The Netherlands.

36 Endocrinology-Diabetology Unit, Corbeil-Essonnes Hospital, Corbeil-Essonnes, France.

37 Collaborative Studies Coordinating Center, Department of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA.

38 Diabetes Research Centre, Biomedical Research Institute, University of Dundee, Ninewells Hospital, Dundee, UK.

39 Pharmacogenomics Centre, Biomedical Research Institute, University of Dundee, Ninewells Hospital, Dundee, UK.

40 Department of Genetics, Harvard Medical School, Boston, Massachusetts, USA.

41 Division of Endocrinology, Diabetes, and Hypertension, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts, USA.

42 Centre for Population Health Sciences, University of Edinburgh, Edinburgh, UK.

43 Hagedorn Research Institute, Gentofte, Denmark.

44 Centre Hospitalier Universitaire de Poitiers, Endocrinologie Diabetologie, CIC INSERM 0801, INSERM U927, Université de Poitiers, UFR, Médecine Pharmacie, Poitiers Cedex, France.

45 Institute for Clinical Diabetology, German Diabetes Center, Leibniz Center for Diabetes Research at Heinrich Heine University Düsseldorf, Düsseldorf, Germany.

46 Folkhälsan Research Center, Helsinki, Finland.

47 Malmska Municipal Health Center and Hospital, Jakobstad, Finland.

48 Diabetes Research and Wellness Foundation Human Islet Isolation Facility and Oxford Islet Transplant Programme, University of Oxford, Oxford, UK.

49 Research Centre for Prevention and Health, Glostrup University Hospital, Glostrup, Denmark.

50 Faculty of Health Science, University of Copenhagen, Copenhagen, Denmark.

51 Department of Epidemiology, Johns Hopkins University, Baltimore, Maryland, USA.

52 Department of Medicine and Welch Center for Prevention, Epidemiology and Clinical Research, Johns Hopkins University, Baltimore, Maryland, USA.

53 Department of Medicine, University of Kuopio and Kuopio University Hospital, Kuopio, Finland.

54 Department of General Medical Practice, University of Aarhus, Aarhus, Denmark.

55 Department of Internal Medicine, Maxima Medical Center, Eindhoven, The Netherlands.

56 Department of Clinical Sciences, Diabetes and Endocrinology Research Unit, University Hospital Malmö, Lund University, Malmö, Sweden.

57 Department of Endocrinology, Diabetology and Nutrition, Bichat-Claude Bernard University Hospital, Assistance Publique des Hôpitaux de Paris, Paris, France.

58 INSERM U695, Université Paris 7, Paris, France.

59 Institute of Human Genetics, Helmholtz Zentrum Muenchen, Neuherberg, Germany.

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

61 Nord-Trøndelag Health Study (HUNT) Research Center, Department of Community Medicine and General Practice, Norwegian University of Science and Technology, Trondheim, Norway.

62 Genetics of Complex Traits, Institute of Biomedical and Clinical Science, Peninsula Medical School, University of Exeter, Exeter, UK.

63 Diabetes Genetics, Institute of Biomedical and Clinical Science, Peninsula Medical School, University of Exeter, Exeter, UK.

64 Institute of Biometrics and Epidemiology, German Diabetes Center, Leibniz Center for Diabetes Research at Heinrich Heine University Düsseldorf, Düsseldorf, Germany.

65 Department of Human Genetics, McGill University, Montreal, Canada.

66 Department of Medicine, Faculty of Medicine, McGill University, Montreal, Canada.

67 McGill University and Genome Quebec Innovation Centre, Montreal, Canada.

68 Department of Metabolic Diseases, Heinrich Heine University Düsseldorf, Düsseldorf, Germany.

69 Department of Endocrinology and Diabetes, Norfolk and Norwich University Hospital National Health Service Trust, Norwich, UK.

70 General Medicine Division, Massachusetts General Hospital, Boston, Massachusetts, USA.

71 Institut interrégional pour la Santé (IRSA), La Riche, France.

72 Department of Medicine, Helsinki University Hospital, University of Helsinki, Helsinki, Finland.

73 Department of Internal Medicine, University Medical Center Utrecht, Utrecht, The Netherlands.

74 Molecular Genetics, Medical Biology Section, Department of Pathology and Medical Biology, University Medical Center Groningen and University of Groningen, Groningen, The Netherlands.

75 Department of Genetics, University Medical Center Groningen and University of Groningen, Groningen, The Netherlands.

76 Department of Physiology and Biophysics, University of Southern California School of Medicine, Los Angeles, California, USA.

77 National Institute of Health, Bethesda, Maryland, USA.

78 Department of Genetics and Pathology, Rudbeck Laboratory, Uppsala University, Uppsala, Sweden.

79 University of Southern Denmark, Odense, Denmark.

80 Centre for Diabetes, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, UK.

81 Department of Medicine, The Hospital of Levanger, Levanger, Norway.

82 Department of Genetics, University of North Carolina, Chapel Hill, North Carolina, USA.

83 Institute of Genetic Medicine, European Academy Bozen/Bolzano (EURAC), Bolzano, Italy.

84 Croatian Centre for Global Health, Faculty of Medicine, University of Split, Split, Croatia.

85 Institute for Clinical Medical Research, University Hospital 'Sestre Milosrdnice', Zagreb, Croatia.

86 Department of Public Health, University of Helsinki, Helsinki, Finland.

87 South Ostrobothnia Central Hospital, Seinäjoki, Finland.

88 Red RECAVA Grupo RD06/0014/0015, Hospital Universitario La Paz, Madrid, Spain.

89 Diabetes Research Group, Institute of Cellular Medicine, Newcastle University, Newcastle upon Tyne, UK.

90 Department of Preventative Medicine, Keck Medical School, University of Southern California, Los Angeles, California, USA.

91 Division of Epidemiology and Community Health, University of Minnesota, Minneapolis, Minnesota, USA.

92 Department of Biomedical Science, Panum, Faculty of Health Science, University of Copenhagen, Copenhagen, Denmark.

93 Faculty of Health Science, University of Aarhus, Aarhus, Denmark.

94 Klinikum Grosshadern, Munich, Germany.

95 Diabetes Unit, Massachusetts General Hospital, Boston, Massachusetts, USA.

96 Faculty of Medicine, University of Iceland, Reykjavík, Iceland.

97 Genomic Medicine, Imperial College London, Hammersmith Hospital, London, UK.

98 Oxford National Institute for Health Research Biomedical Research Centre, Churchill Hospital, Oxford, UK.

99 A full list of members is provided in the supplementary Note of the original publication.

100 These authors contributed equally

Members of the MuTHER Consortium are:

Kourosh R. Ahmadi[1], Chrysanthi Ainali[2], Amy Barrett[3], Veronique Bataille[1], Jordana T. Bell[1,4], Alfonso Buil[5], Panos Deloukas[6], Emmanouil T. Dermitzakis[5], Antigone S. Dimas[4,5], Richard Durbin[6], Daniel Glass[1], Elin Grundberg[1,6], Neelam Hassanali[3], Åsa K. Hedman[4], Catherine Ingle[6], David Knowles[7], Maria Krestyaninova[8], Cecilia M. Lindgren[4], Christopher E. Lowe[9,10], Mark I. McCarthy[3,4,11], Eshwar Meduri [1,6], Paola di Meglio[12], Josine L. Min[4], Stephen B. Montgomery[5], Frank O. Nestle[12], Alexandra C. Nica[5], James Nisbet[6], Stephen O'Rahilly[9,10], Leopold Parts[6], Simon Potter[6], Magdalena Sekowska[6], So-Youn Shin[6], Kerrin S. Small[1,6], Nicole Soranzo[1,6], Tim D. Spector[1], Gabriela Surdulescu[1], Mary E. Travers[3], Loukia Tsaprouni[6], Sophia Tsoka[2], Alicja Wilk[6], Tsun-Po Yang[6], Krina T. Zondervan[4]

1. Department of Twin Research and Genetic Epidemiology, King's College London, London, UK
2. Department of Informatics, School of Natural and Mathematical Sciences, King's College London, Strand, London, UK
3. Oxford Centre for Diabetes, Endocrinology & Metabolism, University of Oxford, Churchill Hospital, Oxford, UK
4. Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK
5. Department of Genetic Medicine and Development, University of Geneva Medical School, Geneva, Switzerland
6. Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, UK
7. University of Cambridge, Cambridge, UK
8. European Bioinformatics Institute, Hinxton, UK
9. University of Cambridge Metabolic Research Labs, Institute of Metabolic Science Addenbrooke's Hospital Cambridge, UK
10. Cambridge NIHR Biomedical Research Centre, Addenbrooke's Hospital, Cambridge, UK
11. Oxford NIHR Biomedical Research Centre, Churchill Hospital, Oxford, UK
12. St. John's Institute of Dermatology, King's College London, London, UK

Author Contributions

Conceived and designed the experiments: MCYN DS XG JLiu LRY MAN LJRT YVS SRP LW BMS CDL KLK BIF LJR HAT DC JSP EB AA GC ABS ABZ EKK DSS YLi LFB MAP EPB OG QC WZhe WJB WLL DCC SSR MGH XOS RJFL IBB PAP SRC BMP MF MKE DMB WHLK JGW JIR MMS SL CNR DWB. Performed the experiments: MAN NDP AA AD ABS YLi EPB JIR. Analyzed the data: MCYN DS BHC JLi WMC XG JLiu LRY MAN MEC RAJ DSE YVS PA SRP YLu JLo LLA LY BMS NDP PM CDL RAM DV ABS BM KR WCH WZha AK OG JAP EG MGH RJFL. Contributed reagents/materials/analysis tools: SJB LW BIF JCM UN LJR MOG YDIC HAT AC MS DC AA AD ABZ RPI JRS LFB EPB QC WZhe WLL DCC XOS PAP BMP SKI MKE JGW JIR MMS SL CNR DWB. Wrote the paper: MCYN DS.

References

1. Centers for Disease Control and Prevention (2011) National diabetes fact sheet: National estimates and general information on diabetes and prediabetes in the United States. Atlanta, GA: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention.
2. McCarthy MI (2010) Genomics, type 2 diabetes, and obesity. N Engl J Med 363: 2339–2350.
3. Voight BF, Scott LJ, Steinthorsdottir V, Morris AP, Dina C, et al. (2010) Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. Nat Genet 42: 579–589.
4. Kooner JS, Saleheen D, Sim X, Sehmi J, Zhang W, et al. (2011) Genome-wide association study in individuals of South Asian ancestry identifies six new type 2 diabetes susceptibility loci. Nat Genet: 984–989.
5. Cho YS, Chen CH, Hu C, Long J, Hee Ong RT, et al. (2011) Meta-analysis of genome-wide association studies identifies eight new loci for type 2 diabetes in East Asians. Nat Genet 44: 67–72.
6. Palmer ND, McDonough CW, Hicks PJ, Roh BH, Wing MR, et al. (2012) A genome-wide association search for type 2 diabetes genes in African Americans. PLoS ONE 7: e29202.
7. Morris AP, Voight BF, Teslovich TM, Ferreira T, Segre AV, et al. (2012) Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. Nat Genet 44: 981–990.
8. DIAbetes Genetics Replication And Meta-analysis (DIAGRAM) Consortium, Asian Genetic Epidemiology Network Type 2 Diabetes (AGEN-T2D) Consortium, South Asian Type 2 Diabetes (SAT2D) Consortium, Mexican American Type 2 Diabetes (MAT2D) Consortium, Type 2 Diabetes Genetic Exploration by Next-generation sequencing in multi-Ethnic Samples (T2D-GENES) Consortium (2014) Genome-wide trans-ancestry meta-analysis provides insight into the genetic architecture of type 2 diabetes susceptibility. Nat Genet 46: 234–244.
9. Rotimi C, Cooper R, Cao G, Sundarum C, McGee D (1994) Familial aggregation of cardiovascular diseases in African-American pedigrees. Genet Epidemiol 11: 397–407.
10. Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, et al. (2012) An integrated map of genetic variation from 1,092 human genomes. Nature 491: 56–65.
11. Devlin B, Roeder K (1999) Genomic control for association studies. Biometrics 55: 997–1004.
12. Willer CJ, Li Y, Abecasis GR (2010) METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics 26: 2190–2191.
13. Hara K, Fujita H, Johnson TA, Yamauchi T, Yasuda K, et al. (2014) Genome-wide association study identifies three novel loci for type 2 diabetes. Hum Mol Genet 23: 239–246.
14. Genomes Project Consortium (2010) A map of human genome variation from population-scale sequencing. Nature 467: 1061–1073.
15. Unoki H, Takahashi A, Kawaguchi T, Hara K, Horikoshi M, et al. (2008) SNPs in *KCNQ1* are associated with susceptibility to type 2 diabetes in East Asian and European populations. Nat Genet 40: 1098–1102.
16. Yasuda K, Miyake K, Horikawa Y, Hara K, Osawa H, et al. (2008) Variants in *KCNQ1* are associated with susceptibility to type 2 diabetes mellitus. Nat Genet 40: 1092–1097.
17. Takeuchi F, Serizawa M, Yamamoto K, Fujisawa T, Nakashima E, et al. (2009) Confirmation of multiple risk loci and genetic impacts by a genome-wide association study of type 2 diabetes in the Japanese population. Diabetes 58: 1690–1699.
18. Parra EJ, Below JE, Krithika S, Valladares A, Barta JL, et al. (2011) Genome-wide association study of type 2 diabetes in a sample from Mexico City and a meta-analysis of a Mexican-American sample from Starr County, Texas. Diabetologia 54: 2038–2046.
19. Cui B, Zhu X, Xu M, Guo T, Zhu D, et al. (2011) A genome-wide association study confirms previously reported loci for type 2 diabetes in Han Chinese. PLoS ONE 6: e22353.
20. Tsai FJ, Yang CF, Chen CC, Chuang LM, Lu CH, et al. (2010) A genome-wide association study identifies susceptibility variants for type 2 diabetes in Han Chinese. PLoS Genet 6: e1000847.
21. Stoy J, Steiner DF, Park SY, Ye H, Philipson LH, et al. (2010) Clinical and molecular genetics of neonatal diabetes due to mutations in the insulin gene. Rev Endocr Metab Disord 11: 205–215.
22. Petrik J, Pell JM, Arany E, McDonald TJ, Dean WL, et al. (1999) Overexpression of insulin-like growth factor-II in transgenic mice is associated with pancreatic islet cell hyperplasia. Endocrinology 140: 2353–2363.
23. Calderari S, Gangnerau MN, Thibault M, Meile MJ, Kassis N, et al. (2007) Defective IGF2 and IGF1R protein production in embryonic pancreas precedes beta cell mass anomaly in the Goto-Kakizaki rat model of type 2 diabetes. Diabetologia 50: 1463–1471.
24. Ramos E, Chen G, Shriner D, Doumatey A, Gerry NP, et al. (2011) Replication of genome-wide association studies (GWAS) loci for fasting plasma glucose in African-Americans. Diabetologia 54: 783–788.
25. Frayling TM, Timpson NJ, Weedon MN, Zeggini E, Freathy RM, et al. (2007) A common variant in the *FTO* gene is associated with body mass index and predisposes to childhood and adult obesity. Science 316: 889–894.
26. Hertel JK, Johansson S, Sonestedt E, Jonsson A, Lie RT, et al. (2011) *FTO*, type 2 diabetes, and weight gain throughout adult life: a meta-analysis of 41,504 subjects from the Scandinavian HUNT, MDC, and MPP studies. Diabetes 60: 1637–1644.
27. Li H, Kilpelainen TO, Liu C, Zhu J, Liu Y, et al. (2012) Association of genetic variation in *FTO* with risk of obesity and type 2 diabetes with data from 96,551 East and South Asians. Diabetologia 55: 981–995.
28. Binh TQ, Phuong PT, Nhung BT, Thoang DD, Lien HT, et al. (2013) Association of the common *FTO*-rs9939609 polymorphism with type 2 diabetes, independent of obesity-related traits in Vietnamese population. Gene 513: 31–35.
29. Dunham I, Kundaje A, Aldred SF, Collins PJ, Davis CA, et al. (2012) An integrated encyclopedia of DNA elements in the human genome. Nature 489: 57–74.
30. Ward LD, Kellis M (2012) HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. Nucleic Acids Res 40: D930–934.
31. Bell AC, Felsenfeld G (2000) Methylation of a CTCF-dependent boundary controls imprinted expression of the *Igf2* gene. Nature 405: 482–485.
32. Oshel KM, Knight JB, Cao KT, Thai MV, Olson AL (2000) Identification of a 30-base pair regulatory element and novel DNA binding protein that regulates the human *GLUT4* promoter in transgenic mice. J Biol Chem 275: 23666–23673.
33. Wellcome Trust Case Control Consortium (2007) Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 447: 661–678.
34. Barrett JC, Clayton DG, Concannon P, Akolkar B, Cooper JD, et al. (2009) Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. Nat Genet 41: 703–707.
35. Plagnol V, Howson JM, Smyth DJ, Walker N, Hafler JP, et al. (2011) Genome-wide association analysis of autoantibody positivity in type 1 diabetes cases. PLoS Genet 7: e1002216.
36. Dabelea D, Pihoker C, Talton JW, D'Agostino RB, Jr., Fujimoto W, et al. (2011) Etiological approach to characterization of diabetes type: the SEARCH for Diabetes in Youth Study. Diabetes Care 34: 1628–1633.

37. Barinas-Mitchell E, Pietropaolo S, Zhang YJ, Henderson T, Trucco M, et al. (2004) Islet cell autoimmunity in a triethnic adult population of the Third National Health and Nutrition Examination Survey. Diabetes 53: 1293–1302.
38. Hakonarson H, Grant SF, Bradfield JP, Marchand L, Kim CE, et al. (2007) A genome-wide association study identifies *KIAA0350* as a type 1 diabetes gene. Nature 448: 591–594.
39. Erlich H, Valdes AM, Noble J, Carlson JA, Varney M, et al. (2008) *HLA DR-DQ* haplotypes and genotypes and type 1 diabetes risk: analysis of the type 1 diabetes genetics consortium families. Diabetes 57: 1084–1092.
40. Howson JM, Walker NM, Clayton D, Todd JA (2009) Confirmation of *HLA* class II independent type 1 diabetes associations in the major histocompatibility complex including *HLA-B* and *HLA-A*. Diabetes Obes Metab 11 Suppl 1: 31–45.
41. Eike MC, Becker T, Humphreys K, Olsson M, Lie BA (2009) Conditional analyses on the T1DGC *MHC* dataset: novel associations with type 1 diabetes around *HLA-G* and confirmation of *HLA-B*. Genes Immun 10: 56–67.
42. Howson JM, Roy MS, Zeitels L, Stevens H, Todd JA (2013) *HLA* class II gene associations in African American Type 1 diabetes reveal a protective *HLA-DRB1*03* haplotype. Diabet Med 30: 710–716.
43. Noble JA, Johnson J, Lane JA, Valdes AM (2013) *HLA* class II genotyping of African American type 1 diabetes patients reveals associations unique to African haplotypes. Diabetes 62: 3292–3299.
44. Odegaard JI, Chawla A (2012) Connecting type 1 and type 2 diabetes through innate immunity. Cold Spring Harb Perspect Med 2: a007724.
45. Rich SS, French LR, Sprafka JM, Clements JP, Goetz FC (1993) *HLA*-associated susceptibility to type 2 (non-insulin-dependent) diabetes mellitus: the Wadena City Health Study. Diabetologia 36: 234–238.
46. Cervin C, Lyssenko V, Bakhtadze E, Lindholm E, Nilsson P, et al. (2008) Genetic similarities between latent autoimmune diabetes in adults, type 1 diabetes, and type 2 diabetes. Diabetes 57: 1433–1437.
47. Raj SM, Howson JM, Walker NM, Cooper JD, Smyth DJ, et al. (2009) No association of multiple type 2 diabetes loci with type 1 diabetes. Diabetologia 52: 2109–2116.
48. Lukacs K, Hosszufalusi N, Dinya E, Bakacs M, Madacsy L, et al. (2012) The type 2 diabetes-associated variant in *TCF7L2* is associated with latent autoimmune diabetes in adult Europeans and the gene effect is modified by obesity: a meta-analysis and an individual study. Diabetologia 55: 689–693.
49. Li Y, Willer CJ, Ding J, Scheet P, Abecasis GR (2010) MaCH: using sequence and genotype data to estimate haplotypes and unobserved genotypes. Genet Epidemiol 34: 816–834.
50. Howie BN, Donnelly P, Marchini J (2009) A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. PLoS Genet 5: e1000529.
51. Browning SR, Browning BL (2007) Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. Am J Hum Genet 81: 1084–1097.
52. Chen MH, Yang Q (2010) GWAF: an R package for genome-wide association analyses with family data. Bioinformatics 26: 580–581.
53. Almasy L, Blangero J (1998) Multipoint quantitative-trait linkage analysis in general pedigrees. Am J Hum Genet 62: 1198–1211.
54. Skol AD, Scott LJ, Abecasis GR, Boehnke M (2006) Joint analysis is more efficient than replication-based analysis for two-stage genome-wide association studies. Nat Genet 38: 209–213.
55. Magi R, Morris AP (2010) GWAMA: software for genome-wide association meta-analysis. BMC Bioinformatics 11: 288.
56. Lin S, Chakravarti A, Cutler DJ (2004) Exhaustive allelic transmission disequilibrium tests as a new approach to genome-wide association studies. Nat Genet 36: 1181–1188.
57. Wray NR, Yang J, Goddard ME, Visscher PM (2010) The genetic interpretation of area under the ROC curve in genomic profiling. PLoS Genet 6: e1000864.
58. Barrett JC, Fry B, Maller J, Daly MJ (2005) Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 21: 263–265.
59. Purcell S, Cherny SS, Sham PC (2003) Genetic Power Calculator: design of linkage and association genetic mapping studies of complex traits. Bioinformatics 19: 149–150.
60. Nica AC, Parts L, Glass D, Nisbet J, Barrett A, et al. (2011) The architecture of gene regulatory variation across multiple human tissues: the MuTHER study. PLoS Genet 7: e1002003.
61. Aulchenko YS, Ripke S, Isaacs A, van Duijn CM (2007) GenABEL: an R library for genome-wide association analysis. Bioinformatics 23: 1294–1296.
62. Aulchenko YS, Struchalin MV, van Duijn CM (2010) ProbABEL package for genome-wide association analysis of imputed data. BMC Bioinformatics 11: 134.
63. Stranger BE, Forrest MS, Dunning M, Ingle CE, Beazley C, et al. (2007) Relative impact of nucleotide and copy number variation on gene expression phenotypes. Science 315: 848–853.
64. Boyle AP, Hong EL, Hariharan M, Cheng Y, Schaub MA, et al. (2012) Annotation of functional variation in personal genomes using RegulomeDB. Genome Res 22: 1790–1797.
65. Wellcome Trust Case Control Consortium (2007) Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 447: 661–678.