View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Elsevier - Publisher Connector

ARTICLE

Meta-analysis of Dense Genecentric Association Studies Reveals Common and Uncommon Variants Associated with Height

Matthew B. Lanktree,[1,115] Yiran Guo,[2,3,115] Muhammed Murtaza,[4,66] Joseph T. Glessner,[2] Swneke D. Bailey,[6] N. Charlotte Onland-Moret,[21] Guillaume Lettre,[5] Halit Ongen,[8] Ramakrishnan Rajagopalan,[10] Toby Johnson,[9] Haiqing Shen,[11] Christopher P. Nelson,[15,86] Norman Klopp,[12] Jens Baumert,[12] Sandosh Padmanabhan,[54] Nathan Pankratz,[24,83] James S. Pankow,[83] Sonia Shah,[87] Kira Taylor,[13] John Barnard,[14] Bas J. Peters,[108] Cliona M. Maloney,[30] Maximilian T. Lobmeyer,[16] Alice Stanton,[58] M. Hadi Zafarmand,[18,109] Simon P.R. Romaine,[23] Amar Mehta,[25] Erik P.A. van Iperen,[22,82] Yan Gong,[16] Tom S. Price,[20] Erin N. Smith,[31] Cecilia E. Kim,[2] Yun R. Li,[2] Folkert W. Asselbergs,[18,21,109] Larry D. Atwood,[35] Kristian M. Bailey,[23] Deepak Bhatt,[99] Florianne Bauer,[21] Elijah R. Behr,[45] Tushar Bhangale,[43] Jolanda M.A. Boer,[28] Bernhard O. Boehm,[92] Jonathan P. Bradfield,[2] Morris Brown,[95] Peter S. Braund,[15,86] Paul R. Burton,[32] Cara Carty,[19] Hareesh R. Chandrupatla,[29] Wei Chen,[105] John Connell,[38] Chrysoula Dalgeorgou,[46] Anthonius de Boer,[108] Fotios Drenos,[27] Clara C. Elbers,[21] James C. Fang,[51] Caroline S. Fox,[35] Edward C. Frackelton,[2] Barry Fuchs,[36] Clement E. Furlong,[10] Quince Gibson,[11] Christian Gieger,[12] Anuj Goel,[8,72] Diederik E. Grobbee,[104] Claire Hastie,[54] Philip J. Howard,[9] Guan-Hua Huang,[52] W. Craig Johnson,[34] Qing Li,[111] Marcus E. Kleber,[88] Barbara E.K. Klein,[17] Ronald Klein,[17] Charles Kooperberg,[19] Bonnie Ky,[50] Andrea LaCroix,[19] Paul Lanken,[36] Mark Lathrop,[96] Mingyao Li,[29] Vanessa Marshall,[94] Olle Melander,[55] Frank D. Mentch,[2] Nuala J. Meyer,[36] Keri L. Monda,[40] Alexandre Montpetit,[42] Gurunathan Murugesan,[33] Karen Nakayama,[10] Dave Nondahl,[17] Abiodun Onipinla,[9] Suzanne Rafelt,[15,86] Stephen J. Newhouse,[9] F. George Otieno,[2] Sanjey R. Patel,[41] Mary E. Putt,[102] Santiago Rodriguez,[53] Radwan N. Safa,[49] Douglas B. Sawyer,[48] Pamela J. Schreiner,[39] Claire Simpson,[111] Suthesh Sivapalaratnam,[26] Sathanur R. Srinivasan,[105] Christine Suver,[30] Gary Swergold,[112] Nancy K. Sweitzer,[47] Kelly A. Thomas,[2] Barbara Thorand,[12] Nicholas J. Timpson,[53]

[1]Department of Medicine and Biochemistry, University of Western Ontario, London, Ontario, N6A 5C1, Canada; [2]Center for Applied Genomics, Abramson Research Center, The Children's Hospital of Philadelphia, Philadelphia, PA 19104, USA; [3]Beijing Genomics Institute at Shenzhen, Shenzhen, China; [4]Wellcome Trust Sanger Institute, Hinxton, Cambridge CB10 1SA, UK; [5]Montréal Heart Institute, Université de Montréal, Montréal, Québec, H1T 1C8, Canada; [6]Department of Human Genetics, McGill University, Montréal, Québec, H3A 1B1, Canada; [7]Genetic Epidemiology Group, Department of Epidemiology and Public Health, University College London, London WC1E 6BT, UK; [8]The Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford OX3 7BN, UK; [9]Clinical Pharmacology and Barts and the London Genome,Centre, William Harvey Research Institute, Barts and the London School of Medicine, Queen Mary University of London, London EC1M 6BQ, UK; [10]Department of Medicine, Division of Medical Genetics, University of Washington, Seattle, WA, 98195, USA; [11]Division of Endocrinology, Diabetes and Nutrition, University of Maryland School of Medicine, Baltimore, MD, 21201, USA; [12]Institute of Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg 85764, Germany; [13]Department of Epidemiology, School of Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC 27514, USA; [14]Department of Quantitative Health Sciences, The Cleveland Clinic Foundation, 9500 Euclid Avenue, Cleveland, OH 44195, USA; [15]Department of Cardiovascular Sciences, University of Leicester, Glenfield Hospital, Leicester LE3 9QP, UK; [16]Center for Pharmacogenomics, College of Pharmacy, University of Florida, FL 32610 USA; [17]Department of Ophthalmology and Visual Sciences, School of Medicine and Public Health, University of Wisconsin-Madison, Madison, WI 53705 USA; [18]Department of Cardiology, Division Heart and Lungs, University Medical Center Utrecht, Utrecht, The Netherlands; [19]Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA 98109, USA; [20]MRC SGDP Centre, Institute of Psychiatry, London SE5 8AF, UK; [21]Complex Genetics Section, Department of Medical Genetics (DBG) University Medical Center Utrecht, Utrecht STR 6, The Netherlands; [22]Durrer Center for Cardiogenetic Research, Amsterdam, The Netherlands; [23]Leeds Institute of Genetics Health & Therapeutics, University of Leeds, Leeds LS2 9JT, UK; [24]Department of Medical and Molecular Genetics, Indiana University, School of Medicine, 410 West 10th Street, HS4000, Indianapolis, IN 46202, USA; [25]Department of Environmental Health, Environmental and Occupational Medicine and Epidemiology Program, Harvard School of Public Health, Boston, MA 02115 USA; [26]Department of Cardiology and Vascular Medicine, Academic Medical Center, 1105 AZ Amsterdam; [27]Centre for Cardiovascular Genetics, Department of Medicine, University College London, 5 University Street, London WC1E 6JF, UK; [28]National Institute for Public Health and the Environment (RIVM), 3720 BA Bilthoven, The Netherlands; [29]Cardiovascular Institute, University of Pennsylvania School of Medicine, Philadelphia, PA 19104, USA; [30]Department of Genetics, Rosetta Inpharmatics, Seattle, WA 98109-5234, USA; [31]Scripps Genomic Medicine and Scripps Translational Science Institute, 3344 N. Torrey Pines Ct. Ste 300, La Jolla, CA 92037, USA; [32]Department of Health Sciences, University of Leicester, Adrian Building, University Rd., Leicester LE1 7RH, UK; [33]Department of Clinical Pathology, Pathology and Laboratory Medicine Institute, Cleveland Clinic, Cleveland, OH 44106, USA; [34]Department of Biostatistics, University of Washington, Seattle, WA 98195 USA; [35]Framingham Heart Study, Boston University School of Medicine, Boston, MA 02118-2526, USA; [36]University of Pennsylvania Medical Center, Pulmonary, Allergy & Critical Care Division, Philadelphia, PA 19104-6160, USA; [37]Department of Epidemiology, Graduate School of Public Health, University of Pittsburgh, 130 DeSoto St., Pittsburgh, PA 15261, USA; [38]University of Dundee, Medical School, Ninewells Hospital and Medical School, DD1 9SY Dundee, UK; [39]Division of Epidemiology & Community Health, University of Minnesota, Minneapolis, MN 55454, USA; [40]Department of Genetics, University of North Carolina School of Medicine at Chapel Hill, Chapel Hill, NC 27514, USA; [41]Harvard Medical School, Brigham and Women's Hospital and Beth Israel Deaconess Medical Center, Boston, MA 02215, USA; [42]McGill University and Genome Québec Innovation Centre, Montréal, Québec H3A 1A4 Canada; [43]Department of Bioinformatics and Computational Biology, Genentech Inc, South San Francisco; [44]Broad Institute of MIT and Harvard, Cambridge, MA 02115, USA; [45]Division of Cardiovascular Sciences, St George's University of London, London SW17 0RE, UK; [46]Division of Clinical Developmental Sciences, St George's University of London SW17 0RE, London, UK; [47]Cardiovascular Medicine, University of Wisconsin, Madison, WI 53792, USA; [48]Cardiovascular Division, Vanderbilt University Medical Center, Nashville, TN

Sam Tischfield,[44] Martin Tobin,[32] Maciej Tomaszweski,[15,86] W.M. Monique Verschuren,[28] Chris Wallace,[97] Bernhard Winkelmann,[93] Haitao Zhang,[2] Dongling Zheng,[46] Li Zhang,[14] Joseph M. Zmuda,[37] Robert Clarke,[107] Anthony J. Balmforth,[23] John Danesh,[65] Ian N. Day,[53] Nicholas J. Schork,[31] Paul I.W. de Bakker,[62,44,21] Christian Delles,[54] David Duggan,[59] Aroon D. Hingorani,[7,71] Joel N. Hirschhorn,[44,77,78] Marten H. Hofker,[63] Steve E. Humphries,[27] Mika Kivimaki,[7] Debbie A. Lawlor,[53] Kandice Kottke-Marchant,[100] Jessica L. Mega,[60] Braxton D. Mitchell,[11] David A. Morrow,[60] Jutta Palmen,[27] Susan Redline,[41] Denis C. Shields,[57] Alan R. Shuldiner,[11,80] Patrick M. Sleiman,[2] George Davey Smith,[53] Martin Farrall,[8,72] Yalda Jamshidi,[46] David C. Christiani,[25,81] Juan P. Casas,[110] Alistair S. Hall,[23] Pieter A. Doevendans,[18] Jason D. Christie,[36] Gerald S. Berenson,[105] Sarah S. Murray,[31] Thomas Illig,[12] Gerald W. Dorn, II,[85] Thomas P. Cappola,[50] Eric Boerwinkle,[68] Peter Sever,[101] Daniel J. Rader,[29,74] Muredach P. Reilly,[29,74] Mark Caulfield,[9] Philippa J. Talmud,[27] Eric Topol,[98] James C. Engert,[67] Kai Wang,[2] Anna Dominiczak,[56] Anders Hamsten,[106] Sean P. Curtis,[113] Roy L. Silverstein,[61] Leslie A. Lange,[40] Marc S. Sabatine,[60] Mieke Trip,[26] Danish Saleheen,[65,66] John F. Peden,[8,72] Karen J. Cruickshanks,[17,79] Winfried März,[89,90,91] Jeffrey R. O'Connell,[11] Olaf H. Klungel,[108] Cisca Wijmenga,[69] Anke Hilse Maitland-van der Zee,[108] Eric E. Schadt,[84] Julie A. Johnson,[64] Gail P. Jarvik,[10] George J. Papanicolaou,[70] Hugh Watkins on behalf of PROCARDIS,[72] Struan F.A. Grant,[2,75] Patricia B. Munroe,[9] Kari E. North,[13,76] Nilesh J. Samani,[15,86]

37232, USA; [49]Department of Molecular Medicine, Boston University, Boston, MA 02118, USA; [50]Penn Cardiovascular Institute, University of Pennsylvania School of Medicine, Philadelphia, PA 19104, USA; [51]Cardiovascular Medicine, Case Western Reserve University, Cleveland, OH 44106,USA; [52]Institute of Statistics, National Chiao Tung University, Hsinchu 30010, Taiwan; [53]MRC Centre for Causal Analyses in Translational Epidemiology, School of Social and Community Medicine, University of Bristol, Oakfield House, Oakfield Grove, Bristol BS8 2BN, UK; [54]BHF Glasgow Cardiovascular Research Centre, Division of Cardiovascular and Medical Sciences, University of Glasgow, Western Infirmary, Glasgow G12 8TA, UK; [55]Clinical Research Center (CRC), Malmö University Hospital, SE-205 02 Malmö, Sweden; [56]Institute of Cardiovascular and Medical Sciences, College of Medical, Veterinary and Life Sciences, University of Glasgow, University Place, Glasgow G12 8QQ, UK; [57]Conway Institute of Biomolecular & Biomedical Research, University College Dublin, Dublin 4, Ireland; [58]Molecular and Cellular Therapeutics, Royal College of Surgeons in Ireland, Dublin 2, Ireland; [59]Translational Genomics Research Institute, Phoenix, AZ 85004, USA; [60]TIMI Study Group, Cardiovascular Division, Brigham and Women's Hospital, Boston, MA 02115, USA; [61]Department of Cell Biology, Lerner Research Institute, Cleveland Clinic Foundation, Department of Molecular Medicine, Cleveland Clinic Lerner College of Medicine, Case Western Reserve University, 9500 Euclid Ave./NC10, Cleveland, OH 44195, USA; [62]Division of Genetics, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA; [63]Molecular Genetics, University Medical Center Groningen, Groningen University, Groningen 9700 RB, the Netherlands; [64]Department of Pharmacotherapy and Translational Research and Center for Pharmacogenomics, University of Florida College of Pharmacy, Gainesville, FL 32610, USA; [65]Department of Public Health and Primary Care, University of Cambridge, Cambridge, CB1 8RN, UK; [66]Center for Non-Communicable Diseases, Karachi, Pakistan; [67]Departments of Medicine and Human Genetics, McGill University, Montréal, Québec H3A 1B1, Canada; [68]Human Genetics Center and Div. of Epidemiology, 1200 Herman Pressler, Suite E-447, Houston, TX 77030, USA; [69]Department of Genetics, University Medical Center Groningen and Groningen University, 9700 RB Groningen, The Netherlands; [70]National Heart, Lung, and Blood Institute (NHLBI), Division of Cardiovascular Sciences, Bethesda, MD 20892, USA; [71]Centre for Clinical Pharmacology, Department of Medicine, University College London, London WC1E 6JF, UK; [72]Department of Cardiovascular Medicine, University of Oxford, Level 6 West Wing, John Radcliffe Hospital, Headley Way, Headington, Oxford OX3 9DU, UK; [73]Department of Medicine and Clinical Epidemiology and Biostatistics, Population Genomics Program, McMaster University, Hamilton Health Sciences, Hamilton General Hospital, Hamilton, Ontario L8L 2X2, Canada; [74]The Institute for Translational Medicine and Therapeutics, School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA; [75]Department of Pediatrics, University of Pennsylvania, Philadelphia PA 19104, USA; [76]Carolina Center for Genome Sciences, School of Public Health, University of North Carolina, Chapel Hill, NC 27514, USA; [77]Divisions and Endocrinology and Genetics and Program in Genomics, Children's Hospital, Boston, MA 02115, USA; [78]Department of Genetics, Harvard Medical School, Boston, MA 02115, USA; [79]Department of Population Health Sciences, School of Medicine and Public Health, University of Wisconsin-Madison, Madison, WI 53705, USA; [80]Geriatric Research and Education Clinical Center, Veterans Administration Medical Center, Baltimore, MD 21201, USA; [81]Pulmonary and Critical Care Unit, Massachusetts General Hospital, Department of Medicine, Harvard Medical School, Boston, MA 02114, USA; [82]Department of Clinical Epidemiology, Biostatistics and Bioinformatics, Academic Medical Center, University of Amsterdam, 1007 MB Amsterdam, The Netherlands; [83]Division of Epidemiology & Community Health, University of Minnesota, Minneapolis, MN 55454, USA; [84]Sage Bionetworks, Seattle, WA 98109, USA; [85]Washington University Center for Pharmacogenetics, 660 S. Euclid Ave., Campus Box 8220, St. Louis, MO 63110-1093, USA; [86]Leicester National Institute of Health Research Biomedical Research Unit in Cardiovascular Disease, Glenfield Hospital, Groby Road, Leicester LE3 9QP, UK; [87]UCL Genetic Institute, University College London, London WC1E 6BT, UK; [88]LURIC Nonprofit LLC, Freiburg, Germany; [89]Synlab Center of Laboratory Diagnostics Heidelberg, Heidelberg D-58509, Germany; [90]Institute of Public Health, Social Medicine and Epidemiology Medical Faculty, University of Heidelberg, D-68167 Mannheim, Germany; [91]Clinical Institute of Medical and Chemical Laboratory Diagnostics, Medical University of Graz, A-8036 Graz, Austria; [92]Division of Endocrinology, Diabetes and Metabolism, Centre of Excellence Baden-Wuerttemberg, Metabolic Diseases, Ulm University, D - 89081 Ulm, Germany; [93]Cardiology Group Frankfurt-Sachsenhausen, Frankfurt, Germany; [94]Drug Safety Research Unit, Southampton SO31 1AA, UK; [95]Clinical Pharmacology and the Cambridge Institute of Medical Research, University of Cambridge, Addenbrooke's Hospital, Cambridge CB2 0SP, UK; [96]Centre National de Genotypage, CP 5721, 91 057 Evry Cedex, France; [97]JDRF/WT Diabetes and Inflammation Laboratory, Cambridge Institute for Medical Research, Wellcome Trust/MRC Building, Addenbrooke's Hospital, Cambridge CB2 0XY, UK; [98]Department of Molecular and Experimental Medicine, Scripps Research Institute, La Jolla, CA 92037, US; [99]Harvard Medical School, Cambridge, MA 02115, USA; [100]Pathology and Laboratory Medicine Institute, Cleveland Clinic Cleveland, OH 44195; [101]International Centre for Circulatory Health, National Heart & Lung Institute, Imperial College London, London W2 1NY, UK; [102]Department of Biostatistics and Epidemiology, University of Pennsylvania, Philadelphia, PA 19104, USA; [103]Department of Internal Medicine II – Cardiology, University of Ulm Medical Center, Ulm Konto Nr. 5050, Germany; [104]Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, 3508 GA Utrecht, The Netherlands; [105]Department of Epidemiology, 1440 Canal Street, Suite 1829, Tulane University, New Orleans, LA 70112-2750, USA; [106]Atherosclerosis Research Unit, Department of Medicine, Karolinska Institutet SE-171 77 Stockholm, Sweden; [107]Clinical Trial Service Unit, Richard Doll Building, Old Road Campus, Roosevelt Drive, Oxford OX37LF, UK; [108]Division of Pharmacoepidemiology and Clinical Pharmacology, Utrecht Institute for Pharmaceutical Sciences, Utrecht University, 3508 TC Utrecht, The Netherlands; [109]Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, 3508 GA Utrecht, The Netherlands; [110]Department of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, WC1E 7HT, UK; [111]National Human Genome Research Institute, National Institutes of Health, 333 Cassell Drive, Suite 1200, Baltimore, MD 21224, USA; [112]Regeneron Pharmaceuticals, 777 Old Saw Mill River Road, Tarrytown, NY 10591, USA; [113]Merck Research Laboratories, P.O. Box 2000, Rahway, NJ 07065, USA

Wolfgang Koenig,[103] Tom R. Gaunt,[53] Sonia S. Anand,[73] Yvonne T. van der Schouw,[104] Meena Kumari on behalf of the Whitehall II Study and the WHII 50K Group,[7] Nicole Soranzo,[4] Garret A. FitzGerald,[74] Alex Reiner,[19] Robert A. Hegele,[1] Hakon Hakonarson,[2,75,*] and Brendan J. Keating[29,74,114,*]

Height is a classic complex trait with common variants in a growing list of genes known to contribute to the phenotype. Using a genecentric genotyping array targeted toward cardiovascular-related loci, comprising 49,320 SNPs across approximately 2000 loci, we evaluated the association of common and uncommon SNPs with adult height in 114,223 individuals from 47 studies and six ethnicities. A total of 64 loci contained a SNP associated with height at array-wide significance ($p < 2.4 \times 10^{-6}$), with 42 loci surpassing the conventional genome-wide significance threshold ($p < 5 \times 10^{-8}$). Common variants with minor allele frequencies greater than 5% were observed to be associated with height in 37 previously reported loci. In individuals of European ancestry, uncommon SNPs in IL11 and SMAD3, which would not be genotyped with the use of standard genome-wide genotyping arrays, were strongly associated with height ($p < 3 \times 10^{-11}$). Conditional analysis within associated regions revealed five additional variants associated with height independent of lead SNPs within the locus, suggesting allelic heterogeneity. Although underpowered to replicate findings from individuals of European ancestry, the direction of effect of associated variants was largely consistent in African American, South Asian, and Hispanic populations. Overall, we show that dense coverage of genes for uncommon SNPs, coupled with large-scale meta-analysis, can successfully identify additional variants associated with a common complex trait.

Introduction

Although complex processes such as age at puberty, perinatal environment, and nutritional intake affect attained adult height, up to 90% of its variation has been attributed to heritable factors.[1,2] Height is an attractive model phenotype to study in an attempt to provide insights into the genetic architecture of complex traits: it is relatively stable over the course of adulthood, it is usually measured in relation to body mass index in large-scale population-based and case-control studies, it is easily and accurately measured, and it is easily harmonized across cohorts. More than 90 years ago, Fisher proposed that many variants with small individual effects explain the heritability of continuous, normally distributed traits, such as height.[3] Recent findings from a number of genome-wide association studies (GWAS) support this hypothesis, as common variation in over 180 loci have now been associated with height,[4,5] but the common variants within the loci explain less than 10% of the population variation in height.[5–14] Recent work suggests that about 300,000 SNPs can explain up to 45% of the variance in height across the population,[15] but it provides no insight into the responsible genes. Given that all of the variants needed to explain height have not been identified to date, the additional genetic variants are likely to be uncommon in the population or of very small effect, requiring extremely large samples to be confidently identified.

Multiple genecentric genotyping arrays have been developed for replication and fine mapping of loci with known or putative roles in specific phenotypes or disease areas. One of the first such arrays, the ITMAT-Broad-CARe or "IBC array" (also referred to as the CardioChip or the Human Cardiovascular Disease [HumanCVD] BeadChip [Illumina]), incorporates about 50,000 SNPs to efficiently capture genetic diversity across over 2000 genic regions related to cardiovascular, inflammatory, and metabolic phenotypes. Genetic variation within the majority of these regions is captured at density equal to or greater than that afforded by genome-wide genotyping products.[16] The IBC array has content derived from the International HapMap Consortium and resequencing data from the SeattleSNPs and National Institute of Environmental Health Sciences (NIEHS) SNPs consortia, with a focus upon inclusion of lower-frequency variants and variants with a higher likelihood of functionality.

Using phased haplotype data from panels of densely genotyped individuals, such as those provided by the HapMap and the 1,000 Genomes Project, imputation is often performed to increase the number of queried SNPs in GWAS efforts.[17] However, SNPs that are ungenotyped or monomorphic in reference panels are not imputable, and imputation quality drops for lower-frequency variants. Until large-scale sequencing projects in large population-sized cohorts become affordable, direct genotyping of previously discovered uncommon variants is the only method available for querying the impact of uncommon moderate- to small-effect-sized variants.

Genes were selected for inclusion on the IBC array on the basis of pathway analysis and previous candidate-gene and GWAS reports of a variety of cardiovascular disease (CVD)-related phenotypes. Although the IBC array content is primarily biased toward CVD phenotypes,[16] of 87 loci reported to be associated with adult height in GWAS performed before 2010,[4] 27 are present on the array, with 20 of the known loci covered at a density equal or greater than that of conventional GWAS platforms. Additionally, many genes with plausible biological hypotheses for involvement in height without prior evidence for association are found on the IBC array, such as genes with

[114]Current address: The Children's Hospital of Philadelphia Research Institute, 3516 Civic Center Blvd, Philadelphia, Pennsylvania 19104, USA
[115]These authors contributed equally to this work.
*Correspondence: hakonarson@chop.edu (H.H.), bkeating@chop.edu (B.J.K.)
DOI 10.1016/j.ajhg.2010.11.007. ©2011 by The American Society of Human Genetics. All rights reserved.

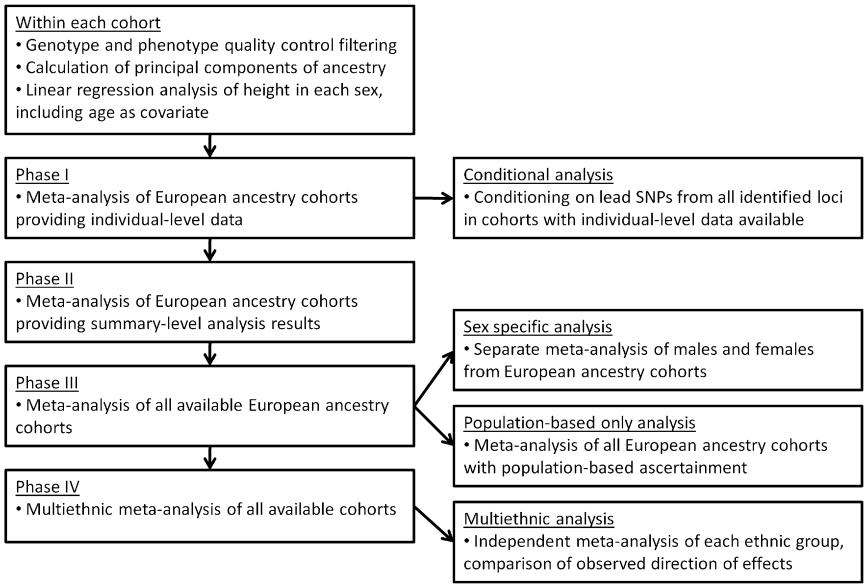

Figure 1. IBC Array Height Meta-Analysis Overview

involvement in endocrine pathways and energy metabolism.

In this study, we performed meta-analysis of 90,446 individuals of European ancestry and 23,777 individuals from an additional five ethnicities, including a total of 47 cohorts genotyped on the IBC array with available adult height data. We aimed to replicate previous genetic associations with height and to find loci not previously described to underpin this highly polygenic trait. Special attention was paid to variants of lower minor allele frequency (MAF) that would go undetected in studies relying on imputation or in studies with fewer participants. Using conditional analyses, we sought to identify multiple independent association signals from within significant loci. Finally, multiethnic meta-analysis was performed, including all available participants, and the concordance of direction of effect across ethnicities was evaluated.

Material and Methods

Participating Studies

Investigators provided either individual-level genotype data with height, age, and sex or summary-level statistics, following analysis guidelines. Data sets included population-based cohorts, collections of cases and controls for a variety of metabolic and cardiovascular phenotypes, and individuals participating in clinical trials. A detailed description of the cohorts included in this study is found in Tables S1 and S2 (available online). All participants were required to have a measured height and to have their age at time of measurement noted, and individuals younger than 21 years or older than 80 years were excluded from analysis. In total, 25 cohorts contributed individual-level phenotype and genotype data for a total of 65,574 participants, forming the individual-level phase I cohort (Figure 1, Table S1). An additional 22 cohorts contributed summary-level results, representing an additional 48,649 subjects (Table S2), creating a total sample size of 114,223. All participating studies were required to obtain informed consent for DNA analysis and to have received approval from local institutional review boards or ethics committees.

Genotyping and Quality Control

Genotyping was performed with the genecentric IBC array, of which the design and coverage compared to conventional genome-wide genotyping arrays has been described in detail elsewhere.[16] In brief, the density of tagging SNPs for over 2000 loci of interest was chosen via a priority strategy, with a "cosmopolitan tagging" approach employed to capture known variation in HapMap populations. Available resequencing data were used to identify lower-frequency variants, with emphasis on nonsynonymous SNPs and known or putative functional variants. Approximately 17,000 SNPs included on the IBC array have an MAF < 0.05 in individuals of European descent. For the majority of regions, SNPs were designed to be inclusive of the intronic, exonic, and flanking untranslated regions (UTRs), as well as to provide coverage of the proximal promoter regions designed for the higher-priority loci. Of the SNPs included on the IBC array, 65% are intronic, 9.9% are in exonic, and 7.7% are nonsynonymous. Further details regarding SNP coverage for each locus can be found in an online reference (see Web Resources).

From the IBC array, a total of 49,320 SNPs were clustered into genotypes with the use of the Illumina Beadstudio software and were subjected to quality control filters at the sample and SNP levels, separately within each cohort. Samples with individual call rates $< 90\%$, gender mismatch, or duplicate discordance were excluded. SNPs with a call rate $< 95\%$ or Hardy Weinberg Equilibrium $p < 10^{-7}$ were removed. Because of the low-frequency SNPs included in the design of the human IBC array and the large sample size of the current study, no filtering was performed on MAF.

Statistical Analyses

Evaluation of Cryptic Relatedness

Within cohorts with family structure, only founders were included in the analysis, with the exception of the GRAPHIC, Amish, and PROCARDIS studies, in which family structure was maintained and utilized in the association analysis (see Table S2). To ensure removal of cryptic relatedness and duplicate samples, calculation of pi_hat ($\widehat{\pi}$), a measure of identity by descent, was estimated from identity by state and sample allele frequencies via the method implemented in PLINK.[18] For each set of duplicates or monozygotic twins and for those with a pairwise $\widehat{\pi} > 0.3$, the sample with the highest genotyping call rate was retained for analysis.

Evaluation of Population Stratification

For the primary analysis of both individual-level and summary-level data, only individuals of European ancestry were included. Self-reported ethnicity was verified by multidimensional-scaling analysis of identity-by-state distances as implemented in PLINK, including HapMap panels as reference standards. After pruning

of SNPs in linkage disequilibrium (LD) ($r^2 > 0.3$), EIGENSOFT was used to compute principal components for use as covariates in the regression analyses.[19,20] Additional self-reported ethnicities (African American, South Asian, East Asian, Hispanic, and Native American) were then examined independently via the same methodology.

Association Testing

Within all cohorts, including those with individual-level data available and those analyzed by studies providing summary-level results, association analysis was performed with the use of linear regression, with height used as a continuous trait, and an additive genetic model, including correction for age and a minimum of the top three principal components of ancestry (described above) for individuals of European descent and ten principal components for all other ethnicities, as implemented in the software package PLINK.[18] Age was included as a covariate in regression analyses for minimization of generation effects. All analyses were performed with stratification by gender and race. In the GRAPHIC, Amish, and PROCARDIS studies, which contained some family relationships (n = 9466 total individuals), association was performed with the Mixed Model Analysis for Pedigrees (MMAP) software (C.J. O'Donnell, 2008, Am. Soc. Hum. Genet., abstract). The genomic control inflation factor was calculated in each cohort and used for within-study correction. For obtaining effect-size estimates, phase I results were obtained by a fixed-effect inverse-variance weighted meta-analysis in METAL. In phase II (including summary-level cohorts of European ancestry), phase III (including all individuals of European ancestry), and phase IV (a multiethnic meta-analyses), meta-analysis was performed with the use of a fixed-effect sample-size weighted Z-score meta-analysis in METAL.[21] All reported p values are two-sided and uncorrected for multiple testing. It is important to note that although we label regions with either the gene nearest to the lead SNP in the locus or a nearby known growth-related gene for convenience, it is not possible to provide evidence of which gene in the region is functionally responsible through association analysis alone. Thus, it is possible that one or more genetic variants within one or more surrounding genes could be contributing to the association signal.

Calculating an appropriate significance threshold is challenging in the context of an array containing hypothesis-driven, densely covered loci, enriched for functional and nonsynonymous polymorphisms. Previous studies using the IBC array have used significance thresholds of $p < 1 \times 10^{-5}$ and $p < 1 \times 10^{-6}$. The Candidate gene Association Resource (CARe) IBC array studies determined that after accounting for LD, the effective number of independent tests was 26,500 for African Americans and 20,500 for European Americans, producing an "array-wide" statistical threshold of $p = 1.9 \times 10^{-6}$ and $p = 2.4 \times 10^{-6}$, respectively, to maintain a false-positive rate of 5%.[22] We also highlight loci that are significantly associated at a more conventional genome-wide significance threshold of $p < 5.0 \times 10^{-8}$. In genetic association studies, power drops dramatically at low allele frequencies. For a SNP with an effect of 7 mm per risk allele, the phase III meta-analysis of 90,446 individuals of European descent yields greater than 95% power to detect a SNP with an MAF of 5% and 80% power for a SNP with an MAF of 3% ($\alpha = 2.4 \times 10^{-6}$).

All loci harboring significant evidence for association were examined for additional signals via conditional analyses in PLINK.[18] A term was added to the regression model, including the lead SNP as a covariate, and surrounding SNPs were evaluated for maintaining array-wide significance. Conditional analysis was performed only in European ancestry cohorts in which individual-level data were available (n = 53,394 from phase I).

After association tests were performed in the sex- and ethnicity-stratified cohorts containing additional ethnicities, a multiethnic meta-analysis including all available participants was performed. Additionally, the direction of effect of the lead SNPs from the previously identified loci was evaluated for consistency in the three additional ethnicities with more than 1000 participants available: African Americans (n = 11,357), South Asians (n = 6003), and Hispanics (n = 4934).

Results

Phase I of investigation into height with the use of the IBC array involved testing for association in participants of European ancestry in which individual-level data were available (n = 53,394). In phase II of the analysis, we sought replication in cohorts providing summary-level data for participants of European ancestry (n = 37,052), followed by a meta-analysis of all participants of European ancestry (n = 90,446) in phase III. To ensure the validity of our study design, we began by examining evidence in loci previously reported to be associated with height in GWAS. The lead SNP (rs4272) in cyclin-dependent kinase 6 (*CDK6* [MIM 603368]) demonstrated strong evidence for association in both phase I and phase II of analysis, confirming a strong positive control for a previously described height signal (phase I, $p = 2.5 \times 10^{-20}$; phase II, $p = 3.2 \times 10^{-17}$; meta-analysis phase III, $p = 4.1 \times 10^{-36}$). Additional previously identified height genes were also significantly associated in phase I analysis, with the same SNP replicating in the same direction in phase II, in the following genes: high mobility group AT-hook 1 (*HMGA1* [MIM 600701]), high mobility group AT-hook 2 (*HMGA2* [MIM 600698]), T-box 2 (*TBX2* [MIM 600747]), suppressor of cytokine signaling 2 (*SOCS2* [MIM 605117]), aggrecan (*ACAN* [MIM 155760]) and patched *Drosophila* homolog 1 (PTCH1 [MIM 601309]) (phase I, $p < 7 \times 10^{-7}$; phase II, same direction and $p < 2 \times 10^{-6}$; meta-analysis, $p < 1 \times 10^{-13}$; Table 1).

A total of 34 genes were significantly associated with height in phase I at array-wide significance ($p < 2.4 \times 10^{-6}$). All associated SNPs in phase I were replicated with the same direction of effect in phase II ($p < 0.05$). In phase III, 64 loci were significantly associated with height at array-wide significance ($p < 2.4 \times 10^{-6}$) with 42 loci surpassing the traditional genome-wide significance threshold ($p < 5.0 \times 10^{-8}$). Of 87 GWAS-identified loci reported before 2010,[4] 27 had SNPs that were present on the IBC array and 20 of them surpassed array-wide significance, 17 of them surpassing genome-wide significance (Table 1). A SNP in strong LD with the previously reported lead SNP in only three of the seven nonreplicated loci ($r^2 > 0.3$) was present (Table 2). Marginal association was observed for all of the loci reported before 2010 that did not reach array-wide significance ($p \leq 0.05$). Of the 64 associated loci in the current study, 33 were identified in

a recent height meta-analysis including 183,727 individuals of European ancestry by Lango Allen and colleagues.[5]

Some of the associated regions without previous reports of association with height containing genes with interesting biological roles include the following: myosin heavy chain 7b (*MYH7B* [MIM 609928]), growth hormone receptor (*GHR* [MIM 600946]), collagen type 11 alpha 1 (*COL11A1* [MIM 120280]), collagen type 25 alpha 1 (*COL25A1* [MIM 610025]), glucokinase regulatory protein (*GCKR* [MIM 600842]), circadian locomotor output cycles kaput (*CLOCK* [MIM 601851]), re1-silencing transcription factor (*REST* [MIM 600571]) and Bardet-Biedl syndrome 7 (*BBS7* [MIM 607590]).

A total of 22 uncommon SNPs (MAF < 5%) that were observed to be significantly associated with human height were found in eight loci: *HMGA1*, *ACAN*, peroxisome proliferator-activated receptor delta (*PPARD* [MIM 600409]), potassium channel voltage-gated KQT-like subfamily member 1 (KCNQ1 [MIM 607542]), insulin-like growth factor 1 receptor (*IGF1R* [MIM 147370]), mitogen-activated protein kinase 14 (*MAPK14* [MIM 600289]), interleukin-11 (*IL11* [MIM 147681]), and mothers against decapentaplegic drosphila homolog 3 (*SMAD3* [MIM 603109]). In two of these genes, the uncommon allele showed the strongest evidence for association (*ACAN* and *KCNQ1*; Table 1), whereas in an additional two genes the uncommon SNP was the only associated variant: *IL11* (phase I, $p = 5.4 \times 10^{-10}$; phase II, $p = 2.8 \times 10^{-5}$; meta-analysis, $p = 1.5 \times 10^{-13}$) and *SMAD3* (phase I, $p = 1.8 \times 10^{-8}$; phase II, $p = 4.5 \times 10^{-3}$; meta-analysis, $p = 1.0 \times 10^{-9}$).

With the use of the empirical results of the current meta-analysis, a plot of the effect size of associated variants as a function of MAF was produced (Figure 2). Alleles in the top right corner of the plot would be common in the population and of large effect, making them easy to identify, but are unobserved for height. Alleles in the bottom right corner of the plot are of small effect but can be identified because of their high frequency in the population. Conversely, alleles in the top left are rarer in the population but may be identified through their large effect sizes.

Conditional analysis was performed for the identification of loci harboring multiple variants independently influencing adult height. Regression was repeated in the phase I study cohorts, conditioned upon the lead SNP for each of the 64 associated loci. In five loci, a second variant obtained array-wide significance after being conditioned upon the lead SNP (Table 3).

Male-only and female-only meta-analyses were performed and tested for significant heterogeneity, which provided no evidence of gender-specific signals of adult height (Table S3). Because a number of the studies included in the meta-analysis comprise CVD-related studies, we restricted an analysis to 47,451 individuals of European ancestry collected as healthy controls or included in studies with a population-based ascertainment scheme. The directions of effect for all of the 64 lead SNPs were consistent with the observations in phase III, and all SNPs remained at least marginally significant ($p < 0.05$; Table S4).

Association testing in African American (n = 11,357), South Asian (n = 6003), Hispanic (n = 4934), East Asian (n = 984), and Native American (n = 499) populations independently revealed no loci with array-wide significance. In the phase IV multiethnic meta-analysis of all available individuals (n = 114,223), the significance of seven loci fell below the array-wide significance threshold, whereas five loci showed array-wide significance (Table S5). Remarkable concordance of the direction of effect was observed between ethnicities: 48 out of 64 SNPs between Europeans and African Americans ($p = 3.9 \times 10^{-5}$), 49 out of 61 SNPs between Europeans and South Asians ($p = 9.8 \times 10^{-8}$), and 53 out of 64 SNPs between Europeans and Hispanics ($p = 5.0 \times 10^{-8}$; Table S6). In total, 35 out of 64 SNPs were concordant across all four ethnic groups examined ($p = 8.0 \times 10^{-16}$).

Discussion

In a meta-analysis of genecentric association studies of height, including 114,223 individuals from 47 studies and six ethnicities, significant association was identified for SNPs within 64 loci. Twenty previously identified height-associated loci were replicated, providing validation of our study as positive control loci. Thirty-three of 64 associated loci reported here were identified in a recent meta-analysis of individuals of European ancestry.[5] Two loci, *IL11* and *SMAD3*, were uncovered via direct genotyping of uncommon nonsynonymous SNPs, which would not have been identified with the use of standard genome-wide genotyping arrays. Biological hypotheses exist for many of the associated loci identified here, with many previously unreported loci falling into known biological pathways such as energy metabolism, insulin and growth hormone signaling, heart morphogenesis, cellular growth and apoptosis, circadian rhythm, and collagen formation.[5] Previously unreported common variants were identified as being associated with adult height in or near genes known to be mutated in monogenic diseases involving abnormal growth or height, such as *COL11A1* and *BBS7*. Additionally, loci containing genes with no known role in growth or height were identified, such as ribosomal protein S6 kinase 1 (RPS6KA1 [MIM 601684]) and CDK5 regulatory subunit-associated protein 1 (*CDK5RAP1* [MIM 608200]).

GWAS were conceived for testing of the hypothesis that common genetic variants are associated with heritable traits. Efforts to identify uncommon SNPs (MAF < 5%) have generally been limited to the identification of variants with large effect via deep resequencing. However, it is rational to hypothesize that lower-frequency variants could also be associated with moderate to small effects. Resequencing studies that identify uncommon variants

Table 1. Sixty-Four Loci Showing Significant Evidence for Association with Adult Height, Identified with the Use of the IBC Array

Locus Rank	Chr.	Candidate Gene[a]	SNP[a]	Effect Allele	MAF	European Ancestry Phase I (up to 53,394)			European Ancestry Phase II (up to 37,052)		European Ancestry Phase III (up to 90,446)	Multiethnic Phase IV[b] (up to 114,223)	Reported before 2010	In Lango Allen et al. (2010)[5]
						Effect	p	I^2	(+/−)	p	p	p		
1	7q22	*CDK6*	rs4272	A	0.21	−0.46	2.5×10^{-20}	0	−	3.2×10^{-17}	1.8×10^{-36}	4.1×10^{-36}	Yes	Yes
2	6p21	*HMGA1*	rs1150781	C	0.09	0.73	2.2×10^{-24}	0	+	3.3×10^{-10}	7.3×10^{-32}	2.0×10^{-39}	Yes	Yes
3	12q15	*HMGA2*	rs867633	A	0.41	−0.39	1.6×10^{-20}	0	−	4.1×10^{-12}	5.6×10^{-31}	1.7×10^{-30}	Yes	Yes
4	20q11	*MMP24*	rs2425019	A	0.46	−0.32	4.9×10^{-14}	7	−	6.7×10^{-14}	2.4×10^{-26}	6.4×10^{-26}	Yes	
5	17q23	*MAP3K3*	rs8081612	T	0.28	0.37	6.2×10^{-12}	3	+	1.6×10^{-7}	3.2×10^{-20}	1.3×10^{-22}	Yes	
6	17q24	*GH1-GH2*	rs7921	A	0.25	0.34	2.0×10^{-13}	8	+	6.2×10^{-8}	3.3×10^{-20}	3.0×10^{-21}		Yes
7	1p36	*MFAP2*	rs2284746	C	0.49	−0.30	2.7×10^{-12}	0	−	1.9×10^{-8}	1.1×10^{-19}	9.2×10^{-19}	Yes	Yes
8	15q26	*IGF1R*	rs2871865	C	0.11	0.44	7.2×10^{-12}	0	+	3.5×10^{-8}	1.3×10^{-18}	7.9×10^{-19}		Yes
9	7p22	*GNA12*	rs1636255	A	0.26	−0.39	7.8×10^{-12}	19	−	3.6×10^{-8}	3.0×10^{-18}	7.0×10^{-19}	Yes	Yes
10	17q23	*TBX2*	rs9892365	A	0.33	0.25	4.4×10^{-9}	1	+	2.6×10^{-10}	1.4×10^{-17}	1.4×10^{-17}	Yes	Yes
11	12q22	*SOCS2*	rs3782415	T	0.21	−0.39	7.1×10^{-15}	0	−	2.1×10^{-4}	1.2×10^{-16}	8.3×10^{-16}	Yes	Yes
12	9q22	*PTCH1*	rs10512248	T	0.33	−0.21	6.7×10^{-7}	13	−	2.1×10^{-9}	1.1×10^{-14}	5.3×10^{-14}	Yes	Yes
13	14q11	*NFATC4*	rs12590407	T	0.29	−0.27	2.9×10^{-9}	0	−	1.9×10^{-6}	1.5×10^{-14}	9.4×10^{-13}		Yes
14	15q26	*ACAN*	rs16942341	T	0.03	−0.73	1.8×10^{-9}	0	−	1.1×10^{-6}	2.4×10^{-14}	9.6×10^{-16}	Yes	Yes
15	2q24	*NPPC*	rs2679178	T	0.09	−0.44	1.3×10^{-9}	3	−	9.8×10^{-6}	4.4×10^{-14}	5.8×10^{-14}	Yes	Yes
16	6p21	*PPARD*	rs3734254	T	0.22	0.27	3.2×10^{-7}	26	+	1.7×10^{-7}	1.1×10^{-13}	4.7×10^{-11}		Yes
17	20q11	*MYH7B*	rs2425012	A	0.43	−0.25	8.0×10^{-9}	2	−	5.2×10^{-5}	3.4×10^{-13}	5.2×10^{-12}		
18	19q13	*IL11*	rs4252548	T	0.03	−0.81	5.4×10^{-10}	0	−	8.8×10^{-5}	7.1×10^{-13}	2.8×10^{-12}		
19	3q26	*GHSR*	rs572169	T	0.30	0.25	1.8×10^{-8}	33	+	4.2×10^{-6}	8.3×10^{-13}	9.9×10^{-13}		Yes
20	2p23	*POMC*	rs1866146	A	0.34	−0.23	6.5×10^{-8}	0	−	7.4×10^{-6}	2.5×10^{-12}	1.5×10^{-11}	Yes	
21	5p14	*NPR3*	rs1173736	A	0.26	−0.26	1.1×10^{-7}	0	−	1.5×10^{-4}	7.3×10^{-12}	1.4×10^{-10}	Yes	Yes
22	5p13	*GHR*	rs6180	A	0.46	0.18	1.8×10^{-5}	0	+	6.8×10^{-8}	1.8×10^{-11}	3.1×10^{-12}		
23	15q22	*SMAD3*	rs35874463	A	0.05	−0.59	1.8×10^{-8}	0	−	1.1×10^{-4}	2.5×10^{-11}	3.4×10^{-13}		
24	11p15	*SPTY2D1*	rs11024739	A	0.26	−0.16	9.3×10^{-4}	0	−	9.3×10^{-10}	3.8×10^{-11}	1.9×10^{-10}		
25	11p15	*KCNQ1*	rs2075870	A	0.03	0.18	1.8×10^{-5}	0	−	4.3×10^{-5}	9.8×10^{-11}	1.8×10^{-8}		Yes
26	1p21	*COL11A1*	rs4338381	A	0.37	−0.18	3.6×10^{-5}	0	−	9.2×10^{-7}	1.6×10^{-10}	2.9×10^{-10}		
27	9q21	*PCSK5*	rs11144688	A	0.12	−0.32	2.2×10^{-7}	0	−	1.9×10^{-4}	3.3×10^{-10}	5.0×10^{-9}		Yes

Table 1. *Continued*

Locus Rank	Chr.	Candidate Gene[a]	SNP[a]	Effect Allele	MAF	European Ancestry Phase I (up to 53,394) Effect	p	I^2	European Ancestry Phase II (up to 37,052) (+/−)	p	European Ancestry Phase III (up to 90,446) p	Multiethnic Phase IV[b] (up to 114,223) p	Reported before 2010	In Lango Allen et al. (2010)[5]
28	2p23	*GCKR*	rs780094	T	0.41	−0.17	5.8×10^{-5}	0	−	1.1×10^{-6}	6.4×10^{-10}	2.2×10^{-11}		
29	1q41	*TGFB2*	rs900	A	0.28	−0.22	5.6×10^{-7}	0	−	2.7×10^{-4}	8.0×10^{-10}	6.0×10^{-10}		Yes
30	20q11	*CDK5RAP1*	rs291700	T	0.31	−0.22	2.4×10^{-7}	0	−	7.1×10^{-4}	9.9×10^{-10}	4.4×10^{-10}		
31	2p12	*EIF2AK3*	rs867529	C	0.27	0.24	3.2×10^{-7}	0	+	5.5×10^{-4}	1.3×10^{-8}	1.4×10^{-10}		Yes
32	19p13	*INSR*	rs8108622	A	0.23	0.24	9.9×10^{-7}	1	+	4.8×10^{-4}	1.8×10^{-9}	2.5×10^{-10}		Yes
33	6q25	*ESR1*	rs488133	T	0.33	−0.21	1.8×10^{-6}	16	−	2.5×10^{-4}	2.6×10^{-9}	1.2×10^{-10}		Yes
34	2q37	*DIS3L2*	rs3103296	T	0.37	−0.23	1.1×10^{-7}	0	−	3.6×10^{-4}	4.8×10^{-9}	9.4×10^{-7}	Yes	
35	2q35	*PLCD4*	rs611203	A	0.42	0.16	1.0×10^{-4}	0	+	9.8×10^{-6}	5.8×10^{-9}	7.0×10^{-9}		
36	1p36	*RPS6KA1*	rs3816540	A	0.23	0.19	1.1×10^{-4}	0	+	1.4×10^{-5}	8.5×10^{-9}	1.2×10^{-7}		
37	15q21	*CYP19A1*	rs3751591	A	0.17	0.25	6.1×10^{-6}	0	+	3.4×10^{-4}	9.4×10^{-9}	7.4×10^{-9}	Yes	Yes
38	5q31	*SLC22A5*	rs17622208	A	0.47	0.17	5.9×10^{-5}	5	+	2.1×10^{-5}	1.1×10^{-8}	3.2×10^{-12}		Yes
39	7p15	*JAZF1*	rs864745	T	0.50	0.21	1.8×10^{-5}	0	+	1.9×10^{-4}	1.9×10^{-8}	1.7×10^{-9}	Yes	Yes
40	17p13	*POLR2A*	rs8071847	A	0.21	−0.20	6.7×10^{-5}	0	−	7.4×10^{-5}	3.0×10^{-8}	5.0×10^{-9}		
41	1p22	*PKN2*	rs12145922	A	0.43	0.15	2.6×10^{-4}	0	+	2.6×10^{-5}	3.2×10^{-8}	2.7×10^{-8}		Yes
42	7q22	*CNOT4*	rs3812265	T	0.24	0.23	9.2×10^{-7}	0	+	9.3×10^{-3}	3.4×10^{-8}	9.2×10^{-8}		
43	14p11	*REST*	rs3796529	T	0.19	0.26	5.1×10^{-7}	32	+	1.4×10^{-2}	5.7×10^{-8}	1.2×10^{-7}		
44	6p21	*MICA*	rs2516448	A	0.49	0.21	9.3×10^{-4}	0	+	1.2×10^{-5}	7.0×10^{-8}	3.2×10^{-8}		Yes
45	11p11	*PTPRJ*	rs4752805	A	0.25	−0.22	9.7×10^{-6}	0	−	2.6×10^{-3}	7.9×10^{-8}	1.3×10^{-7}		
46	16p13	*CASKIN1*	rs258281	A	0.19	−0.23	1.5×10^{-5}	19	−	2.5×10^{-3}	8.3×10^{-8}	2.0×10^{-9}		Yes
47	3q21	*PCCB*	rs9844666	A	0.24	−0.22	5.8×10^{-6}	9	−	2.9×10^{-3}	8.9×10^{-8}	1.7×10^{-7}		Yes
48	14q22	*SAMD4A*	rs709939	T	0.46	0.15	2.5×10^{-4}	0	+	2.4×10^{-4}	1.8×10^{-7}	2.3×10^{-6}		
49	11q13	*BBS1-CTSF*	rs4630309	A	0.24	0.23	1.8×10^{-6}	20	+	2.1×10^{-2}	2.7×10^{-7}	4.7×10^{-7}		
50	4q27	*BBS7*	rs7659604	T	0.41	0.20	2.2×10^{-6}	0	+	6.7×10^{-3}	3.2×10^{-7}	5.2×10^{-7}		
51	4q12	*CLOCK*	rs4864546	A	0.37	0.21	7.3×10^{-7}	0	+	4.2×10^{-2}	4.0×10^{-7}	6.4×10^{-8}		
52	12p12	*PDE3A*	rs7137534	T	0.32	0.18	4.0×10^{-5}	0	+	4.9×10^{-3}	6.1×10^{-7}	4.3×10^{-7}	Yes	Yes
53	12q24	*MPHOSPH9*	rs1051431	A	0.22	−0.19	3.0×10^{-4}	3	−	1.4×10^{-3}	6.9×10^{-7}	6.2×10^{-6}		

(*Continued on next page*)

Table 1. *Continued*

Locus Rank	Chr.	Candidate Gene[a]	SNP[a]	Effect Allele	MAF	European Ancestry Phase I (up to 53,394)			European Ancestry Phase II (up to 37,052)		European Ancestry Phase III (up to 90,446)	Multiethnic Phase IV[b] (up to 114,223)	Reported before 2010	In Lango Allen et al. (2010)[5]
						Effect	p	I^2	(+/−)	p	p	p		
54	1p22	*COL24A1*	rs2046159	A	0.16	0.23	3.8×10^{-5}	25	+	7.0×10^{-3}	7.1×10^{-7}	1.1×10^{-5}		
55	1q23	*DUSP23*	rs1129923	A	0.10	−0.25	2.7×10^{-4}	0	−	8.0×10^{-4}	7.4×10^{-7}	5.9×10^{-5}		
56	10q22	*MAT1A*	rs7087728	A	0.18	0.22	2.2×10^{-4}	0	+	1.4×10^{-3}	9.1×10^{-7}	1.4×10^{-6}		
57	2p15	*PPP3R1*	rs1822469	T	0.41	−0.14	7.8×10^{-4}	9	−	2.2×10^{-4}	9.3×10^{-7}	1.4×10^{-5}		
58	7q36	*ATG9B*	rs1800783	A	0.38	−0.16	2.0×10^{-4}	0	−	2.1×10^{-3}	1.2×10^{-6}	1.9×10^{-6}		
59	14q11	*BCL2L2*	rs3210043	A	0.16	0.25	9.7×10^{-6}	0	+	2.0×10^{-2}	1.3×10^{-6}	5.4×10^{-8}		
60	4p14	*RFC1*	rs11096991	T	0.35	0.15	3.6×10^{-4}	0	+	1.9×10^{-3}	1.5×10^{-6}	1.5×10^{-5}		
61	6p21	*HLA-B*	rs2596494	C	0.17	0.24	1.5×10^{-3}	12	+	2.9×10^{-4}	1.8×10^{-6}	4.9×10^{-6}	Yes	Yes
62	6q21	*ZBTB24*	rs1476387	T	0.42	−0.12	3.8×10^{-3}	8	−	3.0×10^{-4}	1.9×10^{-6}	2.9×10^{-6}		Yes
63	17q24	*GRB2*	rs959260	T	0.18	0.16	4.1×10^{-3}	15	+	2.8×10^{-4}	2.1×10^{-6}	2.1×10^{-6}		
64	19p13	*ADAMTS10*	rs8111085	T	0.07	−0.30	3.2×10^{-4}	0	−	2.6×10^{-3}	2.2×10^{-6}	4.0×10^{-6}	Yes	Yes

Phase I employed an inverse-variance weighted fixed-effect meta-analysis for the estimation of effect size. Phase II, phase III, and the multiethnic meta-analyses used a sample-size weighted Z-score-based fixed-effect meta-analysis. (+/−) indicates the direction of effect in Z-based meta-analysis.

[a] Lead SNP in locus. Nearest gene unless there is a known growth-related gene in the locus.

[b] Meta-analysis results include European-descent (n = 90,446), African American (n = 11,357), South Asian (n = 6003), Hispanic (n = 4934), East Asian (n = 984), and Native American (n = 499).

Table 2. Loci Previously Identified by GWAS that Failed to Replicate at Array-wide Significance in Phase III

Chr.	Candidate Gene	Previous Lead SNP	Lead SNP on IBC Array	r^2 between Previous and Lead SNP on IBC Array	European Ancestry Phase III (up to 90,446) p	In Lango Allen et al. (2010)[5]
4q12	*PDGFRA*	rs17690232	rs7660759	0.07	6.3×10^{-3}	
8q13	*LYN*	rs10958476	rs13249338	0.12	2.2×10^{-3}	
9q33	*PAPPA*	rs789550	rs7020782	0.58	9.2×10^{-6}	Yes
12q23	*IGF1*	rs5742692	rs1019731	0.001	7.1×10^{-4}	
14q32	*FBLN5*	rs7153027	rs3783937	0.52	3.1×10^{-4}	
17q22	*NOG*	rs4794665	rs1076352	0.06	0.03	Yes
20p12	*BMP2*	rs967417	rs6107869	0.06	0.05	Yes

typically use a gene-based approach, totaling the number and category of variants within specific genes, to overcome the low power yielded by rare variants.[23,24] Because the IBC array contains many SNPs with MAFs < 5% and a very large number of individuals have been genotyped on the array, it provides a unique opportunity for a well-powered test for association of lower-frequency variants with relatively small effect sizes directly, without the need for "mutation counting"- or "mutation dosage"-based tests.

In the current meta-analysis, a total of eight genes contained uncommon SNPs (MAF < 5%) significantly associated with height. Perhaps the most important discoveries are the two loci that would not have been identified without direct genotyping of the low-frequency variants. The uncommon SNP in *IL11* (rs4252548) causes an arginine-to-histidine substitution at position 112, replacing a large basic amino acid with a medium-sized polar amino acid. Interleukin 11 (IL-11) is relatively undercharacterized compared to other interleukins; however, it is known that IL-11 signaling induces the proliferation of hematopoietic cells and enhances bone formation and remodeling.[25,26] The uncommon SNP in *SMAD3* (rs35874463) causes an isoleucine-to-valine substitution at residue 126 of SMAD3. SMAD3 is a transcriptional modifier activated by TGF-β,[27] a signaling pathway that has been implicated in height.[5] In *Smad3* knockout mice, a significantly smaller body size is attained and degeneration of the spinal intervertebral discs is observed.[28] For both *IL11* and *SMAD3*, the uncommon alleles are associated with a reduction in attained height with observed effects of 6–8 mm. We cannot conclude from an association study whether the measured allele is functionally responsible for the effect. Further examination of the pleiotropic effect of the alleles can provide clues, but in vitro and in vivo functional analyses are required to concretely establish the effect of the genotyped alleles. Imputation of rs4252548 is not currently possible with conventional GWAS data sets or the use of surrounding SNPs on the IBC array (which contains denser coverage than conventional genome-wide genotyping products), and rs35874463 is not found in HapMap 3. Direct genotyping of these uncommon SNPs is currently the only way to detect their association with height.

A plot of the absolute effect size versus the MAF of genetic variants is often shown to describe the contribution of genetics to complex traits. As meta-analyses grow in size and genetic investigations are modified to include variants of lower allele frequency, our ability to identify less-common, smaller-effect SNPs, closer to the origin of the plot, will improve. The same paradigm is likely to be true in other complex traits, for which improvements in the density of coverage for capturing more of the genetic diversity, including lower-frequency variants, will allow additional signals underpinning complex traits to be identified.

Pleiotropy will become more apparent as the power to detect smaller effect sizes improves in the study of complex traits. Many of the genes identified as being associated with height in the current meta-analysis are also associated with other phenotypes. Of interest, the largest overlap appears to be with type 2 diabetes, as four genes previously reported as being associated with fasting glucose, fasting insulin, insulin resistance, or type 2 diabetes risk are associated with height in the current meta-analysis: *HMGA2*, *GCKR*, *KCNQ1*, and juxtaposed with another zinc finger 1 (*JAZF1* [MIM 66246]).[29] Genetic variation in *GCKR* appears extremely pleiotropic, as the T allele of rs780094 has been associated with numerous traits, including increased plasma triglycerides,[30] increased C-reactive protein,[31] increased uric acid,[32] reduced fasting plasma glucose, and reduced insulin resistance,[33] and the same allele is now associated with reduced adult height. Similarly, SNPs in *KCNQ1* have been associated with not only type 2 diabetes,[29] but also platelet aggregation,[34] QT interval,[35] and now height.

Epidemiological data have provided support of an association between short stature and a small increase in CVD risk.[36] Because a number of the individuals included in the current study were collected in clinical trials, case-only, and case-control studies of CVD, there exists the possibility that the increased prevalence of CVD among study subjects could confound the association with height. To remove this potential source of confounding, we performed a meta-analysis including only the controls from

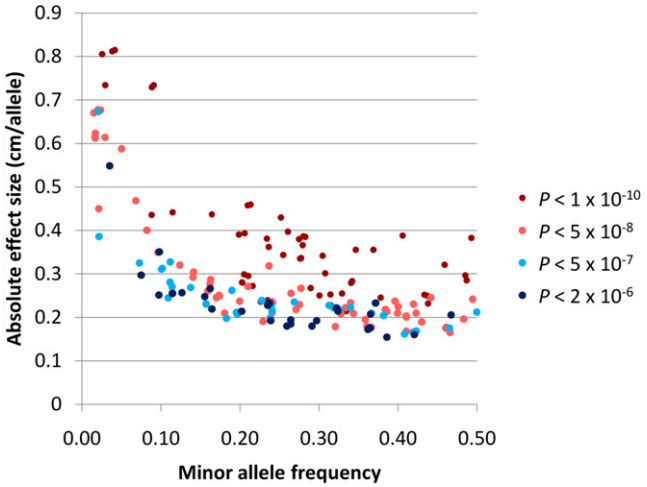

Figure 2. The Effect Size of Identified Height-Associated Genetic Variants as a Function of Minor Allele Frequency
Each point is colored by the strength of association observed in the phase III meta-analysis.

case-control studies and cohorts with population-based ascertainment strategies. Because of the decreased power afforded by the smaller sample size of the restricted analysis, p values were reduced in comparison to those of phase III, but all identified lead SNPs in phase III remained marginally significant with the same direction of effect.

It appears likely that many of the loci associated with variation in adult height in individuals of European ancestry will have the same direction of effect in African American, South Asian, and Hispanic populations. Association results from the additional ethnicities did not independently uncover array-wide or genome-wide significant associations, which is not unexpected given the lower power of the smaller data sets. When replication of the lead SNPs from the European ancestry cohorts in the additional ethnicities was attempted, the direction of effect was concordant more often than would be expected to result from chance. With the same approach used, replication of common variants associated with lipid traits in additional ethnicities showed similar trends, suggesting that many common alleles associated with complex traits are likely to have similar direction of effect across ethnicities.[30]

The effect sizes of the associated variants in this meta-analysis were similar to previous reports, ranging from 0.15 cm to 0.81 cm per allele. Unfortunately, the current study was unable to directly compare the total extent of explained variation to previous reports, because 60 of 87 previously reported height loci were not genotyped in the current study. The current study identified five genes containing two independent signals for association with height. However, conditional analysis was only possible within phase I cohorts with individual-level data available. A total of seven loci reported to be associated with height before 2010 failed to reach array-wide significance in the current study. Marginal association was observed for all of the nonreplicated loci ($p \leq 0.05$). Four reasons exist for a locus to fail to replicate: (1) the first report was a false positive; (2) the current report is a false negative; (3) differences in study design or phenotype exist; or (4) differences in study populations exist. In all cases except insulin-like growth factor 1 (*IGF1* [MIM 147440]), the previously reported lead SNP was not directly genotyped, leaving inadequate coverage as a likely reason for nonreplication. Additionally, heterogeneity of study design between cohorts contributing to the meta-analysis may have reduced the signal-to-noise ratio for less robust signals. Interestingly, only three of the seven nonreplicated loci were found to be associated in another large meta-analysis of height that was recently reported.[5]

In conclusion, meta-analyses of up to 114,223 individuals across six ethnic groups from 47 studies genotyped on the genecentric IBC array identified 64 height-associated loci. Association between height and either *IL11* or *SMAD3* would not have been observed without the inclusion of direct genotyping of uncommon SNPs and large sample size. The direction of effect of common variants associated

Table 3. Loci with Significant Evidence of Two Independent Height Association Signals

Gene	SNP	Position[a]	MAF	Phase I p	Conditional p	r^2 with Lead SNP	D′ with Lead SNP
NPR3	rs1173736	32807695	0.26	1.1×10^{-7}	-	-	-
	rs1421811	32750027	0.39	1.1×10^{-4}	1.9×10^{-6}	0.01	0.22
PROCR-MMP24	rs2425019	33282831	0.46	4.9×10^{-14}	-	-	-
	rs8115394	33353764	0.30	9.1×10^{-15}	1.1×10^{-6}	0.20	0.61
NPPC	rs2679178	232506105	0.09	1.3×10^{-9}	-	-	-
	rs3107179	232496569	0.40	4.9×10^{-8}	9.6×10^{-10}	0	0.03
PPARD	rs3734254	35502988	0.22	3.2×10^{-7}	-	-	-
	rs7751726	35479602	0.03	1.1×10^{-6}	2.5×10^{-7}	0.01	0.35
ACAN	rs16942341	87189909	0.03	1.8×10^{-9}	-	-	-
	rs938609	87199635	0.36	7.6×10^{-5}	4.6×10^{-9}	0.05	1.00

[a] National Center for Biotechnology Information (NCBI) build 36.

with height was significantly concordant across individuals of European, African American, South Asian, and Hispanic ancestries. The increased power to identify variants of small effect, afforded by large sample size and the dense genetic coverage including low-frequency SNPs within loci of interest, has resulted in the identification of association between previously unreported genes and height.

Supplemental Data

Supplemental Data include six tables and Supplemental Acknowledgments and can be found with this article online at http://www.cell.com/AJHG/.

Acknowledgments

We thank the researchers, staff, and participants of all of the studies that contributed data. Specific cohort acknowledgements are cited in the Supplemental Acknowledgments. Matthew B. Lanktree is supported by a Canadian Institutes of Health Research (CIHR) M.D.-Ph.D. Studentship Award. Robert A. Hegele is funded by CIHR grant 79533 and by Genome Canada through the Ontario Genomics Institute.

Received: September 14, 2010
Revised: October 22, 2010
Accepted: November 12, 2010
Published online: December 30, 2010

Web Resources

The URLs for data presented herein are as follows:

EIGENSOFT, http://genepath.med.harvard.edu/~reich/Software.htm
HapMap, http://www.hapmap.org
IBC array reference, http://bmic.upenn.edu/cvdsnp
Illumina, http://www.illumina.com/products/humancvd_whole_genome_genotyping_kits.ilmn
METAL, http://www.sph.umich.edu/csg/abecasis/metal/
National Human Genome Resource Institute GWAS database, http://www.genome.gov/26525384#1
Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim/
PLINK, http://pngu.mgh.harvard.edu/~purcell/plink/

References

1. Silventoinen, K., Sammalisto, S., Perola, M., Boomsma, D.I., Cornes, B.K., Davis, C., Dunkel, L., De Lange, M., Harris, J.R., Hjelmborg, J.V., et al. (2003). Heritability of adult body height: a comparative study of twin cohorts in eight countries. Twin Res. *6*, 399–408.
2. Perola, M., Sammalisto, S., Hiekkalinna, T., Martin, N.G., Visscher, P.M., Montgomery, G.W., Benyamin, B., Harris, J.R., Boomsma, D., Willemsen, G., et al; GenomEUtwin Project. (2007). Combined genome scans for body stature in 6,602 European twins: evidence for common Caucasian loci. PLoS Genet. *3*, e97.
3. Fisher, R. (1918). The correlation between relatives on the supposition of Mendelian inheritance. Trans. R. Soc. Edinb. *52*, 399–433.
4. Hindorff, L.A., Sethupathy, P., Junkins, H.A., Ramos, E.M., Mehta, J.P., Collins, F.S., and Manolio, T.A. (2009). Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. Proc. Natl. Acad. Sci. USA *106*, 9362–9367.
5. Lango Allen, H., Estrada, K., Lettre, G., Berndt, S.I., Weedon, M.N., Rivadeneira, F., Willer, C.J., Jackson, A.U., Vedantam, S., Raychaudhuri, S., et al. (2010). Hundreds of variants clustered in genomic loci and biological pathways affect human height. Nature *467*, 832–838.
6. Cho, Y.S., Go, M.J., Kim, Y.J., Heo, J.Y., Oh, J.H., Ban, H.J., Yoon, D., Lee, M.H., Kim, D.J., Park, M., et al. (2009). A large-scale genome-wide association study of Asian populations uncovers genetic factors influencing eight quantitative traits. Nat. Genet. *41*, 527–534.
7. Estrada, K., Krawczak, M., Schreiber, S., van Duijn, K., Stolk, L., van Meurs, J.B., Liu, F., Penninx, B.W., Smit, J.H., Vogelzangs, N., et al. (2009). A genome-wide association study of northwestern Europeans involves the C-type natriuretic peptide signaling pathway in the etiology of human height variation. Hum. Mol. Genet. *18*, 3516–3524.
8. Gudbjartsson, D.F., Walters, G.B., Thorleifsson, G., Stefansson, H., Halldorsson, B.V., Zusmanovich, P., Sulem, P., Thorlacius, S., Gylfason, A., Steinberg, S., et al. (2008). Many sequence variants affecting diversity of adult human height. Nat. Genet. *40*, 609–615.
9. Kim, J.J., Lee, H.I., Park, T., Kim, K., Lee, J.E., Cho, N.H., Shin, C., Cho, Y.S., Lee, J.Y., Han, B.G., et al. (2010). Identification of 15 loci influencing height in a Korean population. J. Hum. Genet. *55*, 27–31.
10. Lettre, G., Jackson, A.U., Gieger, C., Schumacher, F.R., Berndt, S.I., Sanna, S., Eyheramendy, S., Voight, B.F., Butler, J.L., Guiducci, C., et al; Diabetes Genetics Initiative; FUSION; KORA; Prostate, Lung Colorectal and Ovarian Cancer Screening Trial; Nurses' Health Study; SardiNIA. (2008). Identification of ten loci associated with height highlights new biological pathways in human growth. Nat. Genet. *40*, 584–591.
11. Liu, J.Z., Medland, S.E., Wright, M.J., Henders, A.K., Heath, A.C., Madden, P.A., Duncan, A., Montgomery, G.W., Martin, N.G., and McRae, A.F. (2010). Genome-wide association study of height and body mass index in Australian twin families. Twin Res. Hum. Genet. *13*, 179–193.
12. Okada, Y., Kamatani, Y., Takahashi, A., Matsuda, K., Hosono, N., Ohmiya, H., Daigo, Y., Yamamoto, K., Kubo, M., Nakamura, Y., and Kamatani, N. (2010). A genome-wide association study in 19 633 Japanese subjects identified LHX3-QSOX2 and IGF1 as adult height loci. Hum. Mol. Genet. *19*, 2303–2312.
13. Soranzo, N., Rivadeneira, F., Chinappen-Horsley, U., Malkina, I., Richards, J.B., Hammond, N., Stolk, L., Nica, A., Inouye, M., Hofman, A., et al. (2009). Meta-analysis of genome-wide scans for human adult stature identifies novel Loci and associations with measures of skeletal frame size. PLoS Genet. *5*, e1000445.
14. Weedon, M.N., Lango, H., Lindgren, C.M., Wallace, C., Evans, D.M., Mangino, M., Freathy, R.M., Perry, J.R., Stevens, S., Hall, A.S., et al; Diabetes Genetics Initiative; Wellcome Trust Case Control Consortium; Cambridge GEM Consortium. (2008). Genome-wide association analysis identifies 20 loci that influence adult height. Nat. Genet. *40*, 575–583.
15. Yang, J., Benyamin, B., McEvoy, B.P., Gordon, S., Henders, A.K., Nyholt, D.R., Madden, P.A., Heath, A.C., Martin, N.G., Montgomery, G.W., et al. (2010). Common SNPs explain

a large proportion of the heritability for human height. Nat. Genet. *42*, 565–569.
16. Keating, B.J., Tischfield, S., Murray, S.S., Bhangale, T., Price, T.S., Glessner, J.T., Galver, L., Barrett, J.C., Grant, S.F., Farlow, D.N., et al. (2008). Concept, design and implementation of a cardiovascular gene-centric 50 k SNP array for large-scale genomic association studies. PLoS ONE *3*, e3583.
17. Li, Y., Willer, C., Sanna, S., and Abecasis, G. (2009). Genotype imputation. Annu. Rev. Genomics Hum. Genet. *10*, 387–406.
18. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., de Bakker, P.I., Daly, M.J., and Sham, P.C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. *81*, 559–575.
19. Price, A.L., Patterson, N.J., Plenge, R.M., Weinblatt, M.E., Shadick, N.A., and Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. Nat. Genet. *38*, 904–909.
20. Price, A.L., Butler, J., Patterson, N., Capelli, C., Pascali, V.L., Scarnicci, F., Ruiz-Linares, A., Groop, L., Saetta, A.A., Korkolopoulou, P., et al. (2008). Discerning the ancestry of European Americans in genetic association studies. PLoS Genet. *4*, e236.
21. Willer, C.J., Li, Y., and Abecasis, G.R. (2010). METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics *26*, 2190–2191.
22. Musunuru, K., Lettre, G., Young, T., Farlow, D.N., Pirruccello, J.P., Ejebe, K.G., Keating, B.J., Yang, Q., Chen, M.H., Lapchyk, N., et al; NHLBI Candidate Gene Association Resource. (2010). Candidate gene association resource (CARe): design, methods, and proof of concept. Circ. Cardiovasc. Genet *3*, 267–275.
23. Johansen, C.T., Wang, J., Lanktree, M.B., Cao, H., McIntyre, A.D., Ban, M.R., Martins, R.A., Kennedy, B.A., Hassell, R.G., Visser, M.E., et al. (2010). Excess of rare variants in genes identified by genome-wide association study of hypertriglyceridemia. Nat. Genet. *42*, 684–687.
24. Price, A.L., Kryukov, G.V., de Bakker, P.I., Purcell, S.M., Staples, J., Wei, L.J., and Sunyaev, S.R. (2010). Pooled association tests for rare variants in exon-resequencing studies. Am. J. Hum. Genet. *86*, 832–838.
25. Sims, N.A., Jenkins, B.J., Nakamura, A., Quinn, J.M., Li, R., Gillespie, M.T., Ernst, M., Robb, L., and Martin, T.J. (2005). Interleukin-11 receptor signaling is required for normal bone remodeling. J. Bone Miner. Res. *20*, 1093–1102.
26. Takeuchi, Y., Watanabe, S., Ishii, G., Takeda, S., Nakayama, K., Fukumoto, S., Kaneta, Y., Inoue, D., Matsumoto, T., Harigaya, K., and Fujita, T. (2002). Interleukin-11 as a stimulatory factor for bone formation prevents bone loss with advancing age in mice. J. Biol. Chem. *277*, 49011–49018.
27. Lorda-Diez, C.I., Montero, J.A., Garcia-Porrero, J.A., and Hurle, J.M. (2010). Tgfbeta2 and 3 are coexpressed with their extracellular regulator Ltbp1 in the early limb bud and modulate mesodermal outgrowth and BMP signaling in chicken embryos. BMC Dev. Biol. *10*, 69.
28. Li, C.G., Liang, Q.Q., Zhou, Q., Menga, E., Cui, X.J., Shu, B., Zhou, C.J., Shi, Q., and Wang, Y.J. (2009). A continuous observation of the degenerative process in the intervertebral disc of Smad3 gene knock-out mice. Spine *34*, 1363–1369.
29. Voight, B.F., Scott, L.J., Steinthorsdottir, V., Morris, A.P., Dina, C., Welch, R.P., Zeggini, E., Huth, C., Aulchenko, Y.S., Thorleifsson, G., et al; MAGIC investigators; GIANT Consortium. (2010). Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. Nat. Genet. *42*, 579–589.
30. Teslovich, T.M., Musunuru, K., Smith, A.V., Edmondson, A.C., Stylianou, I.M., Koseki, M., Pirruccello, J.P., Ripatti, S., Chasman, D.I., Willer, C.J., et al. (2010). Biological, clinical and population relevance of 95 loci for blood lipids. Nature *466*, 707–713.
31. Ridker, P.M., Pare, G., Parker, A., Zee, R.Y., Danik, J.S., Buring, J.E., Kwiatkowski, D., Cook, N.R., Miletich, J.P., and Chasman, D.I. (2008). Loci related to metabolic-syndrome pathways including LEPR,HNF1A, IL6R, and GCKR associate with plasma C-reactive protein: the Women's Genome Health Study. Am. J. Hum. Genet. *82*, 1185–1192.
32. Kolz, M., Johnson, T., Sanna, S., Teumer, A., Vitart, V., Perola, M., Mangino, M., Albrecht, E., Wallace, C., Farrall, M., et al; EUROSPAN Consortium; ENGAGE Consortium; PROCARDIS Consortium; KORA Study; WTCCC. (2009). Meta-analysis of 28,141 individuals identifies common variants within five new loci that influence uric acid concentrations. PLoS Genet. *5*, e1000504.
33. Dupuis, J., Langenberg, C., Prokopenko, I., Saxena, R., Soranzo, N., Jackson, A.U., Wheeler, E., Glazer, N.L., Bouatia-Naji, N., Gloyn, A.L., et al; DIAGRAM Consortium; GIANT Consortium; Global BPgen Consortium; Anders Hamsten on behalf of Procardis Consortium; MAGIC investigators. (2010). New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. Nat. Genet. *42*, 105–116.
34. Johnson, A.D., Yanek, L.R., Chen, M.H., Faraday, N., Larson, M.G., Tofler, G., Lin, S.J., Kraja, A.T., Province, M.A., Yang, Q., et al. (2010). Genome-wide meta-analyses identifies seven loci associated with platelet aggregation in response to agonists. Nat. Genet. *42*, 608–613.
35. Holm, H., Gudbjartsson, D.F., Arnar, D.O., Thorleifsson, G., Thorgeirsson, G., Stefansdottir, H., Gudjonsson, S.A., Jonasdottir, A., Mathiesen, E.B., Njølstad, I., et al. (2010). Several common variants modulate heart rate, PR interval and QRS duration. Nat. Genet. *42*, 117–122.
36. Paajanen, T.A., Oksala, N.K., Kuukasjärvi, P., and Karhunen, P.J. (2010). Short stature is associated with coronary heart disease: a systematic review of the literature and a meta-analysis. Eur. Heart J. *31*, 1802–1809.