Lango Allen, Hana; Estrada, Karol; Lettre, Guillaume; Berndt, Sonja I; Weedon, Michael N; Rivadeneira, Fernando; Willer, Cristen J; Jackson, Anne U; Vedantam, Sailaja; Raychaudhuri, Soumya; +283 more... Ferreira, Teresa; Wood, Andrew R; Weyant, Robert J; Segrè, Ayellet V; Speliotes, Elizabeth K; Wheeler, Eleanor; Soranzo, Nicole; Park, Ju-Hyun; Yang, Jian; Gudbjartsson, Daniel; Heard-Costa, Nancy L; Randall, Joshua C; Qi, Lu; Vernon Smith, Albert; Mägi, Reedik; Pastinen, Tomi; Liang, Liming; Heid, Iris M; Luan, Jian'an; Thorleifsson, Gudmar; Winkler, Thomas W; Goddard, Michael E; Sin Lo, Ken; Palmer, Cameron; Workman, Tseggaselassie; Aulchenko, Yuiri S; Johannsson, Asa; Zillikens, M Carola; Feitosa, Mary F; Esko, Tönu; Johnson, Toby; Ketkar, Shamika; Kraft, Peter; Mangino, Massimo; Prokopenko, Inga; Absher, Devin; Albrecht, Eva; Ernst, Florian; Glazer, Nicole L; Hayward, Caroline; Hottenga, Jouke-Jan; Jacobs, Kevin B; Knowles, Joshua W; Kutilik, Zoltán; Monda, Keri L; Polasek, Ozren; Preuss, Michael; Rayner, Nigel W; Robertson, Neil R; Steinhorsdottir, Valdur; Tyrer, Jonathan P; Voight, Benjamin F; Wiklund, Fredrik; Xu, Jian-feng; Zhao, Jing Hua; Nyholt, Dale R; Pellikka, Niina; Perola, Markus; Perry, John RB; Surakka, Ida; Tamnessoo, Mari-Liis; Altmaier, Elizabeth L; Amin, Najaf; Aspelund, Thor; Bhangale, Tushar; Boucher, Gabrielle; Chasman, Daniel I; Chen, Constance; Coin, Lachlan; Cooper, Matthew N; Dixon, Anna L; Gibson, Quince; Grundberg, Elin; Hao, Ke; Juhani Junttila, M; Kaplan, Lee M; Kettunen, Johannes; König, Inke R; Kwan, Tony; Lawrence, Robert W; Levinson, Douglas F; Luetkemeier, James; McIntyre, Barbara; Morris, Andrew P; Müller, Martina; Suh Ngwa, Julius; Purcell, Shaun; Rafelt, Suzanne; Salem, Rany M; Salvi, Erika; Sanna, Serena; Shi, Jianxin; Sovio, Ulla; Thompson, John R; Turchin, Michael C; Vandenput, Liesbeth; Voight, Benjamin; Widen, Elisabeth; Wild, Sarah H; Willemse, Gonneke; Ziegler, Andreas; Almgren, Peter; Balmforth, Anthony J; Campbell, Harry; Citterio, Lorena; De Grandi, Alessandro; Dominiczak, Anna; Duan, Jubao; Elliott, Paul; Elinson, Andri; Eriksson, Johan G; Freimer, Nelson B; Geus, Ebo; Gloriozo, Antonella; Haiqing, Shen; Hartikainen, Anna-Liisa; Havulinna, Aki S; Hicks, Andrew A; Hui, Jennie; Igl, Wilmar; Illig, Thomas; Juntti, Antti; Kajantie, Eero; Kilpeläinen, Tuomas O; Koiranen, Markku; Kolcic, Ivana; Koskiniemi, Seppo; Kovacs, Peter; Laitinen, Naama; Liu, Jianjun; Lokki, Marja-Liisa; Marus, Ana; Maschio, Andrea; Meitinger, Thomas; Mulas, Antonella; Paré, Guillaume; Parker, Alex N; Peden, John F; Petersmann, Astrid; Pichler, Irene; Pieliläinen, Kirsi H; Pouta, Anneli; Ridderstråle, Martin; Rotter, Jerome I; Sambrook, Jennifer G; Sanders, Alan R; Schmidt, Carsten Oliver; Sinisalo, Juha; Smith, Jan H; Stringham, Heather M; Bragi Walters, G; Widen, Elisabeth; Wild, Sarah H; Willemsen, Gonneke; Zagona, Laura; Zgaga, Lina; Zitting, Paavo; Alaverde, Helene; Farrall, Martin; McArdf, Wendy L; Nelis, Mari; Peters, Marjolein J; Ripatti, Samuli; van Meurs, Joyce BJ; Aben, Katja K; Ardlie, Kristin G; Beckmann, Jacques S; Beilby, John P; Bergman, Richard N; Bergmann, Sven; Collins, Francis S; Cusi, Daniele; den Heijer, Martin; Eiriksdottir, Gudny; Gejman, Pablo V; Hall, Alistair S; Hamsten, Anders; Huikuri, Heikki V; Iribarren, Carlos; Kähönen, Mika; Kaprio, Jaakko; Kathiresan, Sekar; Kiemeney, Lambertus; Kocher,
(2010) Hundreds of variants clustered in genomic loci and biological pathways affect human height. Nature, 467 (7317). pp. 832-838. ISSN 0028-0836 DOI: https://doi.org/10.1038/nature09410

Downloaded from: http://researchonline.lshtm.ac.uk/id/eprint/2326/

DOI: https://doi.org/10.1038/nature09410

Usage Guidelines:

Please refer to usage guidelines at https://researchonline.lshtm.ac.uk/policies.html or alternatively contact researchonline@lshtm.ac.uk.

Available under license: http://creativecommons.org/licenses/by-nc-nd/2.5/
Hundreds of variants clustered in genomic loci and biological pathways affect human height

Hana Lango Allen1,*, Karol Estrada1,2,3,4,*, Guillaume Lettre5,6,*, Sonja I. Berndt7,*, Michael N. Weedon1,7, Fernando Rivadeneira1,2,3,4,*, Cristen J. Willer8, Anne U. Jackson9, Sailaja Vedantam10, Soumya Raychaudhuri11,12, Teresa Ferreira13, Andrew R. Wood1, Robert J. Weyant8, Ayellet V. Segre11,14,15, Elizabeth K. Speliotes10,16, Eleanor Wheeler17, Nicole Soranzo18,19, Ju-Hyun Park7, Jian Yang19, Daniel Gudbjartsson20, Nancy L. Heard-Costa21, Joshua C. Randall13, Lu Qi22,23, Albert Vernon Smith24,25, Reedik Mägi13, Tomi Pastinen26,27,28, Liming Liang29, Iris M. Heid30,31, Jian’an Luan32, Gudmar Thorleifsson20, Thomas W. Winkler30, Michael E. Goddard33,34, Ken Sin Lo5, Cameron Palmer3,10, Tsegaselassie Workalemahu22, Yuriiri Aulchenko2,4, Åsa Johansson35,36,1, M.Carola Zillikens3, Mary F. Feitosa37, Tõnu Esko38,39,40, Toby Johnson41,42,43,44, Shamika Ketka37, Peter Kraft45,46, Massimo Mangino18, Inga Prokopenko13,47, Devin Absher48, Eva Albrecht31, Florian Ernst49, Nicole L. Glazer50, Caroline Hayward51, Jouke-Jan Hottenga52, Kevin B. Jacobs53, Joshua W. Knowles54, Zoltán Kutalik41,42, Keri L. Monda55, Ozren Polasek56,57, Michael Preuss58, Nigel W. Rayner13,47, Neil R. Robertson13,47, Valgerdur Steinthorsdottir20, Jonathan F. Tyrer59, Benjamin F. Voight11,14,15, Fredrik Wiklund60, Jianfeng Xu61, Jing Hua Zhao32, Dale R. Nyholt62, Niina Pellikka63,64, Markus Perola63,64, John R.B. Perry1, Ida Surakka63,64, Mari-Liis Tammeoo36, Elizabeth L. Altmayer3,10, Najaf Amin2, Thor Aspelund24,25, Tushar Bhangale49, Gabrielle Boucher6, Daniel I.Chasman66,67, Constance Chen68, Lachlan Coin69, Matthew N. Cooper70, Anna L. Dixon71, Quince Gibson72, Elin Grundberg17,26,27, Ke Hao73, M. Juhani Junttila74, Lee M. Kaplan16,67,75, Johannes Kettunen63,64, Inke R. König58, Tony Kwan26,27, Robert W. Lawrence70, Douglas F. Levinson76, Mattias Lorentzon77, Barbara Mcknight78, Andrew P. Morris13, Martina Müller31,79,80, Julius Suh Ngwa81, Shaun Purcell14,82,83, Suzanne Rafelt84, Rany M. Salem9,10, Erika Salvi85,86, Serena Sanna87, Jianxin Shi7, Ulla Sovio69, John R. Thompson88,89, Michael C. Turchin9,10, Liesbeth Vandenput77, Dominique J. Verlaan26,27, Veronique Vitart51, Charles C. White81, Andreas Ziegler90, Peter Almgren91, Anthony J. Balmforth92, Harry Campbell83, Lorena Citterio94, Alessandro De Grandi95, Anna Dominiczak36, Jubaq Duan97, Paul Elliott97, Roberto Elosua98, Johan G. Eriksson99,100,101,102,103, Nelson B. Freimer104, Eco J.C. Geus52, Nicola Glorioso105, Shen Haïping72, Anna-Liisa Hartikainen106, Aki S. Havulinna107, Andrew A. Hicks95, Jennie Hui70,108,109, Wilmar IgI35, Thomas Illig31, Antti Jula110, Eero Kajantie100, Tuomas O. Kilpeläinen32, Maruku Koiranen111, Ivana Kolicz56, Seppo Koskinen107, Peter Kovacs112, Jaana Laitinen113, Jianjun Liu114, Marja-Liisa Lokki115, Ana Marusic116, Andrea Maschio87, Thomas Meitinger117,118, Antonella Mulas87, Gaëlle Dulac119, Alex N. Parker120, John F. Peden13,121, Astrid Petersmann122, Irene Pichler95, Kirsi H. Pietiläinen123,124, Anneli

Users may view, print, download and text-data mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use: http://www.nature.com/authors/editorial_policies/license.html#terms

Corresponding Authors: Michael Weedon, Genetics of Complex Traits, Peninsula College of Medicine and Dentistry, University of Exeter, Exeter, EX1 2LU, UK; michael.weedon@pms.ac.uk; Gonçalo Abecasis, Department of Biostatistics, Center for Statistical Genetics, University of Michigan, Ann Arbor, Michigan 48109, USA, gonalco@umich.edu; Kari Stefansson, deCODE Genetics, 101 Reykjavik, Iceland, kstefans@decode.is; Timothy Frayling, Genetics of Complex Traits, Peninsula College of Medicine and Dentistry, University of Exeter, Exeter, EX1 2LU, UK; tim.frayling@pms.ac.uk; Joel Hirschhorn, Children’s Hospital, Harvard Medical School, Broad Institute, Boston, Massachusetts 02115, USA, joelh@broadinstitute.org.

These authors contributed equally

Author Contributions: Full author contributions and roles are listed in the Supplementary Information.
Pouta et al.106,125 Martin Ridderstråle126, Jerome I. Rotter127, Jennifer G. Sambrook128,129, Alan R. Sanders97, Carsten Oliver Schmidt130, Juha Sinisalo131, Jan H. Smit132, Heather M. Stringham9, G. Bragi Walters20, Elisabeth Widen63, Sarah H. Wild93, Gonneke Willemsen52, Laura Zagato94, Lina Zgaga56, Paavo Zitting133, Helene Alavere38, Martin Farrall113,121,134, Wendy L. Mc Ardle135, Mari Nelis38,39,40, Marjolein J. Peters3,4, Samuli Ripatti63,64, Joyce B.J. van Meurs2,3,4, Katja K. Aben136, Kristin G Ardlie11, Jacques S. Beckmann41,137, John P. Beilby108,109,138, Richard N. Bergman139, Sven Bergmann41,42, Francis S. Collins140, Daniele Cusi85, Martin den Heijer141, Gudny Eiriksdottir24, Pablo V. Gejman97, Alistair S. Hall92, Anders Hamsten142, Heikki V. Huikuri74,74, Carlos Irizarren143,144, Mika Kähönen145, Jaakko Kaprio63,123,146, Sekar Kathiresan11,14,14,147,14,149, Lambertus Kiemeneij176,150,151, Thomas Kocher152, Lenore J. Launer153, Terho Lehtimäki154, Olle Melander126, Tom H. Mosley Jr155, Arthur W. Musk109,156, Markku S. Nieminen131,131, Christopher J. O’Donnell148,157, Claes Olsson7, Ben Oostra158, Lyle J. Palmer70,109, Olli Raitakari159, Paul M. Ridker86,67, John D. Rialoux5,6, Aila Rissanen124, Carlo Rivolta41,41, Heribert Schunkert160, Alan R. Shuldiner72,161, David S. Siscovick162,163, Michael Stumvoll164,165, Anke Tönjes164,166, Jaakko Tuomilehto167,168,169, Gert-Jan van Ommer170, Jorma Viikari171, Andrew C. Heath172, Nicholas G. Martin173, Grant W. Montgomery174, Michael A. Province37,175, Manfred Kayser176, Alice M. Arnold178,177, Larry D. Atwood21,21, Eric Boerwinkle178, Stephen J. Chanock7, Panos Deloukas17, Christian Gieger31, Henrik Grönberg60, Per Hall60, Andrew T. Hattersley1, Christian Hengstenberg179,180, Wolfgang Hoffman130, G. Mark Lathrop181, Veikko Salomaa107, Stefan Schreiber182, Manuela Uda87, Dawn Waterworth183, Alan F. Wright51, Themistocles L. Assimes54, Inês Barroso17,184, Albert Hofman2,4, Karen L. Mohlke185, Dorret I. Boomsma52, Mark J. Caulfield14, L. Adrienne Cupples81, Jeanette Erdmann160, Caroline S. Fox186, Vilmundur Gudnason24,25, Ulf Gyllensten35, Tamara B. Harris153, Richard B. Hayes187, Marjo-Riitta Jarvelin69,111,125,188, Vincent Mooser183, Patricia B. Munroe44, Willem H. Ouwend87,128,129, Brenda W. Penninx132,188,190, Peter P. Pramstaller95,191,192, Thomas Queeremous54,116, Niles J. Samani84,88, Timothy D. Spector18, Henry Völzk130, Hugh Watkins13,121,121 on behalf of Procardis Consortium, James F. Wilson93, Leif C. Groop91, Talin Haritunians127, Frank B. Hu22,23,45, Robert C. Kaplan193, Metspalu38,39,40, Kari E. North55,194, David Schlessinger195, Nicholas J. Wareham32, David J. Hunter22,23,45, Jeffrey R. O’Connell72, David P. Strachan196, H.-Erich Wichmann31,80,197, Ingrid B. Borecki37,175, Cornelia M. van Duijn2,4, Eric E. Schadt198,199, Unnur Thorsteinsdottir20,200, Leena Peltonen17,63,64,82,201, André Uitterlinden2,3,4, Peter M. Visscher9,109,111,134,135, Nilanjan Chatterjee2, Ruth J. F. Loos32, Michael Boehnke8, Mark I. McCarthy13,47,202, Erik Ingelsson66, Cecilia M. Lindgren136,137,138, Gonzalez R. Abecasis8,9, Kari Stefansson20,200,13, Timothy M. Frayling11, and Joel N Hirschhorn9,10,203,* for the GIANT Consortium

1 Genetics of Complex Traits, Peninsula College of Medicine and Dentistry, University of Exeter, Exeter, EX1 2LU, UK 2 Department of Epidemiology, Erasmus MC, Rotterdam, 3015GE, The Netherlands 3 Department of Internal Medicine, Erasmus MC, Rotterdam, 3015GE, The Netherlands 4 Netherlands Genomics Initiative (NGI)-sponsored Netherlands Consortium for Healthy Aging (NCHA) 5 Montreal Heart Institute, Montreal, Quebec, H1T 1C8, Canada 6 Department of Medicine, Université de Montréal, Montreal, Quebec, H3T 1J4, Canada 7 Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Department of Health and Human Services, Bethesda, Maryland 20892, USA 8 Department of Biostatistics, Center for Statistical Genetics, University of Michigan, Ann Arbor, Michigan 48109, USA 9 Divisions of Genetics and Endocrinology and Program in Genomics, Children’s Hospital, Boston, Massachusetts 02115, USA 10 Metabolism Initiative and Program in Medical and Population Genetics, Broad Institute, Cambridge, Massachusetts 02142, USA 11 Program in Medical and Population Genetics, Broad Institute of Harvard and Massachusetts Institute of Technology, Cambridge, Massachusetts 02142, USA 12 Division of Rheumatology, Immunology

Nature. Author manuscript; available in PMC 2011 April 14.
Stampar School of Public Health, Medical School, University of Zagreb, 10000 Zagreb, Croatia 57
Gen-Info Ltd, 10000 Zagreb, Croatia 58 Universität zu Lübeck, Institut für Medizinische Biometrie und Statistik, 23562 Lübeck, Germany 59 Department of Oncology, University of Cambridge, Cambridge, CB1 8RN, UK 60 Department of Medical Epidemiology and Biostatistics, Karolinska Institute, 171 77 Stockholm, Sweden 61 Center for Human Genomics, Wake Forest University, Winston-Salem, North Carolina 27157, USA 62 Neurogenetics Laboratory, Queensland Institute of Medical Research, Queensland 4006, Australia 63 Institute for Molecular Medicine Finland (FIMM), University of Helsinki, 00014, Helsinki, Finland 64 National Institute for Health and Welfare, Department of Chronic Disease Prevention, Unit of Public Health Genomics, 00014, Helsinki, Finland 65 Department of Genome Sciences, University of Washington, Seattle, 98195 Washington, USA 66 Division of Preventive Medicine, Brigham and Women's Hospital, Boston, Massachusetts 02215, USA 67 Harvard Medical School, Boston, Massachusetts 02115, USA 68 Program in Molecular and Genetic Epidemiology, Harvard School of Public Health, Boston, Massachusetts 02115, USA 69 Department of Epidemiology and Biostatistics, School of Public Health, Faculty of Medicine, Imperial College London, London, W2 1PG, UK 70 Centre for Genetic Epidemiology and Biostatistics, University of Western Australia, Crawley, Western Australia 6009, Australia 71 Royal National Hospital for Rheumatic Diseases and University of Bath, Bath, BA1 1RL, UK 72 Department of Medicine, University of Maryland School of Medicine, Baltimore, Maryland 21201, USA 73 Genetics Department, Rosetta Inpharmatics, a Wholly Owned Subsidiary of Merck & Co. Inc., Seattle, Washington 98109, USA 74 Department of Internal Medicine, University of Oulu, 90014 Oulu, Finland 75 MGH Weight Center, Massachusetts General Hospital, Boston, Massachusetts 02114, USA 76 Stanford University School of Medicine, Stanford, California 94305, USA 77 Department of Internal Medicine, Institute of Medicine, Sahlgrenska Academy, University of Gothenburg, 413 45 Gothenburg, Sweden 78 Departments of Biostatistics, University of Washington, Seattle, Washington 98195, USA 79 Ludwig-Maximilians-University, Department of Medicine I, University Hospital Grosshadern, 81377 Munich, Germany 80 Ludwig-Maximilians-Universität, Institute of Medical Informatics, Biometry and Epidemiology, Chair of Epidemiology, 81377 Munich, Germany 81 Department of Biostatistics, Boston University School of Public Health, Boston, Massachusetts 02118, USA 82 The Broad Institute of Harvard and MIT, Cambridge, Massachusetts 02142, USA 83 Department of Psychiatry, Harvard Medical School, Boston, Massachusetts 02115, USA 84 Department of Cardiovascular Sciences, University of Leicester, Glenfield Hospital, Leicester, LE3 9QP, UK 85 University of Milan, Department of Medicine, Surgery and Dentistry, 20139 Milano, Italy 86 KOS Genetic Srl, 20123 Milan, Italy 87 Istituto di Neurogenetica e Neurofarmacologia del CNR, Monserrato, 09042, Cagliari, Italy 88 Leicester NIHR Biomedical Research Unit in Cardiovascular Disease, Glenfield Hospital, Leicester, LE3 9QP, UK 89 Department of Health Sciences, University of Leicester, University Road, Leicester, LE1 7RH, UK 90 Universität zu Lübeck, Institut für Medizinische Biometrie und Statistik, 23562 Lübeck, Germany 91 Lund University Diabetes Centre, Department of Clinical Sciences, Lund University, 20502 Malmö, Sweden 92 Multidisciplinary Cardiovascular Research Centre (MCRC), Leeds Institute of Genetics, Health and Therapeutics (LIGHT), University of Leeds, Leeds LS2 9JT, UK 93 Centre for Population Health Sciences, University of Edinburgh, Teviot Place, Edinburgh, EH8 9AG, Scotland 94 University Vita-Salute San Raffaele, Division of Nephrology and Dialysis, 20132 Milan, Italy 95 Institute of Genetic Medicine, European Academy Bozen/Bolzano (EURAC), Bolzano/Bozen, 39100, Italy. Affiliated Institute of the University of Lübeck, Lübeck, Germany 96 British Heart Foundation Glasgow Cardiovascular Research Centre, University of Glasgow, Glasgow, G12 8TA, UK 97 North shore University Healthsystem, Evanston, Illinois 60201, USA 98 Cardiovascular Epidemiology and Genetics, Institut Municipal D’investigacio Medica and CIBER Epidemiología y Salud Pública, Barcelona, Spain 99 Department of General Practice and Primary health Care, University of Helsinki, Helsinki, Finland 100 National Institute for Health and Welfare, 00271 Helsinki, Finland 101 Helsinki University Central Hospital, Unit of General Practice, 00280 Helsinki, Finland 102 Folkhalsan
Research Centre, 00250 Helsinki, Finland 103 Vasa Central Hospital, 65130 Vasa, Finland 104 Center for Neurobehavioral Genetics, University of California, Los Angeles, California 90095, USA 105 Hypertension and Cardiovascular Prevention Center, University of Sassari, 07100 Sassari, Italy 106 Department of Clinical Sciences/Obstetrics and Gynecology, University of Oulu, 90014 Oulu, Finland 107 National Institute for Health and Welfare, Department of Chronic Disease Prevention, Chronic Disease Epidemiology and Prevention Unit, 00014, Helsinki, Finland 108 PathWest Laboratory of Western Australia, Department of Molecular Genetics, J Block, QEII Medical Centre, Nedlands, Western Australia 6009, Australia 109 Busselton Population Medical Research Foundation Inc., Sir Charles Gairdner Hospital, Nedlands, Western Australia 6009, Australia 110 National Institute for Health and Welfare, Department of Chronic Disease Prevention, Population Studies Unit, 20720 Turku, Finland 111 Institute of Health Sciences, University of Oulu, 90014 Oulu, Finland 112 Interdisciplinary Centre for Clinical Research, University of Leipzig, 04103 Leipzig, Germany 113 Finnish Institute of Occupational Health, 90220 Oulu, Finland 114 Human Genetics, Genome Institute of Singapore, Singapore 138672, Singapore 115 Transplantation Laboratory, Haartman Institute, University of Helsinki, 00014, Helsinki, Finland 116 Croatian Centre for Global Health, School of Medicine, University of Split, Split 21000, Croatia 117 Institute of Human Genetics, Klinikum rechts der Isar der Technischen Universität München, 81675 Munich, Germany 118 Institute of Human Genetics, Helmholtz Zentrum München - German Research Center for Environmental Health, 85764 Neuherberg, Germany 119 Department of Pathology and Molecular Medicine, McMaster University, Hamilton, Ontario L8N3Z5, Canada 120 Amgen, Cambridge, Massachusetts 02139, USA 121 Department of Cardiovascular Medicine, University of Oxford, Level 6 West Wing, John Radcliffe Hospital, Headington, Oxford, OX3 9DU 122 Institut für Klinische Chemie und Laboratoriumsmedizin, Universität Greifswald, 17475 Greifswald, Germany 123 Finnish Twin Cohort Study, Department of Public Health, University of Helsinki, 00014, Helsinki, Finland 124 Obesity Research unit, Department of Psychiatry, Helsinki University Central Hospital, Helsinki, Finland 125 National Institute for Health and Welfare, 90101 Oulu, Finland 126 Department of Clinical Sciences, Lund University, 20502 Malmö, Sweden 127 Medical Genetics Institute, Cedars-Sinai Medical Center, Los Angeles, California 90048, USA 128 Department of Haematology, University of Cambridge, Cambridge CB2 0PT, UK 129 NHS Blood and Transplant, Cambridge Centre, Cambridge, CB2 0PT, UK 130 Institut für Community Medicine, 17489 Greifswald, Germany 131 Division of Cardiology, Cardiovascular Laboratory, Helsinki University Central Hospital, 00029 Helsinki, Finland 132 Department of Psychiatry/EMGO Institute, VU University Medical Center, 1081 BT Amsterdam, The Netherlands 133 Department of Psychiatrics, Lapland Central Hospital, 96101 Rovaniemi, Finland 134 Cardiovascular Medicine, University of Oxford, Wellcome Trust Centre for Human Genetics, Oxford, OX3 7BN, UK 135 Avon Longitudinal Study of Parents and Children (ALSPAC) Laboratory, Department of Social Medicine, University of Bristol, Bristol, BS8 2BN, UK 136 Comprehensive Cancer Center East, 6501 BG Nijmegen, The Netherlands 137 Service of Medical Genetics, Centre Hospitalier Universitaire Vaudois (CHUV) University Hospital, 1011 Lausanne, Switzerland 138 School of Pathology and Laboratory Medicine, University of Western Australia, Nedlands, Western Australia 6009, Australia 139 Department of Physiology and Biophysics, Keck School of Medicine, University of Southern California, Los Angeles, California 90033, USA 140 National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland 20892, USA 141 Department of Endocrinology, Radboud University Nijmegen Medical Centre, 6500 HB Nijmegen, The Netherlands 142 Atherosclerosis Research Unit, Department of Medicine, Solna, Karolinska Institutet, Karolinska University Hospital, 171 76 Stockholm, Sweden 143 Division of Research, Kaiser Permanente Northern California, Oakland, California 94612, USA 144 Department of Epidemiology and Biostatistics, University of California, San Francisco, San Francisco, California 94107, USA 145 Department of Clinical Physiology, University of Tampere and Tampere University Hospital, 33520 Tampere, Finland 146 National Institute for Health and Welfare, Department of Mental Health and Substance Abuse Services, Unit for Child and

Nature. Author manuscript; available in PMC 2011 April 14.
Abstract

Most common human traits and diseases have a polygenic pattern of inheritance: DNA sequence variants at many genetic loci influence phenotype. Genome-wide association (GWA) studies have identified >600 variants associated with human traits, but these typically explain small fractions of phenotypic variation, raising questions about the utility of further studies. Here, using 183,727 individuals, we show that hundreds of genetic variants, in at least 180 loci, influence adult height, a highly heritable and classic polygenic trait. The large number of loci reveals patterns with important implications for genetic studies of common human diseases and traits. First, the 180 loci are not random, but instead are enriched for genes that are connected in biological pathways (P=0.016), and that underlie skeletal growth defects (P<0.001). Second, the likely causal gene is often located near the most strongly associated variant: in 13 of 21 loci containing a known skeletal growth gene, that gene was closest to the associated variant. Third, at least 19 loci have multiple independently associated variants, suggesting that allelic heterogeneity is a frequent feature of polygenic traits, that comprehensive explorations of already-discovered loci should discover additional variants, and that an appreciable fraction of associated loci may have been identified. Fourth, associated variants are enriched for likely functional effects on genes, being over-represented amongst variants that alter amino acid structure of proteins and expression levels of nearby genes. Our data explain ∼10% of the phenotypic variation in height, and we estimate that unidentified common variants of similar effect sizes would increase this figure to ∼16% of phenotypic variation (∼20% of heritable variation). Although additional approaches are needed to fully dissect the genetic architecture of polygenic human traits, our findings indicate that GWA studies can identify large numbers of loci that implicate biologically relevant genes and pathways.

In Stage 1 of our study, we performed a meta-analysis of GWA data from 46 studies, comprising 133,653 individuals of recent European ancestry, to identify common genetic variation associated with adult height. To enable meta-analysis of studies across different genotyping platforms, we performed imputation of 2,834,208 single nucleotide polymorphisms (SNPs) present in the HapMap Phase 2 European-American reference panel. After applying quality control filters, each individual study tested the association of adult height with each SNP using an additive model (Supplementary Methods). The individual study statistics were corrected using the genomic control (GC) method and then combined in a fixed effects based meta-analysis. We then applied a second GC correction on the meta-analysis statistics, although this approach may be overly conservative when there are many real signals of association (Supplementary Methods). We detected 207 loci (defined as 1Mb on either side of the most strongly associated SNP) as potentially associated with adult height (P<5×10^-6).
To identify loci robustly associated with adult height, we took forward at least one SNP (Supplementary Methods) from each of the 207 loci reaching $P<5\times10^{-6}$ into an additional 50,074 samples (Stage 2) that became available after completion of our initial meta-analysis. In the joint analysis of our Stage 1 and Stage 2 studies, SNPs representing 180 loci reached genome-wide significance ($P<5\times10^{-8}$; Supplementary Figures 1 and 2, Supplementary Table 1). Additional tests, including genotyping of a randomly-selected subset of 33 SNPs in an independent sample of individuals from the 5th-10th and 90th-95th percentiles of the height distribution (N=3,190), provided further validation of our results, with all but two SNPs showing consistent direction of effect (sign test $P<7\times10^{-8}$) (Supplementary Methods, Supplementary Table 2).

Genome wide association studies can be susceptible to false positive associations from population stratification. We therefore performed a family-based analysis, which is immune to population stratification in 7,336 individuals from two cohorts with pedigree information. Alleles representing 150 of the 180 genome-wide significant loci were associated in the expected direction (sign test $P<6\times10^{-20}$; Supplementary Table 3). The estimated effects on height were essentially identical in the overall meta-analysis and the family-based sample. Together with several other lines of evidence (Supplementary Methods), this indicates that stratification is not substantially inflating the test statistics in our meta-analysis.

Common genetic variants have typically explained only a small proportion of the heritable component of phenotypic variation. This is particularly true for height, where >80% of the variation within a given population is estimated to be attributable to additive genetic factors, but over 40 previously published variants explain <5% of the variance. One possible explanation is that many common variants of small effects contribute to phenotypic variation, and current GWA studies remain underpowered to detect the majority of common variants. Using five studies not included in Stage 1, we found that the 180 associated SNPs explained on average 10.5% (range 7.9-11.2%) of the variance in adult height (Supplementary Methods). Including SNPs associated with height at lower significance levels ($0.05>P>5\times10^{-8}$) increased the variance explained to 13.3% (range 9.7-16.8%) (Figure 1a). In addition, we found no evidence that non-additive effects including gene-gene interaction would increase the proportion of the phenotypic variance explained (Supplementary Methods, Supplementary Tables 5 and 6).

As a separate approach, we used a recently developed method to estimate the total number of independent height-associated variants with effect sizes similar to the ones identified. We obtained this estimate using the distribution of effect sizes observed in Stage 2 and the power to detect an association in Stage 1, given these effect sizes (Supplementary Methods). The cumulative distribution of height loci, including those we identified and others as yet undetected, is shown in Figure 1b. We estimate that there are 697 loci (95% confidence interval (CI): 483-1040) with effects equal or greater than those identified, which together would explain approximately 15.7% of the phenotypic variation in height or 19.6% (95% CI: 16.2-25.6) of height heritability (Supplementary Table 4). We estimated that a sample size of 500,000 would detect 99.6% of these loci at $P<5\times10^{-5}$. This figure does not account for variants that have effect sizes smaller than those observed in the current study and, therefore, underestimates the contribution of undiscovered common loci to phenotypic variation.

A further possible source of missing heritability is allelic heterogeneity – the presence of multiple, independent variants influencing a trait at the same locus. We performed genome-wide conditional analyses in a subset of Stage 1 studies, including a total of 106,336 individuals. Each study repeated the primary GWA analysis but additionally adjusted for
SNPs representing the 180 loci associated at $P<5\times10^{-6}$ (Supplementary Methods). We then meta-analysed these studies in the same way as for the primary GWA study meta-analysis. Nineteen SNPs within the 180 loci were associated with height at $P<3.3\times10^{-7}$ (a Bonferroni-corrected significance threshold calculated from the ~15% of the genome covered by the conditioned 2Mb loci; Supplementary Methods, Table 1, Figure 2, Supplementary Figure 3). The distances of the second signals to the lead SNPs suggested that both are likely to be affecting the same gene, rather than being coincidentally in close proximity. At 17 of 17 loci (excluding two contiguous loci in the *HMGA1* region), the second signal occurred within 500kb, rather than between 500kb and 1 Mb, of this lead SNP (binomial test $P=2\times10^{-5}$). Further analyses of allelic heterogeneity may identify additional variants that increase the proportion of variance explained. For example, within the 180 2Mb loci, a total of 45 independent SNPs reached $P<1\times10^{-5}$ when we would expect <2 by chance.

Whilst GWA studies have identified many variants robustly associated with common human diseases and traits, the biological significance of these variants, and the genes on which they act, is often unclear. We first tested the overlap between the 180 height-associated variants and two types of putatively functional variants, nonsynonymous (ns) SNPs and cis-eQTLs (variants strongly associated with expression of nearby genes). Height variants were 2.4-fold more likely to overlap with cis-eQTLs in lymphocytes than expected by chance (47 variants: $P=4.7\times10^{-11}$) (Supplementary Table 7) and 1.7-fold more likely to be closely correlated ($r^2 \geq 0.8$ in HapMap CEU) with nsSNPs (24 variants $P=0.004$) (Supplementary Methods, Supplementary Table 8). Although the presence of a correlated eQTL or nsSNP at an individual locus does not establish the causality of any particular variant, this enrichment shows that common functional variants contribute to the causal variants at height-associated loci. We also noted five loci where the height associated variant was strongly correlated ($r^2>0.8$) with variants associated with other traits and diseases ($P<5\times10^{-8}$), including bone mineral density, rheumatoid arthritis, type 1 diabetes, psoriasis and obesity, suggesting that these variants have pleiotropic effects on human phenotypes (Supplementary Methods; Supplementary Table 9).

We next addressed the extent to which height variants cluster near biologically relevant genes; specifically, genes mutated in human syndromes characterized by abnormal skeletal growth. We limited this analysis to the 652 genes occurring within the recombination hotspot-bounded regions surrounding each of the 180 index SNPs. We showed that the 180 loci associated with variation in normal height contained 21 of 241 genes (8.7%) found to underlie such syndromes (Supplementary Table 10), compared to a median of 8 (range 1-19) genes identified in 1,000 matched control sets of regions ($P<0.001$: 0 observations of 21 or more skeletal growth genes among 1,000 sets of matched SNPs). In 13 of these 21 loci the closest gene to the most associated height SNP in the region is the growth disorder gene, and in 9 of these cases, the most strongly associated height SNP is located within the growth disorder gene itself (Supplementary Methods, Supplementary Table 11). These results suggest that GWA studies may provide more clues about the identity of the functional genes at each locus than previously suspected.

We also investigated whether significant and relevant biological connections exist between the genes within the 180 loci, using two different computational approaches. We used the GRAIL text-mining algorithm to search for connectivity between genes near the associated SNPs, based on existing literature20. Of the 180 loci, 42 contained genes that were connected by existing literature to genes in the other associated loci (the pair of connected genes appear in articles that share scientific terms more often than expected at $P<0.01$). For comparison, when we used GRAIL to score 1,000 sets of 180 SNPs not associated with height (but matched for number of nearby genes, gene proximity, and allele frequency), we only observed 16 sets with 42 or more loci with a connectivity $P<0.01$, thus providing strong
statistical evidence that the height loci are functionally related (P=0.016) (Figure 3a). For the 42 regions with GRAIL connectivity P<0.01, the implicated genes and SNPs are highlighted in Figure 3b. The most strongly connected genes include those in the Hedgehog, TGF-beta, and growth hormone pathways.

As a second approach to find biological connections, we applied a novel implementation of gene set enrichment analysis (GSEA) (Meta-Analysis Gene-set Enrichment of variaNT Associations, MAGENTA21) to perform pathway analysis (Supplementary Methods). This analysis revealed 17 different biological pathways and 14 molecular functions nominally enriched (P<0.05) for associated genes, many of which lie within the validated height loci. These gene-sets include previously reported11,13 (e.g. Hedgehog signaling) and novel (e.g. TGF-beta signaling, histones, and growth and development-related) pathways and molecular functions (Supplementary Table 12). Several SNPs near genes in these pathways narrowly missed genome-wide significance, suggesting that these pathways likely contain additional associated variants. These results provide complementary evidence for some of the genes and pathways highlighted in the GRAIL analysis. For instance, genes such as TGFB2 and LTBP1-3 highlight a role for the TGF-beta signaling pathway in regulating human height, consistent with the implication of this pathway in Marfan syndrome22.

Finally, to examine the evidence for the potential involvement of specific genes at individual loci, we aggregated evidence from our data (eQTLs, proximity to the associated variant, pathway-based analyses), and human and mouse genetic databases (Supplementary Table 13). Of 32 genes with highly correlated (r2>0.8) nsSNPs, several are newly identified strong candidates for playing a role in human growth. Some are in pathways enriched in our study (such as ECM2, implicated in extracellular matrix), while others have similar functions to known growth-related genes, including FGFR4 (FGFR3 underlies several classic skeletal dysplasias23) and STAT2 (STAT5B mutations cause growth defects in humans24). Interestingly, Fgfr4-/- Fgfr3-/- mice show severe growth retardation not seen in either single mutant25, suggesting that the FGFR4 variant might modify FGFR3-mediated skeletal dysplasias. Other genes at associated loci, such as NPPC and NPR3 (encoding the C-type natriuretic peptide and its receptor), influence skeletal growth in mice and will likely also influence human growth17. Many of the remaining 180 loci have no genes with obvious connections to growth biology, but at some our data provide modest supporting evidence for particular genes, including C3orf63, PML, CCDC91, ZNF1, ID4, RYBP, SEPT2, ANKRD13B, FOLH1, LRRC37B, MFAP2, SLBP, SOCS5, and ZBTB24 (Supplementary Table 13).

We have identified >100 novel loci that influence the classic polygenic trait of normal variation in human height, bringing the total to 180. Our results have potential general implications for genetic studies of complex traits. We show that loci identified by GWA studies highlight relevant genes: the 180 loci associated with height are non-randomly clustered within biologically relevant pathways and are enriched for genes that are involved in growth-related processes, that underlie syndromes of abnormal skeletal growth, and that are directly relevant to growth-modulating therapies (GH1, IGF1R, CYP19A1, ESR1). The large number of loci with clearly relevant genes suggests that the remaining loci could provide potential clues to important and novel biology.

We provide the strongest evidence yet that the causal gene will often be located near the most strongly associated DNA sequence variant. At the 21 loci containing a known growth disorder gene, that gene was on average 81 kb from the associated variant, and in over half of the loci it was the closest gene to the associated variant. Despite recent doubts about the benefits of GWA studies26, this finding suggests that GWA studies are useful mapping tools.
to highlight genes that merit further study. The presence of multiple variants within
associated loci could help localize the relevant genes within these loci.

By increasing our sample size to >100,000 individuals, we identified common variants that
account for approximately 10% of phenotypic variation. Although larger than predicted by
some models, this figure suggests that GWA studies, as currently implemented, will not
explain a majority of the estimated 80% contribution of genetic factors to variation in height.
This conclusion supports the idea that biological insights, rather than predictive power, will
be the main outcome of this initial wave of GWA studies, and that new approaches, which
could include sequencing studies or GWA studies targeting variants of lower frequency, will
be needed to account for more of the “missing” heritability. Our finding that many loci
exhibit allelic heterogeneity suggests that many as yet unidentified causal variants, including
common variants, will map to the loci already identified in GWA studies, and that the
fraction of causal loci that have been identified could be substantially greater than the
fraction of causal variants that have been identified.

In our study, many associated variants are tightly correlated with common nsSNPs, which
would not be expected if these associated common variants were proxies for collections of
rare causal variants, as has been proposed. Although a substantial contribution to
heritability by less common and/or quite rare variants may be more plausible, our data are
not inconsistent with the recent suggestion that a large number of common variants of
very small effect mostly explain the regulation of height.

In summary, our findings indicate that additional approaches, including those aimed at less
common variants, will likely be needed to dissect more completely the genetic component to
complex human traits. Our results also strongly demonstrate that GWA studies can identify
large numbers of loci that together implicate biologically relevant pathways and
mechanisms. We envision that thorough exploration of the genes at associated loci through
additional genetic, functional, and computational studies will lead to novel insights into
human height and other polygenic traits and diseases.

Methods summary

The primary meta-analysis (Stage 1) included 46 GWA studies of 133,653 individuals. The
in-silico follow up (Stage 2) included 15 studies of 50,074 individuals. All individuals were
of European ancestry and >99.8% were adults. Details of genotyping, quality control, and
imputation methods of each study are given in Supplementary Methods Table 1-2. Each
study provided summary results of a linear regression of age-adjusted, within-sex Z scores
of height against the imputed SNPs, and an inverse-variance meta-analysis was performed in
METAL (http://www.sph.umich.edu/csg/abecasis/METAL). Validation of selected SNPs
was performed through direct genotyping in an extreme height panel (N=3,190) using
Sequenom iPlexX, and in 492 Stage 1 samples using the KASPar SNP System. Family-based
testing was performed using QFAM, a linear regression-based approach that uses
permutation to account for dependency between related individuals, and FBAT, which
uses a linear combination of offspring genotypes and traits to determine the test statistic.
We used a previously described method to estimate the amount of genetic variance
explained by the nominally associated loci (using significance threshold increments from
P=5×10^{-8} to P=0.05). To predict the number of height susceptibility loci, we took the
height loci that reached a significance level of P=5×10^{-8} in Stage 1 and estimated the
number of height loci that are likely to exist based on the distribution of their effect sizes
observed in Stage 2 and the power to detect their association in Stage 1. Gene-by-gene
interaction, dominant, recessive and conditional analyses are described in Supplementary
Methods. Empirical assessment of enrichment for coding SNPs used permutations of

Nature. Author manuscript; available in PMC 2011 April 14.
random sets of SNPs matched to the 180 height-associated SNPs on the number of nearby genes, gene proximity, and minor allele frequency. GRAIL and GSEA methods have been described previously20-21. To assess possible enrichment for genes known to be mutated in severe growth defects, we identified such genes in the OMIM database (Supplementary Table 10), and evaluated the extent of their overlap with the 180 height-associated regions through comparisons with 1000 random sets of regions with similar gene content (+10%).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

A number of participating studies are members of CHARGE and ENGAGE consortia. We acknowledge funding from the Academy of Finland (104781, 117797, 120315, 121584, 126925, 129494, 129680, 213506); Affymetrix, Inc for genotyping services (N02-HL-6-4278); Agency for Science, Technology and Research of Singapore (A*STAR); ALF/LUA Gothenburg; Althingi (the Icelandic Parliament); Amgen; AstraZeneca AB; Australian National Health and Medical Research Council (241944, 389875, 389891, 389902, 389938, 429215, 4D00336, R01-1096688, 55248, 635672); Austrian Research Council (DK03762); Biocentrum Helsinki; Boston Obesity Nutrition Research Center (DK46200); British Diabetes Association; British Heart Foundation (PG/02/128); British Heart Foundation Centre for Research Excellence, Oxford; CamStrad; Cancer Research UK; Centre for Neurogenomics and Cognitive Research (CNCR-VU); Chief Scientist Office of the Scottish Government (CSO) (CZB/4/279); Council of Health of the Academy of Finland; DIAB Core project of the German Network of Diabetes; Diabetes UK; Donald W. Reynolds Foundation; Emil and Vera Cornell Foundation; Erasmus MC, Estonian Government (SF0180142); European Commission (201413, ECOGENE:205419, BBMRI:212111, OPGENE:245336, ENGAGE:HEALTH-F4:2007-201413, EUWLRT:2001-01254, HEALTH-F2:2008-ENGAGE, HEALTH-F4:2007-201550, LSHG-CT-2006-018947, LSHG-CT-2006-01947, Procaccis:LSHM-CT-2007-037273, POLYGENE:LSHC-CT-2005, QLGI-CT-2000-01643, QLG2-CT-2002-01254, DG XII, Marie Curie Intra-European Fellowship); Epple Appeal; Finish Ministry of Education; Finnish Diabetes Research Foundation; Finnish Diabetes Research Society; Finnish Foundation for Cardiovascular Research; Finnish Medical Society; Finska Läkareassistenten; Folkhälsan Research Foundation; Fondation LeDucq; Foundation for Life and Health in Finland; Foundation for Strategic Research (SSF); GEN-AU-Programme “GOLD”; Genetic Association InformationNetwork (GAIN); German Bundesministerium fuer Forschung und Technologie (01 AK 803 A-H, 01 IG 07015 G); German Federal Ministry of Education and Research (BMBF) (01GS0831); German Ministry for Health, Welfare and Sports; German Ministry of Cultural Affairs; German Ministry of Education, Culture and Science; German National Genome Research Net (NGFN2 and NGFNplus) (01GS0823, 01ZZ1039, 01ZZ20403, 01ZZ6039, 03ZIK001); German Research Council (KFO-152); GlaxoSmithKline; Göteborg Medical Society; Gyllenberg Foundation; Helmholtz Center Munich; Juvenile Diabetes Research Foundation International (JDRF) (U01 DK062418); Karolinska Institute; Knut and Alice Wallenberg Foundation; Lundberg Foundation; March of Dimes (6-FY-09-507); MC Health; Medical Research Council UK (G0000649, G0000934, G0500359, G0600331, G0610216, G9521010D, PrevMetSyn); Microarray Core Facility of the Interdisciplinary Centre for Clinical Research (IZKF) (B27); Mid-Atlantic Nutrition and Obesity Research Center of Maryland (P30 DK072488); Ministry of Health and Department of Educational Assistance (South Tyrol, Italy); Ministry of Science, Education and Sport of the Republic of Croatia (216-1080315-0302); Montreal Heart Institute Foundation; Närpes Health Care Foundation; National Institute for Health and Research (NIHR) Cambridge Biomedical Research Centre; NIHR Oxford Biomedical Research Centre; NIHR comprehensive Biomedical Research Centre; National Institutes of Health (263-MA-410953, AA014041, AA07535, AA10248, AA13320, AA13326, CA047988, CA94499, CA50385, CA65725, CA67262, CA87969, DA128370, DK063491, DK072193, DK079466, DK080145, DK58845, HG002651, HG005214, HG005581, HL043851, HL084729, HL69757, HL71981, K08-AR055688, K23-DK080145, K99-HL094535, M01-RR00425, MH084698, N01-AG12100, N01-AG12109, N01-HC15103, N01-HC25195, N01-HC55129, N01-HC55135, N01-HC55015, N01-HC55016, N01-HC55018 through N01-HC55022, N01-HC55222, N01-HC55150, N01-HC55079 through N01-HC55086, N01-HG65403, N01-AG031890, R01 CA104021, R01-DK08336, R01-DK073490, R01-DK075681, R01-DK075787, R01-HL06694, R01-HL087641, R01-HL087647, R01-HL087652, R01-HL087676, R01-HL087679, R01-HL087700, R01-HL08819, R01-HL095367, R01-MH059160, R01-MH59565, R01-MH59566, R01-MH59571, R01-MH59586, R01-MH59587, R01-MH59588, R01-MH60870, R01-MH60879, R01-MH61675, R01-MH63706, R01-MH67257, R01-MH79469, R01-MH81800, RL1-MH083268, T32-HG000040, U01-CA098233, U01-GM074518, U01-HG00399, U01-HG00402, U01-HL080295, U01-HL084756, U01-HL72515, U01-MH79469, U01-MH79470, U54-RR020278, U1L-RR025005, Z01-AG00675, Z01-AG007380, Z01-HG000024; contract HHSN268200625226C; ADA Mentor-Based Postdoctoral Fellowship; Pew Scholarship for the Biomedical Sciences); Netherlands Genomics Initiative (NGI)/Netherlands Consortium for Healthy Aging (NCHA) (050-060-810); Netherlands Organisation for Scientific Research (NWO) (Investment nr. 175.010.2005.011, 911-03-012); Netherlands Organization for the Health

Nature. Author manuscript; available in PMC 2011 April 14.
References

1. Hindorff LA, et al. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. Proc Natl Acad Sci U S A. 2009; 106:9362–7. [PubMed: 19474294]

2. Galton F. Regression towards mediocrity in hereditary stature. J R Anthropol Inst. 1885; 5:329–348.

3. Fisher RA. The Correlation Between Relatives on the Supposition of Mendelian Inheritance. Transactions of the Royal Society of Edinburgh. 1918; 52:399–433.

4. Frazer KA, et al. A second generation human haplotype map of over 3.1 million SNPs. Nature. 2007; 449:851–61. [PubMed: 17943122]

5. Devlin B, Roeder K. Genomic control for association studies. Biometrics. 1999; 55:997–1004. [PubMed: 11315092]

6. Reich DE, Goldstein DB. Detecting association in a case-control study while correcting for population stratification. 2001; 20:4–16.

7. Campbell CD, et al. Demonstrating stratification in a European-American population. Nature Genet. 2005; 37:868–872. [PubMed: 16041375]

8. Manolio TA, et al. Finding the missing heritability of complex diseases. Nature. 2009; 461:747–53. [PubMed: 19812666]

9. Visscher PM, et al. Assumption-Free Estimation of Heritability from Genome-Wide Identity-by-Descent Sharing between Full Siblings. PLoS Genet. 2006; 2:e41. [PubMed: 16565746]

10. Weedon MN, et al. A common variant of HMGA2 is associated with adult and childhood height in the general population. Nat Genet. 2007; 39:1245–1250. [PubMed: 17767157]

11. Weedon MN, et al. Genome-wide association analysis identifies 20 loci that influence adult height. Nat Genet. 2008; 40:575–83. [PubMed: 18391952]

12. Sanna S, et al. Common variants in the GDF5-UQCC region are associated with variation in human height. Nat Genet. 2008; 40:198–203. [PubMed: 18193045]

13. Lettre G, et al. Identification of ten loci associated with height highlights new biological pathways in human growth. Nat Genet. 2008; 40:584–91. [PubMed: 18391950]

14. Soranzo N, Rivadeneira F, Chinapen-Horsley U, Malkina I. Meta-analysis of genome-wide scans for human adult stature in humans identifies novel loci and associations with measures of skeletal frame size. PLoS Genet. 2009; 5:e1000445. [PubMed: 19343178]

15. Gudbjartsson DF, et al. Many sequence variants affecting diversity of adult human height. Nat Genet. 2008; 40:609–15. [PubMed: 18391951]

16. Johansson A, et al. Common variants in the JAZF1 gene associated with height identified by linkage and genome-wide association analysis. Hum Mol Genet. 2009; 18:373–80. [PubMed: 18952825]
17. Estrada K, et al. 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. 2009; 18:3516–24. [PubMed: 19570815]

18. Purcell SM, et al. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. Nature. 2009; 460:748–52. [PubMed: 19571811]

19. Park JH, et al. Estimation of effect size distribution from genome-wide association studies and implications for future discoveries. Nat Genet. 2010; 42:570–5. [PubMed: 20562874]

20. Raychaudhuri S, et al. Identifying relationships among genomic disease regions: predicting genes at pathogenic SNP associations and rare deletions. PLoS Genet. 2009; 5:e1000534. [PubMed: 19557189]

21. Segrè AV, et al. Common Inherited Variation in Mitochondrial Genes is not Enriched for Associations with Type 2 Diabetes or Related Glycemic Traits. PLoS Genet. 2010 In Press.

22. Neptune ER, et al. Dysregulation of TGF-beta activation contributes to pathogenesis in Marfan syndrome. Nat Genet. 2003; 33:407–11. [PubMed: 12598898]

23. Superti-Furga A, Unger S. Nosology and classification of genetic skeletal disorders: 2006 revision. Am J Med Genet A. 2007; 143:1–18. [PubMed: 17120245]

24. Kofoed EM, et al. Growth hormone insensitivity associated with a STAT5b mutation. N Engl J Med. 2003; 349:1139–47. [PubMed: 13679528]

25. Weinstein M, Xu X, Ohyama K, Deng CX. FGFR-3 and FGFR-4 function cooperatively to direct alveogenesis in the murine lung. Development. 1998; 125:3615–23. [PubMed: 9716527]

26. Goldstein DB. Common genetic variation and human traits. N Engl J Med. 2009; 360:1696–8. [PubMed: 19369660]

27. Dickson SP, Wang K, Krantz I, Hakonarson H, Goldstein DB. Rare variants create synthetic genome-wide associations. PLoS Biol. 2010; 8:e1000294. [PubMed: 20126254]

28. Yang J, et al. Common SNPs explain a large proportion of the heritability for human height. Nat Genet. 2010; 42:565–9. [PubMed: 20562875]

29. Purcell S, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007; 81:559–75. [PubMed: 17701901]

30. Laird NM, Horvath S, Xu X. Implementing a unified approach to family-based tests of association. Genet Epidemiol. 2000; 19:1:S36–42. [PubMed: 11055368]
Figure 1. Phenotypic variance explained by common variants
(a) Variance explained is higher when SNPs not reaching genome-wide significance are included in the prediction model. The y-axis represents the proportion of variance explained at different $P$-value thresholds from Stage 1. Results are given for five studies that were not part of Stage 1. *Proportion of variation explained by the 180 SNPs. (b) Cumulative number of susceptibility loci expected to be discovered, including already identified loci and as yet undetected loci. The projections are based on loci that achieved a significance level of $P<5\times10^{-8}$ in the initial scan and the distribution of their effect sizes in Stage 2. The dotted red line corresponds to expected phenotypic variance explained by the 110 loci that reached genome-wide significance in Stage 1, were replicated in Stage 2 and had at least 1% power.
Figure 2. Example of a locus with a secondary signal before (a) and after (b) conditioning
The plot is centered on the conditioned SNP (purple diamond) at the locus. $r^2$ is based on the
CEU HapMap II samples. The blue line and right hand Y axis represent CEU HapMap II
recombination rates. Created using LocusZoom (http://csg.sph.umich.edu/locuszoom/).
Figure 3. Loci associated with height contain genes related to each other
(a) 180 height-associated SNPs. The y-axis plots GRAIL P-values on a log scale. The
histogram corresponds to the distribution of GRAIL P-values for 1,000 sets of 180 matched
SNPs. The scatter plot represents GRAIL results for the 180 height SNPs (blue dots). The
black horizontal line marks the median of the GRAIL P-values (P=0.14). The top 10
keywords linking the genes were: ‘growth’, ‘kinase’, ‘factor’, ‘transcription’, ‘signaling’,
‘binding’, ‘differentiation’, ‘development’, ‘insulin’, ‘bone’. (b) Graphical representation of
the connections between SNPs and corresponding genes for the 42 SNPs with GRAIL
P<0.01. Thicker and redder lines imply stronger literature-based connectivity.
Table 1

Secondary signals at associated loci after conditional analysis

| Second signal SNP | Conditioned SNP | Chr | Second signal SNP position | Distance of conditioned SNP from index SNP (bp) | HapMap $\text{P}^a$ | Second signal P-value after conditioning | Second signal P-value pre-conditioning | Gene $^b$ |
|-------------------|----------------|-----|-----------------------------|-----------------------------------------------|---------------------|----------------------------------------|----------------------------------------|---------|
| rs2280470         | rs16942341     | 15  | 87196630                    | 6721                                          | 0.009               | $1 \times 10^{-14}$                    | $1 \times 10^{-15}$                   | ACAN    |
| rs10859563        | rs11107116     | 12  | 92644470                    | 141835                                        | 0.003               | $3 \times 10^{-12}$                    | $8 \times 10^{-10}$                   | SOCS2   |
| rs750460          | rs5742915      | 15  | 72028559                    | 95127                                         | 0.004               | $4 \times 10^{-12}$                    | $7 \times 10^{-08}$                   | PML     |
| rs6938239         | rs2780226 $^c$ | 6   | 34791613                    | 484583                                        | 0.019               | $6 \times 10^{-12}$                    | $9 \times 10^{-14}$                   | HMGAI   |
| rs7652177         | rs572169       | 3   | 173451771                   | 196650                                        | 0.006               | $7 \times 10^{-11}$                    | $1 \times 10^{-11}$                   | GHSR    |
| rs7916441         | rs21459988     | 10  | 80595838                    | 196119                                        | 0.112               | $6 \times 10^{-10}$                    | $3 \times 10^{-07}$                   | PPAP    |
| rs3792752         | rs1173727      | 5   | 32804391                    | 61887                                         | 0.02                | $7 \times 10^{-10}$                    | $4 \times 10^{-08}$                   | PRP3    |
| rs10958476        | rs7460090      | 8   | 57258362                    | 98355                                         | 0.02                | $1 \times 10^{-09}$                    | $5 \times 10^{-13}$                   | SDR16C5 |
| rs2353398         | rs7689420      | 4   | 145742208                   | 45594                                         | 0.022               | $2 \times 10^{-09}$                    | $1 \times 10^{-10}$                   | IHIP    |
| rs2724475         | rs6449353      | 4   | 1755530                     | 87096                                         | 0.098               | $2 \times 10^{-09}$                    | $8 \times 10^{-10}$                   | LCORL   |
| rs2070776         | rs2665838      | 17  | 59361230                    | 41033                                         | 0.15                | $9 \times 10^{-09}$                    | $1 \times 10^{-14}$                   | GH region |
| rs1401796         | rs227724       | 17  | 52194758                    | 60942                                         | 0.005               | $2 \times 10^{-08}$                    | $7 \times 10^{-07}$                   | NOG     |
| rs4711336         | rs2780226 $^d$ | 6   | 33767024                    | 540046                                        | 0.111               | $3 \times 10^{-08}$                    | $5 \times 10^{-08}$                   | HMGAI   |
| rs6892884         | rs12153391     | 5   | 17048228                    | 187815                                        | 0                  | $4 \times 10^{-08}$                    | $2 \times 10^{-09}$                   | FBXW11  |
| rs1367226         | rs3791675      | 2   | 55943044                    | 21769                                         | 0.204               | $4 \times 10^{-08}$                    | $0.1245$                             | EFEMP1  |
| rs2421992         | rs1734652      | 1   | 170507874                   | 187964                                        | 0.019               | $5 \times 10^{-08}$                    | $1 \times 10^{-09}$                   | DNM3    |
| rs225694          | rs763064      | 6   | 142568835                   | 270147                                        | 0.001               | $1 \times 10^{-07}$                    | $2 \times 10^{-06}$                   | GPR126  |
| rs10187066        | rs12470505     | 2   | 219223003                   | 393610                                        | 0.022               | $2 \times 10^{-09}$                    | $5 \times 10^{-08}$                   | IHH     |
| rs879882          | rs2256183      | 6   | 31247431                    | 241077                                        | 0.016               | $2 \times 10^{-07}$                    | $8 \times 10^{-08}$                   | MICA    |

$a$ HapMap CEU phase II release 23

$b$ Nearest gene unless there is a known skeletal growth disorder gene in the locus (highlighted blue). Positions are based on NCBI build 36.

$c$ Nearest conditioned SNP where second signal occurs within 1 Mb of two conditioned SNPs.