Title: A low-frequency inactivating *AKT2* variant enriched in the Finnish population is associated with fasting insulin levels and type 2 diabetes risk.

Running title: *AKT2* coding variant affects fasting insulin levels

Corresponding authors

Prof. Anna L Gloyn
Oxford Centre for Diabetes Endocrinology & Metabolism
University of Oxford
Churchill Hospital
Headington
Oxford
OX3 7LE
United Kingdom
anna.gloyn@drl.ox.ac.uk

Prof. Cecilia M Lindgren
The Big Data Institute, Li Ka Shing Centre for Health Information and Discovery
The Wellcome Trust Centre for Human Genetics
University of Oxford
Roosevelt Drive
Oxford
OX3 7BN
United Kingdom
celi@well.ox.ac.uk

Word count: 4,850
Number of figures: 4
Number of tables: 0

Diabetes Publish Ahead of Print, published online March 24, 2017

Author List

Alisa Manning[1,2,3,¶], Heather M Highland[4,5,¶], Jessica Gasser[1,¶], Xueling Sim[6,7,¶], Taru Tukiainen[1,8,9,¶], Pierre Fontanillas[1,10,¶], Niels Grarup[11], Manuel A Rivas[12], Anubha Mahajan[12], Adam E Locke[6], Pablo Cingolani[13,14], Tune H Pers[1,11,15,16], Ana Viñuela[17,18,19], Andrew A Brown[20,21], Ying Wu[22], Jason Flannick[1,23], Christian Fuchsberger[6], Eric R Gamazon[24,25], Kyle J Gaulton[12,26], Hae Kyung Im[24], Tanya M Teslovich[6], Thomas W Blackwell[6], Jette Bork-Jensen[11], Noël P Burtt[1], Yuhui Chen[12], Todd Green[1], Christopher Hartl[1], Hyun Min Kang[6], Ashish Kumar[12,27], Claes Ladenvall[28], Clement Ma[6], Loukas Moutsianas[12], Richard D Pearson[12], John R B Perry[12,29,30], N William Rayner[12,31,32], Neil R Robertson[12,31], Laura J Scott[6], Martijn van de Bunt[12,31], Johan G Eriksson[33,34,35,36,37], Antti Jula[37], Seppo Koskinen[37], Terho Lehtimäki[38], Aarno Palotie[1,2,39], Olli T Raitakari[40,41], Suzanne BR Jacobs[1], Jennifer Wessel[42,43], Audrey Y Chu[44], Robert A Scott[30], Mark O Goodarzi[45,46], Christine Blancher[47], Gemma Buck[47], David Buck[47], Peter S Chines[48], Stacey Gabriel[1], Anette P Gjesing[11], Christopher J Groves[31], Mette Hollensted[11], Jeroen R Huyghe[6], Anne U Jackson[6], Goo Jun[6], Johanne Marie Justesen[11], Massimo Mangino[49], Jacquelyn Murphy[1], Matt Neville[31], Robert Onofrio[1], Kerrin S Small[49], Heather M Stringham[6], Joseph Trakalo[47], Eric Banks[1], Jason Carey[1], Mauricio O Carneiro[1], Mark DePristo[1], Yossi Farjoun[1], Timothy Fennell[1], Jacqueline I Goldstein[1,8], George Grant[1], Martin Hrabé de Angelis[50,51,52], Jared Maguire[1], Benjamin M Neale[1,8], Ryan Poplin[1], Shaun Purcell[1,2,53], Thomas Schwarzmayr[54], Khalid Shakir[1], Joshua D Smith[55], Tim M Strom[54,56], Thomas Wieland[54], Jaana Lindstrom[57], Ivan Brandslund[58,59], Cramer Christensen[60], Gabriela L Surdulescu[49], Timo A Lakka[61,62,63], Alex S F Doney[64], Peter Nilsson[65], Nicholas J Wareham[30], Claudia Langenberg[30], Tibor V Varga[66], Paul W Franks[66,67,68], Olov Rolandsson[68], Anders H Rosengren[28], Vidya S Farook[69], Farook Thameem[70], Sobha Puppala[69], Satish Kumar[69], Donna M Lehman[70], Christopher P Jenkinson[70,71], Joanne E Curran[69], Daniel Esten Hale[72], Sharon P Fowler[70], Rector Arya[72], Ralph A DeFronzo[70], Hanna E Abboud[70], Ann-Christine Syvänen[73], Pamela J Hicks[74,75,76], Nicholette D Palmer[74,75,76], Maggie C Y Ng[74,75], Donald W Bowden[74,75,76], Barry I Freedman[77], Tõnu Esko[1,9,78,79], Reedik Mägi[79], Lili Milani[79], Evelin Mihailov[79], Andres Metspalu[79], Narisu Narisu[48], Leena Kinnunen[37], Lori L Bonnycastle[48], Amy Swift[48], Dorota Pasko[29], Andrew R Wood[29], João Fadista[28], Toni I Pollin[80], Nir Barzilai[81], Gil Atzmon[81], Benjamin Glaser[82], Barbara Thorand[51,83], Konstantin Strauch[84,85], Annette Peters[51,83,86], Michael Roden[87,88], Martina Müller-Nurasyid[84,85,86,89], Liming Liang[90,91], Jennifer Kriebel[51,83,92], Thomas Illig[92,93,94], Harald Grallert[51,83,92], Christian Gieger[84], Christa Meisinger[83], Lars Lannfelt[95], Solomon K Musani[96], Michael Griswold[97], Herman A Taylor Jr[98], Gregory Wilson Sr[99], Adolfo Correa[98], Heikki Oksa[100], William R Scott[101], Uzma Afzal[101], Sian-Tsung Tan[102,103], Marie Loh[101,104,105], John C Chambers[101,103,106], Jobanpreet Sehmi[102,103], Jaspal Singh Kooner[102], Benjamin Lehne[101], Yoon Shin Cho[107], Jong-Young Lee[108], Bok-Ghee Han[109], Annemari Käräjämäki[110,111], Qibin Qi[67,112], Lu Qi[67,113], Jinyan Huang[90], Frank B Hu[67,90], Olle Melander[114], Marju Orho-Melander[115], Jennifer E Below[116], David Aguilar[117], Tien Yin Wong[118,119], Jianjun Liu[7,120], Chiea-Chuen Khor[7,118,119,120,121], Kee Seng Chia[7], Wei Yen Lim[7], Ching-Yu Cheng[7,118,119,122], Edmund Chan[123], E Shyong Tai[7,123,124], Tin Aung[118,119], Allan Linneberg[125,126,127], Bo Isomaa[128,129], Thomas Meitinger[54,56,86], Tiinamaija Tuomi[129,130], Liisa Hakaste[35], Jasmina Kravic[28], Marit E Jørgensen[131], Torsten Lauritzen[132], Panos Deloukas[32],

Kathleen E Stirrups[133,134], Katharine R Owen[31,135], Andrew J Farmer[136], Timothy M Frayling[29], Stephen P O'Rahilly[137], Mark Walker[138], Jonathan C Levy[31], Dylan Hodgkiss[49], Andrew T Hattersley[139], Teemu Kuulasmaa[140], Alena Stančáková[140], Inês Barroso[32,137], Dwaipayan Bharadwaj[141], Juliana Chan[142,143,144], Giriraj R Chandak[145], Mark J Daly[8], Peter J Donnelly[12,146], Shah B Ebrahim[147], Paul Elliott[101,148], Tasha Fingerlin[149], Philippe Froguel[150], Cheng Hu[151], Weiping Jia[151], Ronald C W Ma[142,143,144], Gilean McVean[12], Taesung Park[152,153], Dorairaj Prabhakaran[147], Manjinder Sandhu[32,154], James Scott[102], Rob Sladek[14,155,156], Nikhil Tandon[157], Yik Ying Teo[7,158,159], Eleftheria Zeggini[32], Richard M Watanabe[160,161,162], Heikki A Koistinen[37,163,164], Y Antero Kesaniemi[165], Matti Uusitupa[166], Timothy D Spector[49], Veikko Salomaa[37], Rainer Rauramaa[167], Colin N A Palmer[168], Inga Prokopenko[12,31,169], Andrew D Morris[170], Richard N Bergman[171], Francis S Collins[48], Lars Lind[172], Erik Ingelsson[173,174], Jaakko Tuomilehto[57,175,176,177], Fredrik Karpe[31,135], Leif Groop[28], Torben Jørgensen[125,178], Torben Hansen[11,179], Oluf Pedersen[11], Johanna Kuusisto[140,180], Gonçalo Abecasis[6], Graeme I Bell[181], John Blangero[69], Nancy J Cox[24], Ravindranath Duggirala[69], Mark Seielstad[182,183], James G Wilson[184], Josee Dupuis[185,186], Samuli Ripatti[20,39,187], Craig L Hanis[116], Jose C Florez[1,2,3,188], Karen L Mohlke[22], James B Meigs[1,3,189], Markku Laakso[140,180], Andrew P Morris[12,79,190], Michael Boehnke[6], David Altshuler[1,3,9,23,188,191], Mark I McCarthy[12,31,135], Anna L Gloyn[12,31,135,&], Cecilia M Lindgren[1,12,192,&]

Author Affiliations:
1. Program in Medical and Population Genetics, Broad Institute, Cambridge, Massachusetts, USA.
2. Center for Human Genetic Research, Department of Medicine, Massachusetts General Hospital, Boston, Massachusetts, USA.
3. Department of Medicine, Harvard Medical School, Boston, Massachusetts, USA.
4. Human Genetics Center, The University of Texas Graduate School of Biomedical Sciences at Houston, The University of Texas Health Science Center at Houston, Houston, Texas, USA.
5. Department of Epidemiology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA.
6. Department of Biostatistics and Center for Statistical Genetics, School of Public Health, University of Michigan, Ann Arbor, Michigan, USA.
7. Saw Swee Hock School of Public Health, National University of Singapore, National University Health System, Singapore.
8. Analytic and Translational Genetics Unit, Department of Medicine, Massachusetts General Hospital, Boston, Massachusetts, USA.
9. Department of Genetics, Harvard Medical School, Boston, Massachusetts, USA.
10. 23andMe, Mountain View, California, USA.
11. The Novo Nordisk Foundation Center for Basic Metabolic Research, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark.
12. Wellcome Trust Centre for Human Genetics, Nuffield Department of Medicine, University of Oxford, Oxford, UK.
13. School of Computer Science, McGill University, Montreal, Quebec, Canada.
14. McGill University and Génome Québec Innovation Centre, Montreal, Quebec, Canada.
15. Divisions of Endocrinology and Genetics and Center for Basic and Translational Obesity Research, Boston Children's Hospital, Boston, Massachusetts, USA.
16. Department of Epidemiology Research, Statens Serum Institut, Copenhangen, Denmark.
17. Twin Research and Genetic Epidemiology, King's College London, London, UK.
18. Department of Genetic Medicine and Development, University of Geneva Medical School, Geneva, Switzerland.
19. Institute of Genetics and Genomics in Geneva, University of Geneva, Geneva, Switzerland.
20. Wellcome Trust Sanger Institute, Hinxton, Cambridge, UK.
21. NORMENT, KG Jebsen Center for Psychosis Research, Division of Mental Health and Addiction, Oslo University Hospital, Oslo, Norway.
22. Department of Genetics, University of North Carolina, Chapel Hill, North Carolina, USA.
23. Department of Molecular Biology, Massachusetts General Hospital, Boston, Massachusetts, USA.
24. Department of Medicine, Section of Genetic Medicine, The University of Chicago, Chicago, Illinois, USA.
25. Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands.
26. Department of Pediatrics, University of California San Diego, La Jolla, California, USA.
27. Chronic Disease Epidemiology, Swiss Tropical and Public Health Institute, University of Basel, Basel, Switzerland.

28. Department of Clinical Sciences, Diabetes and Endocrinology, Lund University Diabetes Centre, Malmö, Sweden.
29. Genetics of Complex Traits, University of Exeter Medical School, University of Exeter, Exeter, UK.
30. MRC Epidemiology Unit, Institute of Metabolic Science, University of Cambridge, Cambridge, UK.
31. Oxford Centre for Diabetes, Endocrinology and Metabolism, Radcliffe Department of Medicine, University of Oxford, Oxford, UK.
32. Department of Human Genetics, Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, UK.
33. Department of General Practice and Primary Healthcare, University of Helsinki, Helsinki, Finland.
34. Unit of General Practice, Helsinki University Central Hospital, Finland.
35. Folkhälsan Research Center, Helsinki, Finland.
36. Vaasa Central Hospital, Vaasa, Finland.
37. Department of Health, National Institute of Health and Welfare, Helsinki, Finland.
38. Department of Clinical Chemistry, Fimlab Laboratories, University of Tampere School of Medicine, Tampere, Finland.
39. Institute for Molecular Medicine Finland, University of Helsinki, Helsinki, Finland.
40. Department of Clinical Physiology and Nuclear Medicine, Turku University Hospital, Turku, Finland.
41. Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, Turku, Finland.
42. Department of Epidemiology, Fairbanks School of Public Health, Indianapolis, IN, USA.
43. Department of Medicine, Indiana University School of Medicine, Indianapolis, IN, USA.
44. Division of Preventive Medicine, Brigham and Women's Hospital, Boston, MA, USA.
45. Division of Endocrinology, Diabetes and Metabolism, Department of Medicine, Cedars-Sinai Medical Center, Los Angeles, California, USA.
46. Department of Biomedical Sciences, Cedars-Sinai Medical Center, Los Angeles, California, USA.
47. High Throughput Genomics, Oxford Genomics Centre, Wellcome Trust Centre for Human Genetics, Nuffield Department of Medicine, University of Oxford, Oxford, UK.
48. National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland, USA.
49. Department of Twin Research and Genetic Epidemiology, King's College London, London, UK.
50. Institute of Experimental Genetics, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany.
51. German Center for Diabetes Research (DZD), Neuherberg, Germany.
52. Chair of Experimental Genetics, School of Life Science Weihenstephan, Technische Universität München, Freising, Germany.
53. Department of Psychiatry, Icahn Institute for Genomics and Multiscale Biology, Icahn School of Medicine at Mount Sinai, New York, USA.

54. Institute of Human Genetics, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany.
55. Department of Genome Sciences, University of Washington School of Medicine, Seattle, Washington, USA.
56. Institute of Human Genetics, Technische Universität München, Munich, Germany.
57. Diabetes Prevention Unit, National Institute for Health and Welfare, Helsinki, Finland.
58. Institute of Regional Health Research, University of Southern Denmark, Odense, Denmark.
59. Department of Clinical Biochemistry, Vejle Hospital, Vejle, Denmark.
60. Department of Internal Medicine and Endocrinology, Vejle Hospital, Vejle, Denmark.
61. Institute of Biomedicine, Physiology, University of Eastern Finland, Kuopio, Finland.
62. Kuopio Research Institute of Exercise Medicine, Kuopio, Finland.
63. Department of Clinical Physiology and Nuclear Medicine, Kuopio University Hospital, Kuopio, Finland.
64. Division of Cardiovascular and Diabetes Medicine, Medical Research Institute, Ninewells Hospital and Medical School, Dundee, UK.
65. Department of Clinical Sciences, Medicine, Lund University, Malmö, Sweden.
66. Department of Clinical Sciences, Lund University Diabetes Centre, Genetic and Molecular Epidemiology Unit, Lund University, Malmö, Sweden.
67. Department of Nutrition, Harvard School of Public Health, Boston, Massachusetts, USA.
68. Department of Public Health and Clinical Medicine, Umeå University, Umeå, Sweden.
69. Department of Genetics, Texas Biomedical Research Institute, San Antonio, Texas, USA.
70. Department of Medicine, University of Texas Health Science Center, San Antonio, Texas, USA.
71. Research, South Texas Veterans Health Care System, San Antonio, Texas, USA.
72. Department of Pediatrics, University of Texas Health Science Center, San Antonio, Texas, USA.
73. Department of Medical Sciences, Molecular Medicine and Science for Life Laboratory, Uppsala University, Uppsala, Sweden.
74. Center for Genomics and Personalized Medicine Research, Wake Forest School of Medicine, Winston-Salem, North Carolina, USA.
75. Center for Diabetes Research, Wake Forest School of Medicine, Winston-Salem, North Carolina, USA.
76. Department of Biochemistry, Wake Forest School of Medicine, Winston-Salem, North Carolina, USA.
77. Department of Internal Medicine, Section on Nephrology, Wake Forest School of Medicine, Winston-Salem, North Carolina, USA.
78. Division of Endocrinology, Boston Children's Hospital, Boston, Massachusetts, USA.
79. Estonian Genome Center, University of Tartu, Tartu, Estonia.
80. Department of Medicine, Program in Personalized and Genomic Medicine, University of Maryland, Baltimore, Maryland, USA.
81. Departments of Medicine and Genetics, Albert Einstein College of Medicine, New York, USA.

82. Endocrinology and Metabolism Service, Hadassah-Hebrew University Medical Center, Jerusalem, Israel.
83. Institute of Epidemiology II, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany.
84. Institute of Genetic Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany.
85. Institute of Medical Informatics, Biometry and Epidemiology, Chair of Genetic Epidemiology, Ludwig-Maximilians-Universität, Munich, Germany.
86. Deutsches Forschungszentrum für Herz-Kreislauferkrankungen (DZHK), Partner Site Munich Heart Alliance, Munich, Germany.
87. Institute of Clinical Diabetology, German Diabetes Center, Leibniz Center for Diabetes Research at Heinrich Heine University, Düsseldorf, Germany.
88. German Center for Diabetes Research, Partner Düsseldorf, Germany.
89. Department of Medicine I, University Hospital Grosshadern, Ludwig-Maximilians-Universität, Munich, Germany.
90. Department of Epidemiology, Harvard School of Public Health, Boston, Massachusetts, USA.
91. Department of Biostatistics, Harvard School of Public Health, Boston, Massachusetts, USA.
92. Research Unit of Molecular Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany.
93. Hannover Unified Biobank, Hannover Medical School, Hannover, Germany.
94. Institute of Human Genetics, Hannover Medical School, Hannover, Germany.
95. Department of Public Health and Caring Sciences, Geriatrics, Uppsala University, Uppsala, Sweden.
96. Jackson Heart Study, University of Mississippi Medical Center, Jackson, Mississippi, USA.
97. Center of Biostatistics and Bioinformatics, University of Mississippi Medical Center, Jackson, Mississippi, USA.
98. Department of Medicine, University of Mississippi Medical Center, Jackson, Mississippi, USA.
99. College of Public Services, Jackson State University, Jackson, Mississippi, USA.
100. Pirkanmaa Hospital District, Tampere, Finland.
101. Department of Epidemiology and Biostatistics, Imperial College London, London, UK.
102. National Heart and Lung Institute, Cardiovascular Sciences, Hammersmith Campus, Imperial College London, London, UK.
103. Department of Cardiology, Ealing Hospital NHS Trust, Southall, Middlesex, UK.
104. Institute of Health Sciences, University of Oulu, Oulu, Finland.
105. Translational Laboratory in Genetic Medicine (TLGM), Agency for Science, Technology and Research (A*STAR), Singapore.
106. Imperial College Healthcare NHS Trust, Imperial College London, London, UK.
107. Department of Biomedical Science, Hallym University, Chuncheon, Republic of Korea.
108. Ministry of Health and Welfare, Seoul, Republic of Korea.
109. Center for Genome Science, Korea National Institute of Health, Chungcheongbuk-do, Republic of Korea.
110. Vasa Health Care Center, Vaasa, Finland.

111. Department of Primary Health Care, Vasa Central Hospital, Vasa, Finland.
112. Department of Epidemiology and Population Health, Albert Einstein College of Medicine, New York, USA.
113. Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts, USA.
114. Department of Clinical Sciences, Hypertension and Cardiovascular Disease, Lund University, Malmö, Sweden.
115. Department of Clinical Sciences, Diabetes and Cardiovascular Disease, Genetic Epidemiology, Lund University, Malmö, Sweden.
116. Human Genetics Center, School of Public Health, The University of Texas Health Science Center at Houston, Houston, Texas, USA.
117. Cardiovascular Division, Baylor College of Medicine, Houston, Texas, USA.
118. Singapore Eye Research Institute, Singapore National Eye Centre, Singapore.
119. Department of Ophthalmology, Yong Loo Lin School of Medicine, National University of Singapore, National University Health System, Singapore.
120. Division of Human Genetics, Genome Institute of Singapore, A*STAR, Singapore.
121. Department of Paediatrics, Yong Loo Lin School of Medicine, National University of Singapore, National University Health System, Singapore.
122. Centre for Quantitative Medicine, Office of Clinical Sciences, Duke-NUS Graduate Medical School Singapore, Singapore.
123. Department of Medicine, Yong Loo Lin School of Medicine, National University of Singapore, National University Health System, Singapore.
124. Cardiovascular & Metabolic Disorders Program, Duke-NUS Graduate Medical School Singapore, Singapore.
125. Research Centre for Prevention and Health, Glostrup University Hospital, Glostrup, Denmark.
126. Department of Clinical Experimental Research, Rigshospitalet, Glostrup, Denmark.
127. Department of Clinical Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark.
128. Department of Social Services and Health Care, Jakobstad, Finland.
129. Folkhälsan Research Centre, Helsinki, Finland.
130. Department of Endocrinology, Helsinki University Central Hospital, Helsinki, Finland.
131. Steno Diabetes Center, Gentofte, Denmark.
132. Department of Public Health, Section of General Practice, Aarhus University, Aarhus, Denmark.
133. William Harvey Research Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, UK.
134. Department of Haematology, University of Cambridge, Cambridge, UK.
135. Oxford NIHR Biomedical Research Centre, Oxford University Hospitals Trust, Oxford, UK.
136. Department of Primary Care Health Sciences, University of Oxford, Oxford, UK.
137. Metabolic Research Laboratories, Institute of Metabolic Science, University of Cambridge, Cambridge, UK.
138. The Medical School, Institute of Cellular Medicine, University of Newcastle, Newcastle, UK.
139. University of Exeter Medical School, University of Exeter, Exeter, UK.

140. Faculty of Health Sciences, Institute of Clinical Medicine, Internal Medicine, University of Eastern Finland, Kuopio, Finland.
141. Functional Genomics Unit, CSIR-Institute of Genomics & Integrative Biology (CSIR-IGIB), New Delhi, India.
142. Department of Medicine and Therapeutics, The Chinese University of Hong Kong, Hong Kong, China.
143. Li Ka Shing Institute of Health Sciences, The Chinese University of Hong Kong, Hong Kong, China.
144. Hong Kong Institute of Diabetes and Obesity, The Chinese University of Hong Kong, Hong Kong, China.
145. CSIR-Centre for Cellular and Molecular Biology, Hyderabad, Andhra Pradesh, India.
146. Department of Statistics, University of Oxford, Oxford, UK.
147. Centre for Chronic Disease Control, New Delhi, India.
148. MRC-PHE Centre for Environment and Health, Imperial College London, London, UK.
149. Department of Epidemiology, Colorado School of Public Health, University of Colorado, Aurora, Colorado, USA.
150. Genomics and Molecular Physiology, CNRS (Institut de Biologie de Lille), Lille, France.
151. Department of Endocrinology and Metabolism, Shanghai Diabetes Institute, Shanghai Jiao Tong University Affiliated Sixth People's Hospital, Shanghai, China.
152. Interdisciplinary Program in Bioinformatics, Seoul National University, Seoul, Republic of Korea.
153. Department of Statistics, Seoul National University, Seoul, Republic of Korea.
154. Department of Public Health and Primary Care, Institute of Public Health, University of Cambridge, Cambridge, UK.
155. Department of Human Genetics, McGill University, Montreal, Quebec, Canada.
156. Division of Endocrinology and Metabolism, Department of Medicine, McGill University, Montreal, Quebec, Canada.
157. Department of Endocrinology and Metabolism, All India Institute of Medical Sciences, New Delhi, India.
158. Life Sciences Institute, National University of Singapore, Singapore.
159. Department of Statistics and Applied Probability, National University of Singapore, Singapore.
160. Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, California, USA.
161. Department of Physiology & Biophysics, Keck School of Medicine, University of Southern California, Los Angeles, California, USA.
162. Diabetes and Obesity Research Institute, Keck School of Medicine, University of Southern California, Los Angeles, California, USA.
163. University of Helsinki and Helsinki University Central Hospital, Department of Medicine and Abdominal Center, Endocrinology, Helsinki, Finland.
164. Minerva Foundation Institute for Medical Research, Helsinki, Finland.
165. Institute of Clinical Medicine, Department of Medicine, University of Oulu and Medical Research Center, Oulu University Hospital, Oulu, Finland.
166. Institute of Public Health and Clinical Nutrition, University of Eastern Finland, Kuopio, Finland.

167. Foundation for Research in Health, Exercise and Nutrition, Kuopio Research Institute of Exercise Medicine, Kuopio, Finland.
168. Pat Macpherson Centre for Pharmacogenetics and Pharmacogenomics, Medical Research Institute, Ninewells Hospital and Medical School, Dundee, UK.
169. Department of Genomics of Common Disease, School of Public Health, Imperial College London, London, UK.
170. Clinical Research Centre, Centre for Molecular Medicine, Ninewells Hospital and Medical School, Dundee, UK.
171. Cedars-Sinai Diabetes and Obesity Research Institute, Los Angeles, California, USA.
172. Department of Medical Sciences, Uppsala University, Uppsala, Sweden.
173. Department of Medical Sciences, Molecular Epidemiology and Science for Life Laboratory, Uppsala University, Uppsala, Sweden.
174. Department of Medicine, Division of Cardiovascular Medicine, Stanford University School of Medicine, Stanford, California, USA.
175. Center for Vascular Prevention, Danube University Krems, Krems, Austria.
176. Diabetes Research Group, King Abdulaziz University, Jeddah, Saudi Arabia.
177. Dasman Diabetes Institute, Dasman, 15642 Kuwait.
178. Faculty of Medicine, University of Aalborg, Aalborg, Denmark.
179. Faculty of Health Sciences, University of Southern Denmark, Odense, Denmark.
180. Kuopio University Hospital, Kuopio, Finland.
181. Departments of Medicine and Human Genetics, The University of Chicago, Chicago, Illinois, USA.
182. Department of Laboratory Medicine & Institute for Human Genetics, University of California, San Francisco, San Francisco, California, USA.
183. Blood Systems Research Institute, San Francisco, California, USA.
184. Department of Physiology and Biophysics, University of Mississippi Medical Center, Jackson, Mississippi, USA.
185. Department of Biostatistics, Boston University School of Public Health, Boston, Massachusetts, USA.
186. National Heart, Lung, and Blood Institute's Framingham Heart Study, Framingham, Massachusetts, USA.
187. Hjelt Institute, University of Helsinki, Helsinki, Finland.
188. Diabetes Research Center (Diabetes Unit), Department of Medicine, Massachusetts General Hospital, Boston, Massachusetts, USA.
189. Division of General Internal Medicine, Massachusetts General Hospital, Boston, Massachusetts, USA.
190. Department of Biostatistics, University of Liverpool, Liverpool, UK.
191. Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA.
192. Li Ka Shing Centre for Health Information and Discovery, The Big Data Institute, University of Oxford, Oxford, UK.

¶ These authors contributed equally to this work.
& These authors jointly directed this research.

ABSTRACT

To identify novel coding association signals and facilitate characterization of mechanisms influencing glycemic traits and type 2 diabetes risk, we analyzed 109,215 variants derived from exome array genotyping together with an additional 390,225 variants from exome sequence in up to 39,339 normoglycemic individuals from five ancestry groups. We identified a novel association between the coding variant (p.Pro50Thr) in *AKT2* and fasting insulin, a gene in which rare fully penetrant mutations are causal for monogenic glycemic disorders. The low-frequency allele is associated with a 12% increase in fasting plasma insulin (FI) levels. This variant is present at 1.1% frequency in Finns but virtually absent in individuals from other ancestries. Carriers of the FI-increasing allele had increased 2-hour insulin values, decreased insulin sensitivity, and increased risk of type 2 diabetes (odds ratio=1.05). In cellular studies, the AKT2-Thr50 protein exhibited a partial loss of function. We extend the allelic spectrum for coding variants in *AKT2* associated with disorders of glucose homeostasis and demonstrate bidirectional effects of variants within the pleckstrin homology domain of *AKT2*.

The increasing prevalence of type 2 diabetes is a global health crisis, making it critical to promote development of more efficient strategies for prevention and treatment. Individuals with type 2 diabetes display both pancreatic beta-cell dysfunction and insulin resistance . Genetic studies of surrogate measures of these glycemic traits can identify variants that influence these central features of type 2 diabetes (2) highlighting potential pathways for therapeutic manipulation. Comprehensive surveys of the influence of common genetic variants on fasting plasma glucose (FG) and fasting plasma insulin (FI) have highlighted defects in pathways involved in glucose metabolism, and insulin processing, secretion, and action (3). Recent studies have identified type 2 diabetes-associated alleles that are common in one population but rare or absent in others (4-6). These associations were observed either due to an increase in frequency of older alleles based on population dynamics and demography (5), or the emergence of population-specific alleles (4; 6).

We set out to identify and characterize low-frequency allele (minor allele frequency [MAF]<5%) glycemic trait associations by meta-analysis of exome sequence and exome array genotype data in a multi-ancestry sample. We also performed *in vitro* functional studies of protein expression, localization and activity to understand the consequences of our novel findings.

METHODS

Genetic association studies

Study Samples

The Genetics of Type 2 Diabetes (GoT2D) study and Type 2 Diabetes Genetic Exploration by Next-generation sequencing in multi-Ethnic Samples (T2D-GENES) study were initially designed to evaluate the contribution of coding variants to type 2 diabetes risk (7). We

performed a discovery association analysis to find novel coding variants associated with fasting glycemic traits in 14 studies from GoT2D that contributed exome array information on 33,231 non-diabetic individuals of European ancestry. Further discovery analysis was performed with GoT2D and T2D-GENES studies with exome sequence data (average 80x coverage) in five ancestral groups comprised of 12,940 individuals (6,504 with type 2 diabetes, 6,436 without) with measured FG or FI levels available in 2,144 European, 508 South Asian, 1,104 East Asian, 844 Hispanic, and 508 African American non-diabetic individuals. We performed a replication analysis and an assessment of allele frequency distributions in 5,747 individuals from four Finnish cohorts: Cardiovascular Risk in Young Finns Study (YFS) (8), Helsinki Birth Cohort (HBCS) (9), Health 2000 GenMets Study (GenMets) (10), and National FINRISK Study 1997 and 2002 (FR) (11). We also assessed the allele frequencies of novel findings in 46,658 individuals from CHARGE studies with available exome array data (12), although none of the studies passed our QC filter of a minor allele count greater than 5 for inclusion in our replication analysis. See Supplementary Table 1 for study details, sample characteristics, ascertainment criteria, and detailed genotype calling and quality control procedures for each cohort. The relevant institutional review boards, conducted according to the Declaration of Helsinki, approved all human research and all participants provided written informed consent. A detailed description of ethical permissions is provided in the Supplementary Materials.

Phenotypes

For the discovery and replication analysis, we excluded individuals from the analysis if they had a diagnosis of type 2 diabetes, were currently receiving oral or injected diabetes treatment, had FG measures ≥ 7mmol/L, had 2-hour post-load glucose (2hrG) measures ≥ 11.1mmol/L, or had HbA1c measures ≥ 6.5% (48mmol/mol). Additional exclusions occurring at

the study level included pregnancy, non-fasting at time of exam, type 1 diabetes, or impaired glucose tolerance. See Supplementary Table 1A for details. Within each study, we adjusted FG and log transformed FI levels for age, sex, body mass index (BMI), and additional study specific covariates. We applied rank-based inverse-normal transformations to study- or ancestry-specific residuals to obtain satisfactory asymptotic properties of the exome-wide association tests.

We tested for genetic associations with type 2 diabetes, hypertension, and other related quantitative traits in the Finnish discovery and replication cohorts. We analyzed lipid levels (total cholesterol, high-density lipoprotein cholesterol (HDL), low-density lipoprotein cholesterol (LDL), and triglycerides (TG)), blood pressure (systolic (SBP) and diastolic (DBP) blood pressures and hypertension (HTN)), height, BMI, central adiposity measures (waist-to-hip ratio (WHR), waist circumference, hip circumference), adiponectin level, 2-hour insulin level, and Matsuda index, which is known to correlate with whole-body insulin sensitivity as measured by the hyperinsulinemic euglycemic clamp (r=0.7, $P<1.0\times10^{-4}$) (13). For quantitative traits and HTN, we adjusted for age, sex, BMI (for glycemic, blood pressure, and central adiposity traits), stratified by type 2 diabetes status and sex (for central adiposity measures) within study. We adjusted LDL and total cholesterol for use of lipid-lowering medication, by dividing total cholesterol by 0.8 if on lipid-lowering medication, prior to calculating LDL using the Friedewald equation (14). SBP and DBP were adjusted for use of blood pressure-lowering medication by adding 15 mmHg to SBP and 10 mmHg to DBP measurements if an individual reported taking blood pressure-lowering medication (15). The Matsuda Index was log transformed and analyzed in non-diabetic individuals only. After adjusting for covariates, traits were inverse-normalized within strata. In addition to studying these metabolic outcomes, we used international classification of diseases (ICD) codes to query electronic medical records in the METSIM and

FINRISK 1997 and 2002 cohorts (in all individuals regardless of type 2 diabetes status) and categorized affection status for lipodystrophy, polycystic ovary disease, and ovarian or breast cancer.

Statistical Analysis

Discovery Analysis: We performed association analyses within each study for the exome array data sets and within ancestry for the exome sequence data sets. We used linear mixed models implemented in EMMAX (16) to account for relatedness. Within each study/ancestry, we required variants to have a minor allele count (MAC) greater than or equal to five alleles for single variant association tests. We meta-analyzed the single variant results from the (European-ancestry) exome array studies using the inverse variance meta-analysis approach implemented in METAL (17) and combined these with the European ancestry exome sequence results. Then, we meta-analyzed summary statistics across ancestries. We used $P<5\times10^{-7}$ as exome-wide statistical significance thresholds for the single variant tests (18). We used the binomial distribution to assess enrichment of previously reported associations with FG or FI by calculating a *P*-value for the number of non-significant variants with consistent direction of effects.

Gene based association analysis: We performed gene-based association tests using variants with MAF <1% (including rare variants with MAC≤5), annotating and aggregating variants based on predicted deleteriousness using previously described methods (7). Briefly, we defined four different variant groupings: "PTV-only", containing only variants predicted to severely impair protein function, "PTV+missense", containing PTV and NS variants with MAF <1%, "PTV+NS_{strict}" composed of PTV and NS variants predicted damaging by five algorithms (SIFT, LRT, MutationTaster, polyphen2 HDIV, and polyphen2 HVAR), and "PTV+NS_{broad}" composed of PTV and NS variants with MAF<1% and predicted damaging by at least one

prediction algorithm above. We used the sequence kernel association test (SKAT) (19) and a frequency-weighted burden test to conduct exome array meta-analyses in an unrelated subset of individuals using RareMETAL (20). We conducted exome sequence gene-based analyses within ancestry using a linear mixed model to account for relatedness and combined results across ancestries with MetaSKAT (21), which accounts for heterogeneous effects. We further combined gene-based results from exome array and exome sequences using Stouffer's method with equal weights. For gene-based tests, we considered $P<2.5\times10^{-6}$ as exome-wide significant, corresponding to Bonferroni correction for 20,000 genes in the genome (18).

Replication Analysis: The *AKT2* p.Pro50Thr variant was observed at sufficient frequency in the independent Finnish cohorts to perform single-variant association test of association with FI. We tested association in SNPTEST (22) (v.2.4.0) in each study with the same additive linear model used in the discovery analysis. Covariate adjustments for FI levels were sex, age, and ten principal components (PCs), and models were run with and without adjustment for BMI.

Estimate of effect on raw FI level and variance explained: To characterize the association between *AKT2* p.Pro50Thr and FI, we examined full regression models with raw FI in three studies (FUSION, METSIM, and YFS). We estimated the raw effect on log-transformed FI levels with a fixed-effects meta-analysis. The variance in log-transformed FI explained by *AKT2* p.Pro50Thr was estimated by a weighted average of the narrow-sense heritability of *AKT2* p.Pro50Thr seen in these three studies.

Population genetics and constraint: We used the Exome Aggregation Consortium (ExAC) for constraint metrics and allele frequencies (23). We obtained sequence alignments for AKT proteins and mRNAs in 100 vertebrates from the UCSC Genome Browser (24), used

Shannon's entropy (normalized K=21) as a conservation score (25) and plotted the sequence logos in R using the RWebLogo library (26).

Associations with other traits: We conducted association tests for traits other than FI and FG within studies for both discovery studies as well as the independent Finnish studies used for replication. *P*-values for type 2 diabetes and HTN came from EMMAX (16) or the Wald test from logistic regression (Finnish replication data sets) and meta-analyzed using an N weighted meta-analysis (17). Odds ratios (OR) were obtained from logistic regression adjusting for age, sex, with and without BMI, and PCs and meta-analyzed using an inverse variance meta-analysis.

Trait distributions and phenotype clustering: We examined distributions of traits among *AKT2* missense allele carriers (p.Pro50Thr, p.Arg208Lys, and p.Arg467Trp) in the T2D-GENES exome sequencing data set. We used non-parametric rank based methods (kruskal.wallis and permKS functions in R) on both the inverse-normalized covariate-adjusted traits used in the genetic association studies and normalized raw trait values (scale function in R). We clustered *AKT2* missense allele carriers on scaled trait values (pheatmap function in R).

In vitro functional studies

Plasmids and cell lines: The generation of the *AKT2* allelic series was initiated by the production of pDONR223-AKT2 through PCR of the human *AKT2* open reading frame with the integration of terminal attR sites using primers (see below). HeLa, HuH7, and 293T cells were obtained at The Broad Institute and maintained in 10% FBS DMEM, 100U/ml penicillin and 100µg/ml streptomycin, and documented mycoplasma-free. HeLa and HuH7 cells were starved for 18 hours and stimulated for 15 minutes with 100nM insulin for activation analyses.

Primers for functional work: The generation of the AKT2 allelic series was initiated by the production of pDONR223- AKT2 through PCR of the human AKT2 open reading frame with the

integration of terminal attR sites using primers FWD: 5’ - GGGGACAAGTTTGTACAAAAAAGTTGGCACCATGAATGAGGTGTCTGTCATC -3’ REV: 5’- GGGGACCACTTTGTACAAGAAAGTTGGCAACTCGCGGATGCTG -3’, and subsequent Gateway BP reaction into pDONR223 obtained from The Broad Institute Genetics Perturbation Platform. Site-directed mutagenesis was then performed to generate AKT2.E17K (AKT2.Lys17), AKT2.P50T (AKT2.Thr50), AKT2.R208K (AKT2.Lys208), AKT2.R274H (AKT2.His274), AKT2.R467W (AKT2.Trp467) with the following primers: AKT2.E17K: FWD: 5'- GGCTCCACAAGCGTGGTAAATACATCAAGACCTGG -3' REV: 5'- CCAGGTCTTGATGTATTTACCACGCTTGTGGAGCC -3'; AKT2.P50T: FWD: 5'- AGGCCCCTGATCAGACTCTAACCCCCTTAAAC -3' REV: 5'- GTTTAAGGGGGTTAGAGTCTGATCAGGGGCCT -3'; AKT2.R208K: FWD: 5'- GTCCTCCAGAACACCAAGCACCCGTTCC -3' REV: 5'- GGAACGGGTGCTTGGTGTTCTGGAGGAC -3'; AKT2.R274H: FWD: 5'- GGGACGTGGTATACCACGACATCAAGCTGGA -3'REV3'REV: 5'- TCCAGCTTGATGTCGTGGTATACCACGTCCC -3'; AKT2.R467W: FWD: 5'- GGAGCTGGACCAGTGGACCCACTTCCC -3' REV: 5'- GGGAAGTGGGTCCACTGGTCCAGCTCC -3'. C-terminal, V5-tagged lentiviral pLX304-AKT2.E17K, pLX304-AKT2.P50T, pLX304- AKT2.R208K, pLX304-AKT2.R274H, and pLX304- AKT2.R467W were each generated by subsequent Gateway LR reactions with pDONR223-AKT2.E17K, pDONR223-AKT2.P50T, pDONR223-AKT2.R208K, pDONR223-AKT2.R274H, and pDONR223-AKT2.R467W, respectively, and pLX304 obtained from The Broad Institute Genetics Perturbation Platform. Control plasmid pLX304- empty vector was additionally acquired from The Broad Institute Genetics Perturbation Platform.

Antibodies: Anti-Akt (#4685), anti-phospho-Akt S473 (#4060), anti-phospho-Akt T308 (#9275), anti-β Actin (#4970), anti-GSK3β (#9315), anti-phospho-GSK3β (#9336), anti-GST (#2625), and anti-V5 (#13202) were purchased from Cell Signaling Technologies (product numbers listed for each). Horseradish peroxidase-conjugated anti-rabbit and anti-mouse immunoglobulin G (IgG) antibodies were purchased from Millipore.

3D modeling: The 3D structure of AKT2 with the full allelic series was predicted using IntFOLD (27) and visualized in PyMOL (28).

In vitro kinase assays: We isolated V5-AKT2, V5-AKT2.Lys17, V5-AKT2.Thr50, V5-AKT2.Lys208, V5-AKT2.His274, and V5-AKT2.Trp467 variants from lentivirally infected and 5μg/mL blasticidin selected HeLa cell lysate with V5 agarose beads (SIGMA) and incubated with 150ng GST-GSK3β substrate peptide (Cell Signaling Technologies) and 250mM cold ATP in kinase assay buffer (Cell Signaling Technologies) for 35 minutes at 30°C.

Proliferation assay: We cultured lentiviral pLX304 V5-AKT2 variants and control empty vector infected and 5μg/mL blasticidin selected HuH7 cells in 24 well plate for 72 hours in 10% FBS /phenol red-free DMEM for 72 hours. We added WST-1 (Takara Clontech) to each well at the manufacture recommended 1:10 ratio and incubated for 4 hours at 37°C prior to absorbance measurement at 450nm with BioTek Synergy H4 plate reader.

Immunoblots: We washed cells with phosphate buffered saline and lysed in EBC buffer (120mM NaCl, 50mM TRIS-HCl (pH7.4), 50nM calyculin, cOmplete protease inhibitor cocktail (Roche), 20mM sodium fluoride, 1mM sodium pyrophosphate, 2mM ethylene glycol tetraacetic acid, 2mM ethylenediaminetetraacetic acid, and 0.5% NP-40) for 20 minutes on ice. To preclear cell lysates, we centrifuged at 12,700 rmp at 4°C for 15 minutes. We measured protein concentration with Pierce BCA protein assay kit using a BioTek Synergy H4 plate reader. We

resolved lysates on BioRad any kD mini-PROTEAN TGX polyacrylamide gels by SDS-PAGE and transferred by electrophoresis to nitrocellulose membrane (Life Technologies) at 100V for 70 minutes. We blocked membranes in 5% nonfat dry milk/ TBST (10mM Tris-HCl, 150mM NaCl, 0.2% Tween 20) buffer pH 7.6 for 30 minutes. We incubated blots with indicated antibody overnight at 4°C. The membrane was then washed in TBST, three times at 15 minute intervals, before 1 hour secondary horseradish peroxidase-conjugated antibody incubation at room temperature. We again washed nitrocellulose membranes in TBST, three times for 15 minutes, prior to enhanced chemiluminescent substrate detection (Pierce).

Statistical analysis

The quantified results of the *in vitro* kinase and proliferation assays were normalized to internal control values for each replicate. We used generalized linear models of the quantified assay results to assess effects of variants within and across replicate rounds, allowing for interaction by replicate. The graphical representation was produced using functions in the effects (v 3.0-3) package in R.

Gene Expression Studies

Study samples

GTEx: We compared the expression pattern of *AKT2* to the two other members of the *AKT* gene family, *AKT1* and *AKT3,* using multi-tissue RNA sequencing (RNA-seq) data from the pilot phase of the GTEx project (dbGaP accession number: phs000424.v3.p1) in 44 tissues with data from more than one individual. Detailed procedures for sample collection, RNA extraction, RNA-seq, and gene and transcript quantifications have been previously described (29). *EuroBATs:* Samples from photo protected subcutaneous adipose tissue from 766 twins were extracted (130 unrelated individuals, 131 monozygotic and 187 dizygotic twin pairs) and

processed as previously described (30; 31). *METSIM*: Subcutaneous fat biopsy samples were obtained from a sample of 770 participants from the METSIM study and processed as previously described (32).

Phenotypes

We studied the association of age, body mass index (BMI) and fasting insulin levels with gene expression levels and with expression-associated SNPs (eQTLs) in the *AKT2* region. Age and sex were available for the GTEx study samples. In additional to age and BMI, fasting insulin level was measured at the same time point as the fat biopsies in the EuroBATs sample data, following a previously described protocol (33). Baseline age, BMI and fasting insulin levels were used for the METSIM study participants (34)

Statistical analysis

The comparison of expression levels of *AKT2* versus *AKT1*, and *AKT2* versus *AKT3* was performed using log2-transformed reads per kilobase per million mapped reads (RPKMs). The percent increase in *AKT2* expression was calculated with the following formula: 2^log-fold-change (*AKT2* vs *AKT1*). We studied BMI, age, and fasting insulin (not available in GTEx data) associations with *AKT2* expression using linear mixed models as implemented in the lme4 package in R. The gene expression RPKM values were inverse variance rank normalized for these analyses. Covariates included study-specific fixed and random effects (see Supplementary Note 4 for additional details on each cohort), using sex, BMI and age as additional fixed effects as appropriate. The expression quantitative trait loci (eQTL) analysis was performed on single nucleotide polymorphisms (SNPs) within a 1 Mb of *AKT2* using linear mixed models to assess the association of the SNPs with the inverse normalized RPKM expression values.

RESULTS

Genetic association studies

We tested the association of FI and FG with 390,225 variants from exome sequence data (GoT2D and T2D-GENES studies) and 109,215 variants derived from exome array genotyping (GoT2D studies) (7) (individual study λ_{GC}<1.06; Supplementary Figure S1). We examined variants that had been previously associated with FG and FI (3; 18). Of 28 FG and 14 FI loci with the reported SNPs or close proxies in our data set, 13 FG and four FI showed directionally consistent significant associations. Among the remaining GWAS loci not significant in our data, we observed directionally consistent associations in 14/15 FG and 9/10 FI loci ($P_{enrichment}$ =5×10^{-4} for FG and 0.01 for FI) (Supplementary Note 1; Supplementary Table 2).

In addition, we identified a novel significant single variant association between rs184042322 and FI (MAF=1.2%, P=1.2×10^{-7}), a coding variant in *AKT2* (*V-AKT Murine Thymoma Viral Oncogene Homolog 2*) where amino acid Pro50 is substituted with a threonine (NP_001617.1:p.Pro50Thr) (Figure 1; Supplementary Figure S1). The same allele drove a significant FI signal for *AKT2* in gene-based analysis (P=6.1×10^{-7}), in which we discovered two additional significant gene-based associations between *GIMAP8* and FG (P_{PTV}=2.3×10^{-6}), and between *NDUFAF1* and FI ($P_{PTV+NSBroad}$=9.2×10^{-7}) (Supplementary Figure S2; Supplementary Table 2D).

In an effort to replicate the single variant association of *AKT2* Pro50Thr with FI, we aggregated the allele frequency estimates of *AKT2* Pro50Thr in our data with data from the CHARGE consortium and the four Finnish studies. In ExAC, rs184042322 is multi-allelic (p.Pro50Thr and p.Pro50Ala) but Pro50Ala is observed only twice in the Latino population sample and not seen in our exome sequencing data, which includes 1,021 individuals of Hispanic ancestry. *AKT2* Pro50Thr was observed at a much higher frequency in Finnish individuals

(MAF=1.1%) than other European (MAF=0.2%), African American (MAF=0.01%), Asian (MAF<0.01%), or Hispanic (MAF<0.01%) individuals (Figure 1). We replicated the association between FI and *AKT2* Pro50Thr by meta-analysis of the association in the four Finnish studies (P=5.4×10^{-4}, N=5,747) with the discovery studies ($P_{combined}$=9.98×10^{-10}, N=25,316). We observed no evidence of effect-size heterogeneity between studies ($P_{Heterogeneity}$=0.76). The minor T allele was associated with a 12% (95% CI=7%-18%) increase in FI levels in the discovery and replication studies, a per allele effect of 10.4pmol/L (95% CI=6.6-14.3pmol/L).

The serine/threonine protein kinases AKT1, AKT2, and AKT3 are conserved across all vertebrates (Figure 2). Pro50 and the seven preceding residues in the pleckstrin homology (PH) domain appear to be specific for the AKT2 isoform. Population genetic studies show a strong intolerance to missense and loss of function variation in *AKT2* (Supplementary Note 2; Supplementary Figure S3; Supplementary Figure S4; Supplementary Table 3). Notably, in ExAC data, *AKT2* contains fewer missense variants than expected (the missense constraint metric, Z=3.5, is in the 94th percentile of all genes) and extreme constraint against loss-of-function (LoF) variation (estimated probability of being LoF intolerant (pLI)=1).

AKT2 is a primary transducer of phosphoinositide 3-kinase (PI3K) signaling downstream of the insulin receptor and is responsible for mediating the physiological effects of insulin in tissues including liver, skeletal muscle, and adipose. *Akt2* null mice are characterized by hyperglycemia and hyperinsulinemia, and some develop diabetes (35; 36). In humans, highly penetrant rare alleles in *AKT2* cause familial partial lipodystrophy and hypoinsulinemic hypoglycemia with hemihypertrophy (Glu17Lys) (37; 38) and a syndrome featuring severe insulin resistance, hyperinsulinemia, and diabetes mellitus (Arg274His) (39). Additional rare

alleles have been observed in individuals with severe insulin resistance (Arg208Lys and Arg467Trp) but no variant has been associated with glycemic traits at the population level (40).

Given the spectrum of diseases and traits associated with *AKT2* (41), we hypothesized that *AKT2* Pro50Thr would be associated with features of metabolic syndrome or lipodystrophy. In quantitative trait analysis in the initial discovery and replication cohorts, we did observe a constellation of features indicative of a milder 'lipodystrophy-like phenotype' associated with the rare allele: associations with increased 2-hour insulin values (effect=0.2 SD of log-transformed 2-hour insulin, 95% CI=0.1-0.4; $P=7.9\times10^{-8}$, N=14,150), lower insulin sensitivity (effect=-0.3 SD of the log-transformed Matsuda index, 95% CI=-0.5 to -0.2, $P=1.2\times10^{-6}$, N=8,566), and increased risk of type 2 diabetes (odds ratio (OR)=1.05 95% CI=1.0-1.1, $P=8.1\times10^{-5}$; 9,783 type 2 diabetes cases; 22,662 controls), with no effects on fasting glucose, postprandial glucose, or fasting lipid levels ($P\geq0.01$; Supplementary Table 4). In the T2D-GENES exome sequencing data where FG and FI levels were available in diabetic individuals, we observed one individual who was homozygous for the P50T allele with FI and FG levels in the 99.8th and 98.8th percentiles, respectively. There was a significant difference in trait distributions by P50T genotype (FI P=0.002; FG P=0.02; Supplementary Figure S5; Supplementary Table 4). Next, we used electronic health records available in the Finnish METSIM and FINRISK cohorts to characterize the impact of *AKT2* Pro50Thr on disease risk. We found no evidence for association with any cancer, polycystic ovary disease, or acanthosis nigricans (Supplementary Table 5); however, these tests are underpowered due to the low number of cases and potential for misclassification. Nor did we find evidence for enrichment of low-frequency associations in any *AKT2* related pathways or genes implicated in monogenic

forms of glycemic disease (Supplementary Note 3; Supplementary Table 6; Supplementary Table 7; Supplementary Figure S6; Supplementary Figure S7).

***In vitro* functional studies**

To understand the functional consequences of the *AKT2* Pro50Thr variant on the protein, we investigated protein expression, activation, kinase activity, and downstream effector phosphorylation.

First, we used *in silico* classifiers that predict potential functional consequences of alleles on protein function. Two of the five classifiers predicted *AKT2* Pro50Thr to be deleterious (Supplementary Table 3). Second, we used 3D models of *AKT2* viewed in the PyMol software, which predicted that the Pro50Thr variant causes a change in the conformations of the lipid binding PH domain (Figure 3, Supplementary Figure S8). We hypothesized that the variant protein is inefficiently recruited to the plasma membrane thereby impacting AKT2 phosphorylation and downstream activity.

To assess the molecular and cellular consequence of the *AKT2* Thr50 variant on protein function, we performed a comparative analysis of AKT2-Thr50 with inactivating and activating alleles implicated in monogenic disorders of insulin signaling. Analysis of AKT2-Thr50 expression showed that while AKT2 protein levels remained unchanged, there was a partial loss of AKT2-Thr50 phosphorylation at its activation sites (Thr308 and Ser473) in HeLa cells, suggesting impaired AKT2 signaling (Figure 3; Supplementary Figure S9). Similar effects were observed in human liver derived HuH7 cells (Supplementary Figure S10). AKT2-Thr50 also showed a reduced ability to phosphorylate its downstream target glycogen synthase kinase 3 beta (GSK3β). These defects in AKT2-Thr50 activity were confirmed through an *in vitro* kinase assay ($P<0.01$) (Figure 3). AKT2-Thr50 showed a similar decrease in kinase function to the

lipodystrophy-causing AKT2-His274 variant. Using a four-hour time course analysis of AKT2 activity, we verified a reduction in both maximally phosphorylated Thr308 and Ser473 in AKT2-Thr50 (Supplementary Figure S11). To understand how this loss of activity could manifest as a defect in a known cellular function of AKT2 (42), we determined the impact of AKT2-Thr50 on cell proliferation in HuH7 cells. While the addition of AKT2 stimulated hepatocyte proliferation, the response to AKT2-Thr50 was reduced (effect=-1.2, $P<1.0\times10^{-3}$) (Figure 3C; Supplementary Figure S12).

Gene expression studies

We queried RNA sequencing data from the Genotype Tissue Expression (GTEx) Project and found that, in agreement with previous studies (43), *AKT2* is highly and ubiquitously expressed across all tissues (44 tissue types, 3-156 individuals/tissue). Notably the *AKT2* Pro50Thr containing exon is expressed in all tissues and individuals (Supplementary Figure S13), suggesting that the PH domain is important to AKT2 function (44)**.** Of the three *AKT* homologs, *AKT2* had 1.4-fold higher expression in skeletal muscle than *AKT1* ($P=1.5\times10^{-19}$) and 11-fold higher expression than *AKT3* ($P=7.8\times10^{-91}$). Skeletal muscle was the only tested tissue displaying such pronounced *AKT2* enrichment (Figure 2; Supplementary Note 4; Supplementary Figure S14; Supplementary Table 8).

Motivated by the age-related loss of adipose tissue in *Akt2* null mice (35; 36) and the growth and lipodystrophy phenotypes in carriers of fully-penetrant alleles (37-40), we examined associations of expression levels of *AKT2* with BMI, FI, and age in the three adipose tissue data sets (Supplementary Table 9). We found an association between lower BMI levels and higher *AKT2* expression in two cohorts (EuroBATS effect=-0.07 SD, $P=6.1\times10^{-28}$; METSIM effect=-0.06 SD, $P=8.1\times10^{-8}$) and also observed that higher *AKT2* expression was associated with lower

log-transformed FI (EuroBATS, effect=-0.04 SD, $P=1.1\times10^{-3}$, METSIM, effect=-0.4 SD, $P=3.3\times10^{-11}$). We next tested for gene expression quantitative trait loci (eQTL) and found an eQTL in the 5'UTR of *AKT2* (rs11880261; MAF=35%; $r^2=0.002$, D'=0.47 in the Finnish 1000 Genomes samples) with the common allele associated with lower *AKT2* expression levels (METSIM $P=6.9\times10^{-14}$; EuroBATS $P=2.3\times10^{-8}$; GTEx $P=0.08$) (Supplementary Figure S15). No association was detected between rs11880261 and FI levels, suggesting that the common variant eQTL does not drive the initial FI association (Supplementary Note 4; Supplementary Table 10).

Discussion

Meta-analyses of exome sequence and array genotyping data in up to 38,339 normoglycemic individuals enabled the discovery, characterization, and functional validation of a FI association with a low-frequency *AKT2* coding variant. Rare, penetrant variants in genes encoding components of the insulin signaling pathway, including *AKT2*, cause monogenic but heterogeneous glycemic disorders (45). In parallel, common alleles in or near many of these genes impact FI levels —the *AKT2* Pro50Thr association shows an effect 5 to 10 times larger than those of these previous published associations (3). This discovery expands both the known genetic architecture of glucose homeostasis and the allelic spectrum for *AKT2* coding variants associated with glucose homeostasis into the low-frequency range, and highlights the effects of both locus and allelic heterogeneity (Figure 4).

Individuals of Finnish ancestry drove the *AKT2* Pro50Thr association signal. This demonstrates the value of association studies in different ancestries where frequencies of rare alleles may increase due to selective pressure or stochastic changes from population bottlenecks and genetic drift. The allele associated with increased FI most likely rose to a higher frequency

due to genetic drift and exists within the spectrum of rare and low-frequency variation observed in Finland, the excess of which facilitates the study of complex trait associations (46).

While the *AKT2* Pro50Thr allele shows a strong effect on all of the insulin measures and modest increased type 2 diabetes risk (OR=1.05) we see no effect on any of the glucose measures in individuals without diabetes. Due to the effects of both type 2 diabetes and its treatment on glucose homeostasis, we have not tested genetic associations of FG and FI in individuals with type 2 diabetes, although we observed a diabetic individual homozygous for P50T with extreme FI and FG levels. The mechanism for such heterogeneous effects is unclear and detailed *in vivo* physiological studies are needed.

We leveraged similar findings to generate hypotheses for future work on *AKT2* and downstream targets to further illuminate tissue-specific mechanisms. All reported carriers of the lipodystrophy causing *AKT2* Arg274His allele are hyperinsulinemic, and three of the four carriers have diabetes mellitus (39). These observations are similar to the ones made for *TBC1D4* (which encodes a protein that acts as a substrate immediately downstream of *AKT2* in the PI3K pathway). In *TBC1D4* a population specific, protein-truncating variant (Arg684Ter) is associated with increased type 2 diabetes risk (OR = 10.3), increased postprandial glucose and insulin levels, and a modest decrease in FI and FG levels (6) (Figure 4). Another stop codon allele in *TBC1D4,* Arg363Ter that is rare (not observed in ExAC) has been reported with a modest elevation in FI levels but extreme postprandial hyperinsulinemia and acanthosis nigricans (47). siRNA-mediated gene knock-down of *AKT2* in human primary myotubes completely abolishes insulin action on glucose uptake and glycogen synthesis (48), which highlights the importance of an intact AKT2-TBC1D4 signaling pathway in the regulation of insulin sensitivity in humans. *TBC1D4* is ubiquitously expressed with adipose and skeletal muscle tissue ranking among the

tissues with highest expression in GTEx. *TBC1D4* Arg363Ter seems to have an effect in adipocytes (47), while Arg684Ter falls in an exon that is exclusively expressed in skeletal and heart muscle (6; 49). This is a likely cause of the *TBC1D4* Arg684Ter tissue specificity, which appears to differ from the other *TBC1D4* Arg363Ter variant as well as the *AKT2* variants.

The phenotypes exhibited by carriers of rare, penetrant *AKT2* alleles reflect differential AKT2 activation with kinetically inactivating variants resulting in hyperinsulinemia and lipodystrophy while kinetically activating variants lead to hypoglycemia (37-39). The decrease of cellular proliferation we observe demonstrates that the downstream signaling changes caused by AKT2-Thr50 are sufficient in hepatocytes to impair AKT2 function at the cellular level while maintaining varying portions of regulatory capacity. Along with the observed association with increased fasting insulin levels in human populations, these results support *AKT2* Pro50Thr as a *partial* loss-of-function variant. The inactivating *AKT2* Pro50Thr variant contrasts with the known activating *AKT2* Glu17Lys mutation and showcases bidirectional effects within the PH domain of AKT2. While the Pro50 residue is conserved in AKT2 throughout all vertebrates, the variant lies within the PH domain that is not conserved between AKT isoforms (Figure 2). These residues, harboring the Pro50 variant, may functionally distinguish AKT2 from AKT1 and AKT3. Although AKT isoforms are activated in the same mechanism within the PI3K pathway downstream of insulin, the *Akt2*$^{-/-}$ mouse is the only knockout of the gene family to be characterized by insulin resistance and diabetes (35; 50-52). A deeper understanding of what makes the AKT2 isoform distinct could offer potential sites for therapeutic intervention and enable more targeted approaches to disease prevention.

Acknowledgements

C.M.L. is the guarantor of this work and, as such, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. The authors have no relevant conflicts of interest to disclose.

Author Contributions:

Sample Collection and Phenotyping: NG, A Mahajan, NPB, C Ladenvall, JB-J, NRR, NWR, RAS, APG, AUJ, CJG, CB, D Buck, GB, GJ, HMS, JRH, J Murphy, JMJ, J Trakalo, KSS, MM, MN, M Hollensted, RO, SG, ARW, ATH, HEA, AC, RAD, A Stančáková, AHR, A Metspalu, AJF, A-CS, A Käräjämäki, YAK, RA, A Swift, TA, BL, BG, BIF, B-GH, C Meisinger, CG, C Langenberg, D Pasko, D Aguilar, D Bowden, DH, EST, EC, C-YC, WYL, EM, SPF, FBH, G Atzmon, GWSr, DEH, HG, HK, HO, HATJr, TI, JSK, J Sehmi, J Lindstrom, J Kravic, JEC, CPJ, JEB, J Kriebel, JH, J Li, J Fadista, JCC, JCL, KRO, KSC, C-CK, LLB, J-YL, LK, DML, LH, L Milani, J Liu, L Liang, M Loh, MO-M, MW, MM-N, TM, MG, MR, MCYN, NDP, NN, LQ, NJW, NB, OM, OR, PJH, PWF, PN, A Peters, QQ, RM, S-TT, S Kumar, SKM, SPO'R, S Puppala, K Strauch, TMF, TK, TE, FT, BT, TVV, TYW, TAL, T Lauritzen, T Forsén, TIP, UA, VSF, WRS, YSC, ADM, ASFD, AL, BI, CNAP, FSC, CC, EI, FK, GLS, I Brandslund, J Tuomilehto, J Kuusisto, L Lannfelt, L Lind, LG, MEJ, MU, OP, RR, RNB, T Tuomi, TDS, TH, TJ, VS, GIB, JGW, JB, NJC, RD, KLM, M Laakso, CLH, APM, MB, D Altshuler, MIM

Replication and Expression Studies: T Tukiainen, AV, AAB, YW, A Palotie, AJ, JGE, OTR, S Koskinen, T Lehtimäki, JW, AYC, RAS, MOG, VS, JD, SR, JCF, JBM, ML, KLM

Data Production (Sequencing and Genotyping): XS, NG, A Mahajan, CF, NPB, C Hartl, C Ladenvall, JB-J, NRR, NWR, APG, AUJ, CJG, CB, D Buck, GB, GJ, HMS, JRH, J Murphy, JMJ, J Trakalo, KSS, MM, MN, M Hollensted, RO, PSC, SG, MOC, MD, EB, YF, MHdA, K Shakir, RP, T Fennell, TS, TW, TMS, K Stirrups, TM, PD, MB, MIM

Variant Calling & Panel Generation: MAR, KJG, HMK, GJ, BMN, GG, J Maguire, J Carey, JDS, JIG, S Purcell

Statistical Analysis: A Manning, HMH, JG, XS, T Tukiainen, P Fontanillas, NG, MAR, A Mahajan, AEL, P Cingolani, T Pers, J Flannick, CF, ERG, KJG, HKI, TMT, A Kumar, NPB, C Hartl, C Ladenvall, HMK, JB-J, YC, JRBP, LJS, C Ma, MvdB, L Moutsianas, NRR, RDP, TWB, TG, NWR, APG, AUJ, CJG, CB, D Buck, GB, GJ, HMS, JRH, J Murphy, JMJ, J Trakalo, KSS, MM, MN, M Hollensted, RO, SG, JBM, APM, MB, MIM, CML

Functional Studies: JG, SJ, ALG

Wrote the paper: A Manning, HMH, JG, XS, T Tukiainen, P Fontanillas, JCF, MB, MIM, ALG, CML

Study Design: XS, LJS, ATH, HEA, RAD, BG, EST, G Atzmon, JSK, CPJ, JCC, KSC, J-YL, DML, TM, TMF, TIP, YSC, C Hu, GRC, D Bharadwaj, PJD, D Prabhakaran, EZ, I Barroso, J Scott, J Chan, GM, MJD, M Sandhu, NT, PE, P Froguel, RCWM, RS, SBE, YYT, T Park, T Fingerlin, WJ, RMW, J Tuomilehto, LG, GIB, G Abecasis, JGW, JB, M Seielstad, NJC, RD, JD, IP, JCF, KLM, M Laakso, JBM, CLH, APM, MB, D Altshuler, MIM, ALG, CML

Study Supervision: GIB, G Abecasis, JGW, JB, M Seielstad, NJC, RD, JCF, KLM, JBM, CLH, APM, MB, D Altshuler, MIM, ALG, CML

We thank the more than 44,412 volunteers who participated in this study. We acknowledge the following funding sources: Academy of Finland (129293, 128315, 129330, 131593, 139635, 139635, 121584, 126925, 124282, 129378, 258753); Action on Hearing Loss (G51); Ahokas Foundation; American Diabetes Association (#7-12-MN-02); Atlantic Canada Opportunities Agency; Augustinus foundation; Becket foundation; Benzon Foundation; Biomedical Research Council; British Heart Foundation (SP/04/002); Canada Foundation for Innovation; Commission of the European Communities, Directorate C-Public Health (2004310); Copenhagen County; Danish Centre for Evaluation and Health Technology Assessment; Danish Council for Independent Research; Danish Heart Foundation (07-10-R61-A1754-B838-22392F); Danish Medical Research Council; Danish Pharmaceutical Association; Emil Aaltonen Foundation; European Research Council Advanced Research Grant; European Union FP7 (EpiMigrant, 279143; FP7/2007-2013; 259749); Finland's Slottery Machine Association; Finnish Cultural Foundation; Finnish Diabetes Research Foundation; Finnish Foundation for Cardiovascular Research; Finnish Foundation of Cardiovascular Research; Finnish Medical Society; Finnish National Public Health Institute; Finska Läkaresällskapet; Folkhälsan Research Foundation; Foundation for Life and Health in Finland; German Center for Diabetes Research (DZD) ; German Federal Ministry of Education and Research; Health Care Centers in Vasa, Närpes and Korsholm; Health Insurance Foundation (2012B233) ; Helsinki University Central Hospital Research Foundation; Hospital districts of Pirkanmaa, Southern Ostrobothnia, North Ostrobothnia, Central Finland, and Northern Savo; Ib Henriksen foundation; Juho Vainio Foundation; Korea Centers for Disease Control and Prevention (4845–301); Korea National Institute of Health (2012-N73002-00); Li Ka Shing Foundation; Liv och Hälsa; Lundbeck Foundation; Marie-Curie Fellowship (PIEF-GA-2012-329156); Medical Research Council (G0601261, G0900747-91070, G0601966, G0700931); Ministry of Education in Finland; Ministry of Social Affairs and Health in Finland; MRC-PHE Centre for Environment and Health;Municipal Heath Care Center and Hospital in Jakobstad; Närpes Health Care Foundation; National Institute for Health Research (RP-PG-0407-10371); National Institutes of Health (U01 DK085526, U01 DK085501, U01 DK085524, U01 DK085545, U01 DK085584, U01 DK088389, RC2-DK088389, DK085545, DK098032, HHSN268201300046C, HHSN268201300047C, HHSN268201300048C, HHSN268201300049C, HHSN, R01MH107666 and K12CA139160268201300050C, U01 DK062370, R01 DK066358, U01DK085501, R01HL102830, R01DK073541, PO1AG027734, R01AG046949, 1R01AG042188, P30AG038072, R01 MH101820, R01MH090937, P30DK020595, R01 DK078616, NIDDK K24 DK080140, 1RC2DK088389, T32GM007753); National Medical Research Council; National Research Foundation of Korea (NRF-2012R1A2A1A03006155); Nordic Center of Excellence in Disease Genetics; Novo Nordisk; Ollqvist Foundation; Orion-Farmos Research Foundation; Paavo Nurmi Foundation; Perklén Foundation; Samfundet Folkhälsan; Signe and Ane Gyllenberg Foundation; Sigrid Juselius Foundation; Social Insurance Institution of Finland; South East Norway Health Authority (2011060); Swedish Cultural Foundation in Finland; Swedish Heart-Lung Foundation; Swedish Research Council; Swedish Research Council (Linné and Strategic Research Grant); The American Federation for Aging Research; The Einstein Glenn Center; The European Commission (HEALTH-F4-2007-201413); The Finnish Diabetes Association; The Folkhälsan Research Foundation; The Påhlssons

Foundation; The provinces of Newfoundland and Labrador, Nova Scotia, and New Brunswick; The Sigrid Juselius Foundation; The Skåne Regional Health Authority; The Swedish Heart-Lung Foundation; Timber Merchant Vilhelm Bang's Foundation; Turku University Foundation; Uppsala University; Wellcome Trust (064890, 083948, 085475, 086596, 090367, 090532, 092447, 095101/Z/10/Z, 200837/Z/16/Z, 095552, 098017, 098381, 098051, 084723, 072960/2/03/2, 086113/Z/08/Z, WT098017, WT064890, WT090532, WT098017, 098051, WT086596/Z/08/A and 086596/Z/08/Z). Detailed acknowledgment of funding sources is provided in the Additional Acknowledgements section of the Supplementary Materials.

References

1. Kahn SE, Hull RL, Utzschneider KM: Mechanisms linking obesity to insulin resistance and type 2 diabetes. Nature 2006;444:840-846
2. Phillips DI, Clark PM, Hales CN, Osmond C: Understanding oral glucose tolerance: comparison of glucose or insulin measurements during the oral glucose tolerance test with specific measurements of insulin resistance and insulin secretion. Diabet Med 1994;11:286-292
3. Manning AK, Hivert MF, Scott RA, Grimsby JL, Bouatia-Naji N, Chen H, Rybin D, Liu CT, Bielak LF, Prokopenko I, Amin N, Barnes D, Cadby G, Hottenga JJ, Ingelsson E, Jackson AU, Johnson T, Kanoni S, Ladenvall C, Lagou V, Lahti J, Lecoeur C, Liu Y, Martinez-Larrad MT, Montasser ME, Navarro P, Perry JR, Rasmussen-Torvik LJ, Salo P, Sattar N, Shungin D, Strawbridge RJ, Tanaka T, van Duijn CM, An P, de Andrade M, Andrews JS, Aspelund T, Atalay M, Aulchenko Y, Balkau B, Bandinelli S, Beckmann JS, Beilby JP, Bellis C, Bergman RN, Blangero J, Boban M, Boehnke M, Boerwinkle E, Bonnycastle LL, Boomsma DI, Borecki IB, Bottcher Y, Bouchard C, Brunner E, Budimir D, Campbell H, Carlson O, Chines PS, Clarke R, Collins FS, Corbaton-Anchuelo A, Couper D, de Faire U, Dedoussis GV, Deloukas P, Dimitriou M, Egan JM, Eiriksdottir G, Erdos MR, Eriksson JG, Eury E, Ferrucci L, Ford I, Forouhi NG, Fox CS, Franzosi MG, Franks PW, Frayling TM, Froguel P, Galan P, de Geus E, Gigante B, Glazer NL, Goel A, Groop L, Gudnason V, Hallmans G, Hamsten A, Hansson O, Harris TB, Hayward C, Heath S, Hercberg S, Hicks AA, Hingorani A, Hofman A, Hui J, Hung J, Jarvelin MR, Jhun MA, Johnson PC, Jukema JW, Jula A, Kao WH, Kaprio J, Kardia SL, Keinanen-Kiukaanniemi S, Kivimaki M, Kolcic I, Kovacs P, Kumari M, Kuusisto J, Kyvik KO, Laakso M, Lakka T, Lannfelt L, Lathrop GM, Launer LJ, Leander K, Li G, Lind L, Lindstrom J, Lobbens S, Loos RJ, Luan J, Lyssenko V, Magi R, Magnusson PK, Marmot M, Meneton P, Mohlke KL, Mooser V, Morken MA, Miljkovic I, Narisu N, O'Connell J, Ong KK, Oostra BA, Palmer LJ, Palotie A, Pankow JS, Peden JF, Pedersen NL, Pehlic M, Peltonen L, Penninx B, Pericic M, Perola M, Perusse L, Peyser PA, Polasek O, Pramstaller PP, Province MA, Raikkonen K, Rauramaa R, Rehnberg E, Rice K, Rotter JI, Rudan I, Ruokonen A, Saaristo T, Sabater-Lleal M, Salomaa V, Savage DB, Saxena R, Schwarz P, Seedorf U, Sennblad B, Serrano-Rios M, Shuldiner AR, Sijbrands EJ, Siscovick DS, Smit JH, Small KS, Smith NL, Smith AV, Stancakova A, Stirrups K, Stumvoll M, Sun YV, Swift AJ, Tonjes A, Tuomilehto J, Trompet S, Uitterlinden AG, Uusitupa M, Vikstrom M, Vitart V, Vohl MC, Voight BF, Vollenweider P, Waeber G, Waterworth DM, Watkins H, Wheeler E, Widen E, Wild SH, Willems SM, Willemsen G, Wilson JF, Witteman JC, Wright AF, Yaghootkar H, Zelenika D, Zemunik T, Zgaga L, Replication DIG, Meta-analysis C, Multiple Tissue Human Expression Resource C, Wareham NJ, McCarthy MI, Barroso I, Watanabe RM, Florez JC, Dupuis J, Meigs JB, Langenberg C: A genome-wide approach accounting for body mass index identifies genetic variants influencing fasting glycemic traits and insulin resistance. Nat Genet 2012;44:659-669
4. SIGMA Type 2 Diabetes Consortium: Association of a low-frequency variant in HNF1A with type 2 diabetes in a Latino population. JAMA 2014;311:2305-2314
5. SIGMA Type 2 Diabetes Consortium: Sequence variants in SLC16A11 are a common risk factor for type 2 diabetes in Mexico. Nature 2014;506:97-101

6. Moltke I, Grarup N, Jorgensen ME, Bjerregaard P, Treebak JT, Fumagalli M, Korneliussen TS, Andersen MA, Nielsen TS, Krarup NT, Gjesing AP, Zierath JR, Linneberg A, Wu X, Sun G, Jin X, Al-Aama J, Wang J, Borch-Johnsen K, Pedersen O, Nielsen R, Albrechtsen A, Hansen T: A common Greenlandic TBC1D4 variant confers muscle insulin resistance and type 2 diabetes. Nature 2014;512:190-193
7. Fuchsberger C, Flannick J, Teslovich TM, Mahajan A, Agarwala V, Gaulton KJ, Ma C, Fontanillas P, Moutsianas L, McCarthy DJ, Rivas MA, Perry JR, Sim X, Blackwell TW, Robertson NR, Rayner NW, Cingolani P, Locke AE, Tajes JF, Highland HM, Dupuis J, Chines PS, Lindgren CM, Hartl C, Jackson AU, Chen H, Huyghe JR, van de Bunt M, Pearson RD, Kumar A, Muller-Nurasyid M, Grarup N, Stringham HM, Gamazon ER, Lee J, Chen Y, Scott RA, Below JE, Chen P, Huang J, Go MJ, Stitzel ML, Pasko D, Parker SC, Varga TV, Green T, Beer NL, Day-Williams AG, Ferreira T, Fingerlin T, Horikoshi M, Hu C, Huh I, Ikram MK, Kim BJ, Kim Y, Kim YJ, Kwon MS, Lee J, Lee S, Lin KH, Maxwell TJ, Nagai Y, Wang X, Welch RP, Yoon J, Zhang W, Barzilai N, Voight BF, Han BG, Jenkinson CP, Kuulasmaa T, Kuusisto J, Manning A, Ng MC, Palmer ND, Balkau B, Stancakova A, Abboud HE, Boeing H, Giedraitis V, Prabhakaran D, Gottesman O, Scott J, Carey J, Kwan P, Grant G, Smith JD, Neale BM, Purcell S, Butterworth AS, Howson JM, Lee HM, Lu Y, Kwak SH, Zhao W, Danesh J, Lam VK, Park KS, Saleheen D, So WY, Tam CH, Afzal U, Aguilar D, Arya R, Aung T, Chan E, Navarro C, Cheng CY, Palli D, Correa A, Curran JE, Rybin D, Farook VS, Fowler SP, Freedman BI, Griswold M, Hale DE, Hicks PJ, Khor CC, Kumar S, Lehne B, Thuillier D, Lim WY, Liu J, van der Schouw YT, Loh M, Musani SK, Puppala S, Scott WR, Yengo L, Tan ST, Taylor HA, Jr., Thameem F, Wilson G, Wong TY, Njolstad PR, Levy JC, Mangino M, Bonnycastle LL, Schwarzmayr T, Fadista J, Surdulescu GL, Herder C, Groves CJ, Wieland T, Bork-Jensen J, Brandslund I, Christensen C, Koistinen HA, Doney AS, Kinnunen L, Esko T, Farmer AJ, Hakaste L, Hodgkiss D, Kravic J, Lyssenko V, Hollensted M, Jorgensen ME, Jorgensen T, Ladenvall C, Justesen JM, Karajamaki A, Kriebel J, Rathmann W, Lannfelt L, Lauritzen T, Narisu N, Linneberg A, Melander O, Milani L, Neville M, Orho-Melander M, Qi L, Qi Q, Roden M, Rolandsson O, Swift A, Rosengren AH, Stirrups K, Wood AR, Mihailov E, Blancher C, Carneiro MO, Maguire J, Poplin R, Shakir K, Fennell T, DePristo M, Hrabe de Angelis M, Deloukas P, Gjesing AP, Jun G, Nilsson P, Murphy J, Onofrio R, Thorand B, Hansen T, Meisinger C, Hu FB, Isomaa B, Karpe F, Liang L, Peters A, Huth C, O'Rahilly SP, Palmer CN, Pedersen O, Rauramaa R, Tuomilehto J, Salomaa V, Watanabe RM, Syvanen AC, Bergman RN, Bharadwaj D, Bottinger EP, Cho YS, Chandak GR, Chan JC, Chia KS, Daly MJ, Ebrahim SB, Langenberg C, Elliott P, Jablonski KA, Lehman DM, Jia W, Ma RC, Pollin TI, Sandhu M, Tandon N, Froguel P, Barroso I, Teo YY, Zeggini E, Loos RJ, Small KS, Ried JS, DeFronzo RA, Grallert H, Glaser B, Metspalu A, Wareham NJ, Walker M, Banks E, Gieger C, Ingelsson E, Im HK, Illig T, Franks PW, Buck G, Trakalo J, Buck D, Prokopenko I, Magi R, Lind L, Farjoun Y, Owen KR, Gloyn AL, Strauch K, Tuomi T, Kooner JS, Lee JY, Park T, Donnelly P, Morris AD, Hattersley AT, Bowden DW, Collins FS, Atzmon G, Chambers JC, Spector TD, Laakso M, Strom TM, Bell GI, Blangero J, Duggirala R, Tai ES, McVean G, Hanis CL, Wilson JG, Seielstad M, Frayling TM, Meigs JB, Cox NJ, Sladek R, Lander ES, Gabriel S, Burtt NP, Mohlke KL, Meitinger T, Groop L, Abecasis G, Florez JC, Scott LJ, Morris AP, Kang HM, Boehnke M, Altshuler D, McCarthy MI: The genetic architecture of type 2 diabetes. Nature 2016;
8. Raitakari OT, Juonala M, Ronnemaa T, Keltikangas-Jarvinen L, Rasanen L, Pietikainen M, Hutri-Kahonen N, Taittonen L, Jokinen E, Marniemi J, Jula A, Telama R, Kahonen M, Lehtimaki T, Akerblom HK, Viikari JS: Cohort profile: the cardiovascular risk in Young Finns Study. Int J Epidemiol 2008;37:1220-1226
9. Eriksson JG: Epidemiology, genes and the environment: lessons learned from the Helsinki Birth Cohort Study. J Intern Med 2007;261:418-425
10. Perttila J, Merikanto K, Naukkarinen J, Surakka I, Martin NW, Tanhuanpaa K, Grimard V, Taskinen MR, Thiele C, Salomaa V, Jula A, Perola M, Virtanen I, Peltonen L, Olkkonen VM: OSBPL10, a novel candidate gene for high triglyceride trait in dyslipidemic Finnish subjects, regulates cellular lipid metabolism. J Mol Med (Berl) 2009;87:825-835

11. Vartiainen E, Laatikainen T, Peltonen M, Juolevi A, Mannisto S, Sundvall J, Jousilahti P, Salomaa V, Valsta L, Puska P: Thirty-five-year trends in cardiovascular risk factors in Finland. Int J Epidemiol 2010;39:504-518
12. Grove ML, Yu B, Cochran BJ, Haritunians T, Bis JC, Taylor KD, Hansen M, Borecki IB, Cupples LA, Fornage M, Gudnason V, Harris TB, Kathiresan S, Kraaij R, Launer LJ, Levy D, Liu Y, Mosley T, Peloso GM, Psaty BM, Rich SS, Rivadeneira F, Siscovick DS, Smith AV, Uitterlinden A, van Duijn CM, Wilson JG, O'Donnell CJ, Rotter JI, Boerwinkle E: Best practices and joint calling of the HumanExome BeadChip: the CHARGE Consortium. PLoS One 2013;8:e68095
13. Matsuda M, DeFronzo RA: Insulin sensitivity indices obtained from oral glucose tolerance testing: comparison with the euglycemic insulin clamp. Diabetes Care 1999;22:1462-1470
14. Peloso GM, Auer PL, Bis JC, Voorman A, Morrison AC, Stitziel NO, Brody JA, Khetarpal SA, Crosby JR, Fornage M, Isaacs A, Jakobsdottir J, Feitosa MF, Davies G, Huffman JE, Manichaikul A, Davis B, Lohman K, Joon AY, Smith AV, Grove ML, Zanoni P, Redon V, Demissie S, Lawson K, Peters U, Carlson C, Jackson RD, Ryckman KK, Mackey RH, Robinson JG, Siscovick DS, Schreiner PJ, Mychaleckyj JC, Pankow JS, Hofman A, Uitterlinden AG, Harris TB, Taylor KD, Stafford JM, Reynolds LM, Marioni RE, Dehghan A, Franco OH, Patel AP, Lu Y, Hindy G, Gottesman O, Bottinger EP, Melander O, Orho-Melander M, Loos RJ, Duga S, Merlini PA, Farrall M, Goel A, Asselta R, Girelli D, Martinelli N, Shah SH, Kraus WE, Li M, Rader DJ, Reilly MP, McPherson R, Watkins H, Ardissino D, Project NGES, Zhang Q, Wang J, Tsai MY, Taylor HA, Correa A, Griswold ME, Lange LA, Starr JM, Rudan I, Eiriksdottir G, Launer LJ, Ordovas JM, Levy D, Chen YD, Reiner AP, Hayward C, Polasek O, Deary IJ, Borecki IB, Liu Y, Gudnason V, Wilson JG, van Duijn CM, Kooperberg C, Rich SS, Psaty BM, Rotter JI, O'Donnell CJ, Rice K, Boerwinkle E, Kathiresan S, Cupples LA: Association of low-frequency and rare coding-sequence variants with blood lipids and coronary heart disease in 56,000 whites and blacks. Am J Hum Genet 2014;94:223-232
15. Tobin MD, Sheehan NA, Scurrah KJ, Burton PR: Adjusting for treatment effects in studies of quantitative traits: antihypertensive therapy and systolic blood pressure. Statistics in medicine 2005;24:2911-2935
16. Kang HM, Sul JH, Service SK, Zaitlen NA, Kong SY, Freimer NB, Sabatti C, Eskin E: Variance component model to account for sample structure in genome-wide association studies. Nat Genet 2010;42:348-354
17. Willer CJ, Li Y, Abecasis GR: METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics 2010;26:2190-2191
18. Mahajan A, Sim X, Ng HJ, Manning A, Rivas MA, Highland HM, Locke AE, Grarup N, Im HK, Cingolani P, Flannick J, Fontanillas P, Fuchsberger C, Gaulton KJ, Teslovich TM, Rayner NW, Robertson NR, Beer NL, Rundle JK, Bork-Jensen J, Ladenvall C, Blancher C, Buck D, Buck G, Burtt NP, Gabriel S, Gjesing AP, Groves CJ, Hollensted M, Huyghe JR, Jackson AU, Jun G, Justesen JM, Mangino M, Murphy J, Neville M, Onofrio R, Small KS, Stringham HM, Syvanen AC, Trakalo J, Abecasis G, Bell GI, Blangero J, Cox NJ, Duggirala R, Hanis CL, Seielstad M, Wilson JG, Christensen C, Brandslund I, Rauramaa R, Surdulescu GL, Doney AS, Lannfelt L, Linneberg A, Isomaa B, Tuomi T, Jorgensen ME, Jorgensen T, Kuusisto J, Uusitupa M, Salomaa V, Spector TD, Morris AD, Palmer CN, Collins FS, Mohlke KL, Bergman RN, Ingelsson E, Lind L, Tuomilehto J, Hansen T, Watanabe RM, Prokopenko I, Dupuis J, Karpe F, Groop L, Laakso M, Pedersen O, Florez JC, Morris AP, Altshuler D, Meigs JB, Boehnke M, McCarthy MI, Lindgren CM, Gloyn AL, consortium TDG, Go TDc: Identification and functional characterization of G6PC2 coding variants influencing glycemic traits define an effector transcript at the G6PC2-ABCB11 locus. PLoS Genet 2015;11:e1004876
19. Wu MC, Lee S, Cai T, Li Y, Boehnke M, Lin X: Rare-variant association testing for sequencing data with the sequence kernel association test. Am J Hum Genet 2011;89:82-93
20. Liu DJ, Peloso GM, Zhan X, Holmen OL, Zawistowski M, Feng S, Nikpay M, Auer PL, Goel A, Zhang H, Peters U, Farrall M, Orho-Melander M, Kooperberg C, McPherson R, Watkins H, Willer CJ,

Hveem K, Melander O, Kathiresan S, Abecasis GR: Meta-analysis of gene-level tests for rare variant association. Nat Genet 2014;46:200-204
21. Lee S, Teslovich TM, Boehnke M, Lin X: General framework for meta-analysis of rare variants in sequencing association studies. Am J Hum Genet 2013;93:42-53
22. Marchini J, Howie B, Myers S, McVean G, Donnelly P: A new multipoint method for genome-wide association studies by imputation of genotypes. Nat Genet 2007;39:906-913
23. Exome Aggregation Consortium (ExAC), Cambridge, MA (URL: http://exac.broadinstitute.org/) Accessed 02/2015.
24. Karolchik D, Barber GP, Casper J, Clawson H, Cline MS, Diekhans M, Dreszer TR, Fujita PA, Guruvadoo L, Haeussler M, Harte RA, Heitner S, Hinrichs AS, Learned K, Lee BT, Li CH, Raney BJ, Rhead B, Rosenbloom KR, Sloan CA, Speir ML, Zweig AS, Haussler D, Kuhn RM, Kent WJ: The UCSC Genome Browser database: 2014 update. Nucleic Acids Res 2014;42:D764-770
25. Valdar WS: Scoring residue conservation. Proteins 2002;48:227-241
26. Crooks GE, Hon G, Chandonia JM, Brenner SE: WebLogo: a sequence logo generator. Genome Res 2004;14:1188-1190
27. Roche DB, Buenavista MT, Tetchner SJ, McGuffin LJ: The IntFOLD server: an integrated web resource for protein fold recognition, 3D model quality assessment, intrinsic disorder prediction, domain prediction and ligand binding site prediction. Nucleic Acids Res 2011;39:W171-176
28. Schrodinger, LLC: The PyMOL Molecular Graphics System, Version 1.3r1. 2010
29. GTEx Consortium: Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. Science 2015;348:648-660
30. Buil A, Brown AA, Lappalainen T, Vinuela A, Davies MN, Zheng H-F, Richards JB, Glass D, Small KS, Durbin R, Spector TD, Dermitzakis ET: Gene-gene and gene-environment interactions detected by transcriptome sequence analysis in twins. Nat Genet 2015;47:88-91
31. Brown AA, Buil A, Viñuela A, Lappalainen T, Zheng H-F, Richards JB, Small KS, Spector TD, Dermitzakis ET, Durbin R: Genetic interactions affecting human gene expression identified by variance association mapping. Elife 2014;3
32. Civelek M, Wu Y, Pan C, Raulerson C, Ko A, He A, Tilford C, Saleem N, Stancakova A, Scott L, Fuchsberger C, Stringham H, Jackson A, Narisu N, Chines P, Small K, Kuusisto J, Parks B, Pajukanta P, Kirchgessner T, Collins F, Gargalovic P, Boehnke M, Laakso M, Mohlke K, Lusis A: Genetic regulation of adipose gene expression and integration with GWAS loci and cardio-metabolic traits. submitted
33. Falchi M, Wilson SG, Paximadas D, Swaminathan R, Spector TD: Quantitative Linkage Analysis for Pancreatic B-cell Function and Insulin Resistance in a Large Twin Cohort. Diabetes 2008;57:1120-1124
34. Stancakova A, Javorsky M, Kuulasmaa T, Haffner SM, Kuusisto J, Laakso M: Changes in insulin sensitivity and insulin release in relation to glycemia and glucose tolerance in 6,414 Finnish men. Diabetes 2009;58:1212-1221
35. Cho H, Mu J, Kim JK, Thorvaldsen JL, Chu Q, Crenshaw EB, 3rd, Kaestner KH, Bartolomei MS, Shulman GI, Birnbaum MJ: Insulin resistance and a diabetes mellitus-like syndrome in mice lacking the protein kinase Akt2 (PKB beta). Science 2001;292:1728-1731
36. Garofalo RS, Orena SJ, Rafidi K, Torchia AJ, Stock JL, Hildebrandt AL, Coskran T, Black SC, Brees DJ, Wicks JR, McNeish JD, Coleman KG: Severe diabetes, age-dependent loss of adipose tissue, and mild growth deficiency in mice lacking Akt2/PKB beta. J Clin Invest 2003;112:197-208
37. Hussain K, Challis B, Rocha N, Payne F, Minic M, Thompson A, Daly A, Scott C, Harris J, Smillie BJ, Savage DB, Ramaswami U, De Lonlay P, O'Rahilly S, Barroso I, Semple RK: An activating mutation of AKT2 and human hypoglycemia. Science 2011;334:474
38. Arya VB, Flanagan SE, Schober E, Rami-Merhar B, Ellard S, Hussain K: Activating AKT2 mutation: hypoinsulinemic hypoketotic hypoglycemia. J Clin Endocrinol Metab 2014;99:391-394
39. George S, Rochford JJ, Wolfrum C, Gray SL, Schinner S, Wilson JC, Soos MA, Murgatroyd PR, Williams RM, Acerini CL, Dunger DB, Barford D, Umpleby AM, Wareham NJ, Davies HA, Schafer AJ,

Stoffel M, O'Rahilly S, Barroso I: A family with severe insulin resistance and diabetes due to a mutation in AKT2. Science 2004;304:1325-1328
40. Tan K, Kimber WA, Luan J, Soos MA, Semple RK, Wareham NJ, O'Rahilly S, Barroso I: Analysis of genetic variation in Akt2/PKB-beta in severe insulin resistance, lipodystrophy, type 2 diabetes, and related metabolic phenotypes. Diabetes 2007;56:714-719
41. Parikh C, Janakiraman V, Wu WI, Foo CK, Kljavin NM, Chaudhuri S, Stawiski E, Lee B, Lin J, Li H, Lorenzo MN, Yuan W, Guillory J, Jackson M, Rondon J, Franke Y, Bowman KK, Sagolla M, Stinson J, Wu TD, Wu J, Stokoe D, Stern HM, Brandhuber BJ, Lin K, Skelton NJ, Seshagiri S: Disruption of PH-kinase domain interactions leads to oncogenic activation of AKT in human cancers. Proc Natl Acad Sci U S A 2012;109:19368-19373
42. Lawlor MA, Alessi DR: PKB/Akt: a key mediator of cell proliferation, survival and insulin responses? J Cell Sci 2001;114:2903-2910
43. Zinda MJ, Johnson MA, Paul JD, Horn C, Konicek BW, Lu ZH, Sandusky G, Thomas JE, Neubauer BL, Lai MT, Graff JR: AKT-1, -2, and -3 are expressed in both normal and tumor tissues of the lung, breast, prostate, and colon. Clin Cancer Res 2001;7:2475-2479
44. Peng XD, Xu PZ, Chen ML, Hahn-Windgassen A, Skeen J, Jacobs J, Sundararajan D, Chen WS, Crawford SE, Coleman KG, Hay N: Dwarfism, impaired skin development, skeletal muscle atrophy, delayed bone development, and impeded adipogenesis in mice lacking Akt1 and Akt2. Genes Dev 2003;17:1352-1365
45. O'Rahilly S: Human genetics illuminates the paths to metabolic disease. Nature 2009;462:307-314
46. Lim ET, Wurtz P, Havulinna AS, Palta P, Tukiainen T, Rehnstrom K, Esko T, Magi R, Inouye M, Lappalainen T, Chan Y, Salem RM, Lek M, Flannick J, Sim X, Manning A, Ladenvall C, Bumpstead S, Hamalainen E, Aalto K, Maksimow M, Salmi M, Blankenberg S, Ardissino D, Shah S, Horne B, McPherson R, Hovingh GK, Reilly MP, Watkins H, Goel A, Farrall M, Girelli D, Reiner AP, Stitziel NO, Kathiresan S, Gabriel S, Barrett JC, Lehtimaki T, Laakso M, Groop L, Kaprio J, Perola M, McCarthy MI, Boehnke M, Altshuler DM, Lindgren CM, Hirschhorn JN, Metspalu A, Freimer NB, Zeller T, Jalkanen S, Koskinen S, Raitakari O, Durbin R, MacArthur DG, Salomaa V, Ripatti S, Daly MJ, Palotie A, Sequencing Initiative Suomi P: Distribution and medical impact of loss-of-function variants in the Finnish founder population. PLoS Genet 2014;10:e1004494
47. Dash S, Sano H, Rochford JJ, Semple RK, Yeo G, Hyden CS, Soos MA, Clark J, Rodin A, Langenberg C, Druet C, Fawcett KA, Tung YC, Wareham NJ, Barroso I, Lienhard GE, O'Rahilly S, Savage DB: A truncation mutation in TBC1D4 in a family with acanthosis nigricans and postprandial hyperinsulinemia. Proc Natl Acad Sci U S A 2009;106:9350-9355
48. Bouzakri K, Zachrisson A, Al-Khalili L, Zhang BB, Koistinen HA, Krook A, Zierath JR: siRNA-based gene silencing reveals specialized roles of IRS-1/Akt2 and IRS-2/Akt1 in glucose and lipid metabolism in human skeletal muscle. Cell Metab 2006;4:89-96
49. Baus D, Heermeier K, De Hoop M, Metz-Weidmann C, Gassenhuber J, Dittrich W, Welte S, Tennagels N: Identification of a novel AS160 splice variant that regulates GLUT4 translocation and glucose-uptake in rat muscle cells. Cell Signal 2008;20:2237-2246
50. Cho H, Thorvaldsen JL, Chu Q, Feng F, Birnbaum MJ: Akt1/PKBalpha is required for normal growth but dispensable for maintenance of glucose homeostasis in mice. J Biol Chem 2001;276:38349-38352
51. Toker A, Marmiroli S: Signaling specificity in the Akt pathway in biology and disease. Adv Biol Regul 2014;55:28-38
52. Tschopp O, Yang ZZ, Brodbeck D, Dummler BA, Hemmings-Mieszczak M, Watanabe T, Michaelis T, Frahm J, Hemmings BA: Essential role of protein kinase B gamma (PKB gamma/Akt3) in postnatal brain development but not in glucose homeostasis. Development 2005;132:2943-2954
53. Savage DB, Tan GD, Acerini CL, Jebb SA, Agostini M, Gurnell M, Williams RL, Umpleby AM, Thomas EL, Bell JD, Dixon AK, Dunne F, Boiani R, Cinti S, Vidal-Puig A, Karpe F, Chatterjee VK, O'Rahilly S: Human metabolic syndrome resulting from dominant-negative mutations in the nuclear receptor peroxisome proliferator-activated receptor-gamma. Diabetes 2003;52:910-917

54. Semple RK, Sleigh A, Murgatroyd PR, Adams CA, Bluck L, Jackson S, Vottero A, Kanabar D, Charlton-Menys V, Durrington P, Soos MA, Carpenter TA, Lomas DJ, Cochran EK, Gorden P, O'Rahilly S, Savage DB: Postreceptor insulin resistance contributes to human dyslipidemia and hepatic steatosis. J Clin Invest 2009;119:315-322

Figure Legends

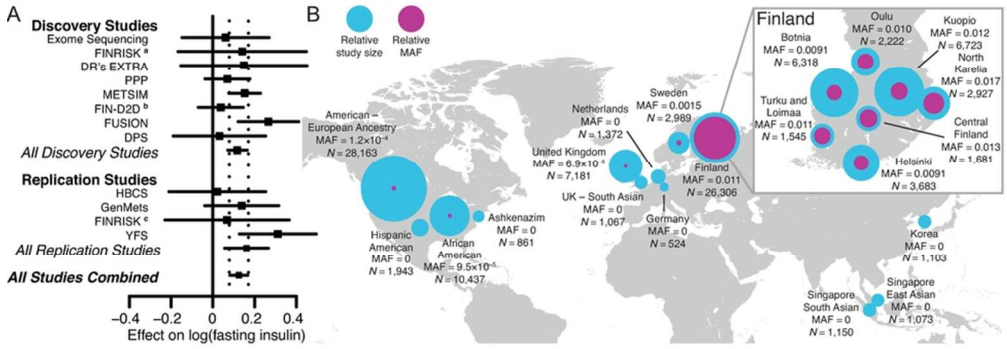
Figure 1. *AKT2* Pro50Thr association with fasting insulin levels. (a) For each study, the square represents the estimate of the additive genetic effect for the association of the *AKT2* Pro50Thr allele with log-transformed fasting insulin (FI) levels and the horizontal line gives the corresponding 95% confidence interval of the estimate. Inverse-variance meta-analyses were performed for *All Discovery Studies*, *All Replication Studies*, and *All Studies Combined*. The vertical dashed lines indicate the 95% confidence interval for the estimate obtained in the meta-analysis of *All Studies Combined*. (b) Minor allele frequency for each available region and ancestry. Across countries the world, the MAF ranges from 0% to 1.1%. The relative sample sizes (N) for each region/ancestry are displayed with the blue circles and the relative minor allele frequencies of *AKT2* Pro50Thr are displayed with the purple circles, with the size of the circles showing comparative differences. Within Finland (inset), where the MAF ranges from 0.9% to 1.7%, birthplace and study center data were used to show the allele distribution across the country. [a] FINRISK 2007; [b] FIN-D2D 2007; [c] FINRISK 1997 and 2002

Figure 2. Expression and conservation properties. (a) Amino acid alignment and conservation of the three AKT proteins in vertebrates. The *x* axis gives the amino acid position and the height of the lines shows the conservation score across 100 vertebrate genome alignments. The functional domains are the pleckstrin homology (PH) domain (blue) and the kinase domain (green). The position of AKT2 Pro50Thr is shown in red while the locations of the other *AKT2* disease-causing mutations (37-40) are shown in orange: Glu17Lys, Arg208Lys, Arg274His, and Arg467Trp. (b) WebLogo plots of amino acids 35-60 are shown for AKT2, AKT1, and AKT3 contrasting the homology of the three isoforms. The height of letters gives the relative frequency of different amino acids across the 100 vertebrate species, with the colors showing amino acids with similar charge. (c) Expression of *AKT1*, *AKT2*, and *AKT3* in eight insulin-sensitive tissues using RNA sequencing data from the GTEx consortium.

Figure 3. Functional properties of AKT2-Thr50 (a) Predicted protein structure of AKT2. Domain and variants are highlighted as in Figure 2. The relative spatial positioning of the AKT2-Pro50 residue is magnified within the inset. (b) HeLa cells were infected with lentiviral V5-AKT2, V5-AKT2-Lys17, V5-AKT2-Thr50, V5-AKT2-Lys208, V5-AKT2-His274, V5-AKT2-Trp467, starved for 18 hours (white bar), and stimulated for 20 minutes with 100nm insulin (grey bar). V5-tagged AKT2 was isolated from cell lysates with anti-V5 agarose beads and incubated with GSK3*β*-GST peptide in an *in vitro* kinase (IVK) assay. Quantification of phosphorylated substrate peptide (pGSK3β) relative to total peptide (GST-GSK3*β*) is shown at the inset. Immunoblots and quantification shown are representative of three independent replicates. Linear model (LM) statistical analyses across all three independent replicates are available in Supplementary Figure S9. The IVK was immunoblotted (IB) with the indicated antibodies. (c) HuH7 cells were infected with lentiviral V5-AKT2, V5-AKT2-Thr50, or control pLX304. At 72

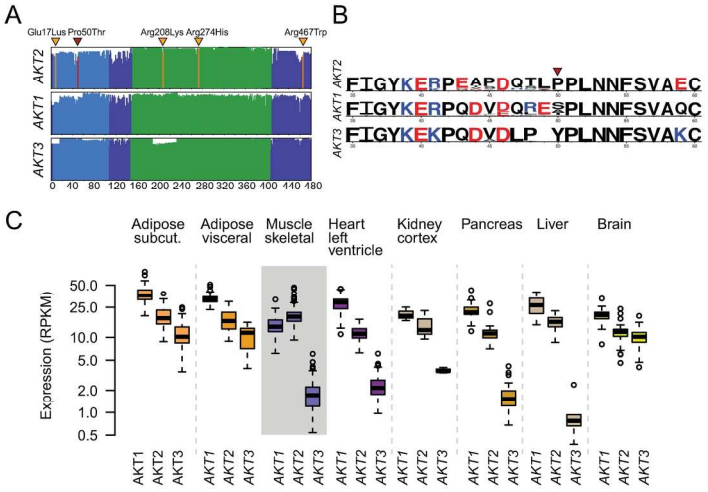
hours relative cellular proliferation was determined with WST-1 assay of HuH7 cells. Error bars represent the standard deviation (SD). *** $P=4.5\times10^{-5}$.

Figure 4. Genetic architecture of rare, low frequency, and common variants associated with FI levels. In this plot, the absolute values of the percent change in fasting insulin level due to rare monogenic mutations (diamonds) and common genetic variants (circles) are plotted against the minor allele frequency of the variant. The extremely rare monogenic mutations (above the dashed line to the left of the *x* axis) were observed in 2 to 18 individuals (3; 37-40; 47; 53; 54) with the height of the point indicating the percent change in fasting insulin levels of mutation carriers from 40 pmol/L, an estimate of population mean fasting insulin level. Mutations in *INSR* and *AKT2* p.Arg274His cause compensatory hyperinsulinemia, individuals with *TBC1D4* p.Arg363Ter show normal fasting insulin levels but postprandial hyperinsulinemia, and mutations in *PTEN* cause enhanced insulin sensitivity providing protection against type 2 diabetes. For common variants, the percent change in fasting insulin levels per insulin-increasing allele is plotted above the solid horizontal axis. These observations are from sequencing (6) and array-based GWAS (3). For several genes, the effects from rare mutations can be compared to the effects of common variants in or near the gene: *PPARG* (blue), *TBC1D4* (green), *PTEN* (orange), and *AKT2* (red). [a] Donohue syndrome: Biallelic loss-of-function mutations in *INSR* (54). [b] Rabson-Mendenhall syndrome: Biallelic loss-of-function mutations in *INSR* (54). [c] Post-pubertal severe IR: Heterozygous or homozygous loss-of-function mutations in *INSR* (54). [d] Loss of function *PTEN* mutations cause Cowden Syndrome in which carriers exhibit a *lowered* fasting insulin level (mean=29 pmol/l) compared to matched controls (3). [e] Carriers with the *AKT2* p.Glu17Lys mutation were described with hypoinsulinemic hypoketotic hypoglycemia and hemihypertrophy with undetectable serum insulin (37; 38).



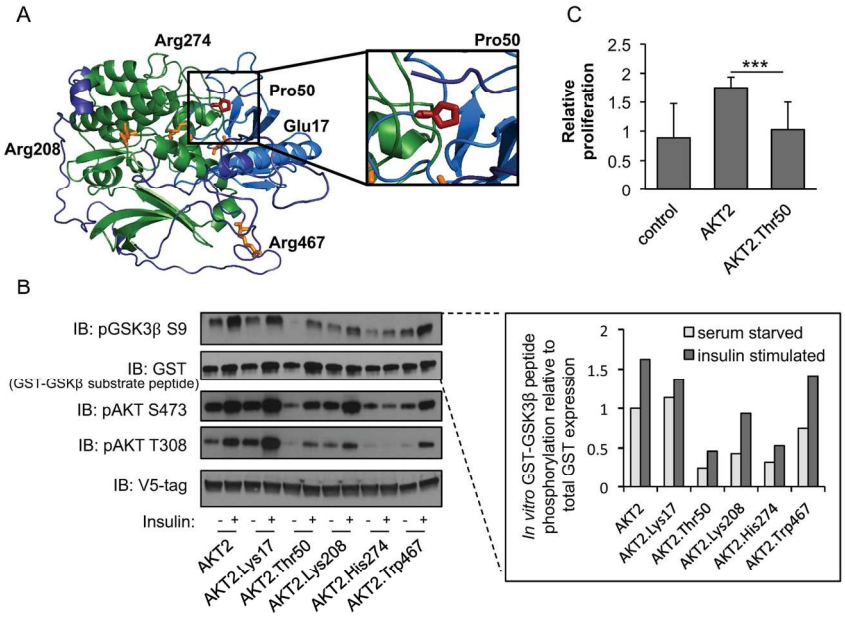
***AKT2* Pro50Thr association with fasting insulin levels.** (A) For each study, the square represents the estimate of the additive genetic effect for the association of the *AKT2* Pro50Thr allele with log-transformed fasting insulin (FI) levels and the horizontal line gives the corresponding 95% confidence interval of the estimate. Inverse-variance meta-analyses were performed for *All Discovery Studies*, *All Replication Studies*, and *All Studies Combined*. The vertical dashed lines indicate the 95% confidence interval for the estimate obtained in the meta-analysis of *All Studies Combined*. (B) Minor allele frequency for each available region and ancestry. Across countries the world, the MAF ranges from 0% to 1.1%. The relative sample sizes (N) for each region/ancestry are displayed with the blue circles and the relative minor allele frequencies of *AKT2* Pro50Thr are displayed with the purple circles, with the size of the circles showing comparative differences. Within Finland (inset), where the MAF ranges from 0.9% to 1.7%, birthplace and study center data were used to show the allele distribution across the country. [a] FINRISK 2007; [b] FIN-D2D 2007; [c] FINRISK 1997 and 2002

Figure 1
82x38mm (300 x 300 DPI)



Expression and conservation properties. (A) Amino-acid alignment and conservation of the three AKT proteins in vertebrates. The *x* axis gives the amino acid position and the height of the lines shows the conservation score across 100 vertebrate genome alignments. The functional domains are the pleckstrin homology (PH) domain (blue) and the kinase domain (green). The position of *AKT2* Pro50Thr is shown in red while the locations of the other *AKT2* disease-causing mutations (34-37) are shown in orange: Glu17Lys, Arg208Lys, Arg274His, and Arg467Trp. (b) WebLogo plots of amino acids 35-60 are shown for AKT2, AKT1, and AKT3 contrasting the homology of the three isoforms. The height of letters gives the relative frequency of different amino acids across the 100 vertebrate species, with the colors showing amino acids with similar charge. (c) Expression of *AKT1*, *AKT2*, and *AKT3* in eight insulin-sensitive tissues using RNA sequencing data from the GTEx consortium.

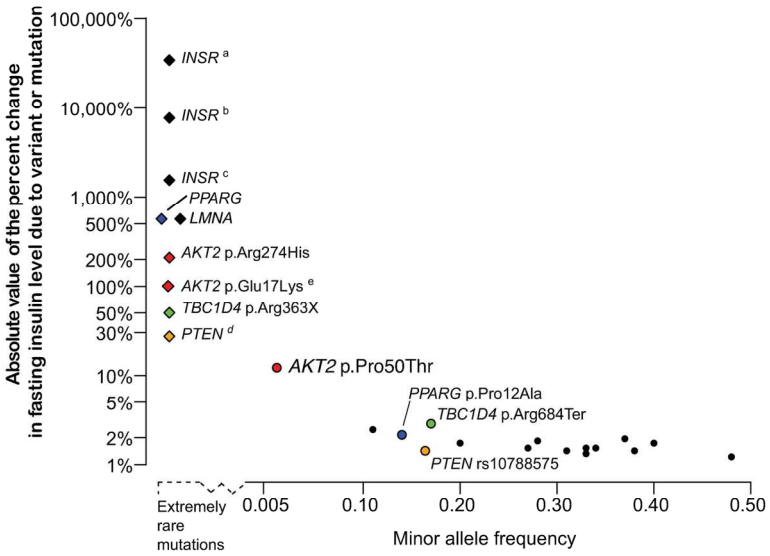
Figure 2
190x142mm (300 x 300 DPI)



Functional properties of AKT2-Thr50 (A) Predicted protein structure of AKT2. Domain and variants are highlighted as in Figure 2. The relative spatial positioning of the AKT2-Pro50 residue is magnified within the inset. (B) HeLa cells were infected with lentiviral V5-AKT2, V5-AKT2-Lys17, V5-AKT2-Thr50, V5-AKT2-Lys208, V5-AKT2-His274, V5-AKT2-Trp467, starved for 18 hours (white bar), and stimulated for 20 minutes with 100nm insulin (grey bar). V5-tagged AKT2 was isolated from cell lysates with anti-V5 agarose beads and incubated with GSK3β-GST peptide in an in vitro kinase (IVK) assay. Quantification of phosphorylated substrate peptide (pGSK3β) relative to total peptide (GST-GSK3β) is shown at the inset. Immunoblots and quantification shown are representative of three independent replicates. Linear model (LM) statistical analyses across all three independent replicates are available in Supplementary Figure 9. The IVK was immunoblotted (IB) with the indicated antibodies. (C) HuH7 cells were infected with lentiviral V5-AKT2, V5-AKT2-Thr50, or control pLX304. At 72 hours relative cellular proliferation was determined with WST-1 assay of HuH7 cells. Error bars represent the standard deviation (SD). *** $P=4.5\times10^{-5}$.

Figure 3

155x118mm (300 x 300 DPI)



Genetic architecture of rare, low frequency, and common variants associated with FI levels. In this plot, the absolute values of the percent change in fasting insulin level due to rare monogenic mutations (diamonds) and common genetic variants (circles) are plotted against the minor allele frequency of the variant. The extremely rare monogenic mutations (above the dashed line to the left of the x axis) were observed in 2 to 18 individuals (3; 34-37; 44; 50; 51) with the height of the point indicating the percent change in fasting insulin levels of mutation carriers from 40 pmol/L, an estimate of population mean fasting insulin level. Mutations in INSR and *AKT2* p.Arg274His cause compensatory hyperinsulinemia, individuals with *TBC1D4* p.Arg363Ter show normal fasting insulin levels but postprandial hyperinsulinemia, and mutations in *PTEN* cause enhanced insulin sensitivity providing protection against type 2 diabetes. For common variants, the percent change in fasting insulin levels per insulin-increasing allele is plotted above the solid horizontal axis. These observations are from sequencing (6) and array-based GWAS (3). For several genes, the effects from rare mutations can be compared to the effects of common variants in or near the gene: *PPARG* (blue), *TBC1D4* (green), PTEN (orange), and *AKT2* (red). [a] Donohue syndrome: Biallelic loss-of-function mutations in INSR (51). [b] Rabson-Mendenhall syndrome: Biallelic loss-of-function mutations in INSR (51). [c] Post-pubertal severe IR: Heterozygous or homozygous loss-of-function mutations in INSR (51). [d] Loss of function *PTEN* mutations cause Cowden Syndrome in which carriers exhibit a lowered fasting insulin level (mean=29 pmol/l) compared to matched controls (3). [e] Carriers with the *AKT2* p.Glu17Lys mutation were described with hypoinsulinemic hypoketotic hypoglycemia and hemihypertrophy with undetectable serum insulin (34; 35).

Figure 4
154x105mm (300 x 300 DPI)

Supplementary Notes

SUPPLEMENTARY NOTE 1: SUMMARY OF ASSOCIATION RESULTS AT KNOWN AND NOVEL LOCI.

The exome-wide single variant association results are displayed in **Supplementary Table 2**. We first partitioned the significant ($P < 5 \times 10^{-7}$) and suggestive ($P < 5 \times 10^{-6}$) single variant association results into two sets: variants in previously reported associated regions (**Supplementary Table 2A**) and variants with potentially novel association signals (**Supplementary Table 2B**).
Of the 57 loci with common variants associated with FG or FI in multiple ancestries (1-13), twenty-one regions contained significant or suggestive association signals in our analysis. Of the seven regions harboring significant associations with non-synonymous variants, five (*GCKR, G6PC2, SLC30A8, PCSK1*, and *GLP1R*) were described previously by our group (13), where, when possible, conditional analyses and functional experiments are utilized to illuminate functional transcripts. In the *MADD* locus, a missense variant *ACP2* p.Arg29Gln showed significant association with FG levels ($P = 1.91 \times 10^{-7}$, MAF = 38%). This variant is in low LD ($r^2 = 0.138$) with the reported variant, rs7944584 ($P = 2.62 \times 10^{-11}$, MAF = 39%), but after conditioning on rs7944584 the association was not significant (P = 0.003). An additional association with a low-frequency variant was observed at the *MTNR1B* locus. A variant upstream of *MTNR1B*, rs7950811, (effect = 0.057; $P = 6.8 \times 10^{-11}$), has a MAF of 4.5% and in low LD with the index SNP, rs10830963 ($r^2 = 0.002$), in 1000 Genomes data (14). After conditioning on the index SNP, the association of rs7950811 with FG remained significant ($P = 3.07 \times 10^{-7}$). For FI, five regions contained significant or suggestive association signals. All of the insulin-associated variants were common with MAF > 25%. Two of these regions, the *GCKR* and *GRB14/COBLL1* loci, harbor significant missense variants and were previously described (13).
Association results at previously reported variants from genome-wide association studies are presented in **Supplementary Table 2C**. Of the 68 previously published common variant associations with FG and FI, we were able to carry out association tests at 36 FG and 16 FI variants. Thirty of the FG association loci showed P < 0.05, with 100 % having a consistent direction of effect. Thirteen FI associated loci had P < 0.05, with 100% demonstrating a consistent direction of effect.

Potentially novel association signals

We observed five and seven variants passing suggestive level of significance for FI and FG, respectively **(Supplementary Table 2B)**. As this analysis focused on coding variation, we took the three coding variants forward to a replication analysis in four independent Finnish studies (N = 5,747) (15-18). The *AKT2* p.Pro50Thr variant in *AKT2* was present and well-imputed in the 1000 Genomes reference panel (imputation score: 0.886 to 0.957). The correlation between imputed and directly genotyped genotypes was high ($r^2 > 0.88$), and the association of this variant with FI levels replicated, ($P_{replication} = 0.00054$, N = 5,747) resulting in a combined (discovery and replication) sample P value of 9.98×10^{-10} (**Supplementary Table 2E**). *MMEL1* p.Glu323Gln, which has a MAF of only 0.2% (seven minor allele carriers in the HBCS subset), was poorly imputed and not tested for association (imputation score: 0.718 to 0.945, $r^2 = 0.57$). *TP53BP1* p.Thr1278Ile was not observed in the studies.

Summary of exome-wide significant gene based association results

The suggestive and significant gene based association signals from each ancestry group in the exome sequencing data and the exome chip data, as well as combined results, are displayed in **Supplementary Table 2D**. The *AKT2* gene based association with FI is described in the main text.
In gene-based tests using the PTV+NS_{broad} mask, *NDUFAF1* was significantly associated with FI levels ($P_{Burden} = 1.10 \times 10^{-6}$). This association was driven by a single missense variant (p.His309Asp, rs199599633, $P = 9.3 \times 10^{-5}$, N = 1,673) that was not associated with FI levels in exome array data (P = 0.018, N = 19,569). NADH dehydrogenase (ubiquinone) complex I, assembly factor 1, or *NDUFAF1*, encodes for a complex I assembly factor protein, which is part of the first step of the respiratory chain. Mutations in both copies of this gene are reported to cause mitochondrial complex I deficiency, which manifests as cardioenphalomypathy or fatal hypertrophic cardiomyopathy while heterozygous parents were reported as healthy(19; 20).
Additionally, a third gene, *GIMAP8*, was associated with FG levels in the PTV-only mask ($P_{Burden} = 2.30 \times 10^{-6}$). This association was driven by singleton and doubleton variants. This gene encodes a GTPase of the immunity-associated protein family (21)

SUPPLEMENTARY NOTE 2: POPULATION GENETICS AND CONSTRAINT

We studied the population genetics properties of *AKT2* and *AKT2* p.Pro50Thr by cataloguing details of all the protein altering variants observed in the T2D-GENES exome sequence data (N=12,940). We phased variants in proteins or genes (including non-coding variants) using SHAPEIT (22) and calculated population statistics and diversity indices with Arlequin (v 3.5) (23), grouped by country of origin. We built the haplotype network using the pegas and igraph libraries in R. dN/dS for Human-Chimpanzee alignments were extracted from ENSEMBL database (24). We computed the "within-human" dN/dS with codeml (PAML) (25) using hg19 sequence as reference and alternative sequence containing all the observed segregating sites. The Mcdonald-Kreitman test (26) for *AKT2* was computed in Bioperl (Bio::PopGen::Statistics) using *AKT3* (hg19) as an outgroup.

There was modest heterogeneity across regions of Finland, with North Karelia (MAF=1.7%) different ($0.001 < \text{pairwise } F_{ST} < 0.003$; $P < 0.01$) from all other tested regions, except Central Finland (MAF=1.3%, pairwise $F_{ST} = 0.0004$, $P = 0.08$). These geographical

differences in Pro50Thr allele frequency are consistent with long-term drift (27) with no evidence of selection pressure differences at *AKT2* across Finland ($dN/dS_{Finland}$=0.1; 0.08<$dN/dS_{European}$<0.4).

In the complete GoT2D and T2D-GENES exome sequence data of 12,940 individuals (6,504 with type 2 diabetes), *AKT2* displayed some evidence of purifying selection (dN/dS<0.01 comparing human and chimpanzee) (**Supplementary Figure S3; Supplementary Figure S4**). We observed 36 non-synonymous variants in *AKT2* (35 with a MAC≤5 and Pro50Thr with MAC=61) (**Supplementary Table 3**). No other protein-altering variants had frequency greater than 0.3% in the 60,706 individuals (including 6,347 from the GoT2D and T2D-GENES studies) in the Exome Aggregation Consortium (ExAC) data.

SUPPLEMENTARY NOTE 3: PATHWAY ANALYSES

We used biological knowledge to test for enrichment of signal in pathways. Pathways and networks were selected from MSigDB (28), which includes Gene Ontology, pathways from KEGG, Ingenuity, Reactome, and Biocarta; and the manually curated monogenic pathways previously considered. We carried out a two-stage enrichment analysis: step one calculates gene aggregation scores using a function of single variant statistics; and step two calculates gene set scores using a function of aggregation scores from each gene in the set. In step one, we make use of a range of gene aggregation functions, including the minimum p-value (or maximum Bayes' factor) for single-variant association (within ancestry or trans-ethnic) in the gene (with correction for the number of variants in the gene). In step two, we apply a pre-ranked GSEA method (28), which consists of a sensitive-improved Kolmogorov-Smirnov (random bridge) statistic, and which provides better correction of the null distribution for highly correlated gene sets (as we see for our hand curated gene sets). Additionally, we performed a biologically enhanced pathway analyses with DEPICT (29), an integrative tool that we used to highlight enriched pathways and identify tissues/cell types where genes from associated loci are highly expressed.

Gene set definitions: We assembled pre-defined, hand-curated lists to create four gene sets: "Monogenic All" (N = 81), including any gene with reported mutations that result in a disease or syndrome leading to either increased prevalence of diabetes or changes in glycemic traits. We further prioritized two subsets of genes, "Monogenic Glucose" (N = 41) and "Monogenic Insulin" (N = 37) including any gene with mutations leading to changes in respective glycemic traits as a primary feature. The list contains genes identified before September 2013. The fourth gene set, "Insulin Receptor Signaling," was created using Ingenuity Pathway Analysis (IPA) tools (30) by merging the insulin receptor signaling, IGF-1 signaling, and PI3K/AKT signaling pathways and adding all downstream phosphylated substrates of AKT.

Association Analysis: SKAT and burden tests were performed after aggregating functional variants (according to the previously described criteria) across all the genes in each gene set. Conditional analyses were performed using features implemented in RareMETALS (31; 32).

Enrichment of association signals: Empirical enrichment for the number of gene based tests with P < 0.001 and the number of single variant tests with P < 0.001 in each gene set was determined by first counting the number of tests below the threshold. For a particular gene set, let $N_{observed}$ denote the number of tests with P < 0.001. A pool of similar genes was assigned to each gene in the gene set, according to the quartile of exon length and quintiles of the number of the nonsynonymous and synonymous variants in the gene. For each gene set, 1,000 matched gene sets were created. An empirical distribution of N_i (the number of tests with P < 0.001 in matched set *i*) was constructed for each of the matched sets. The empirical enrichment P-value was calculated by observing the proportion of matched sets with $N_i \geq N_{observed}$.

Additional traits related to insulin resistance: We examined the single variant association of fasting adiponectin level (log-transformed, age, sex and BMI adjusted, and inverse-normalized), 2 hour glucose level (age, sex and BMI-adjusted, and inverse-normalized) and 2 hour insulin level (log-transformed, age, sex and BMI adjusted, and inverse-normalized) in these pathways using exome array data when available from the discovery cohorts (D2D2007, DPS, DRSEXTRA, FINRISK, FUSION, Health2008, Inter99, METSIM, ULSAM).

Summary of Results

To further assess the evidence of enriched signals in biologically related genes, we looked for enrichment across pathways using both hand curated and publically available pathways. This was conducted using GSEA (28; 33). While no gene-set was significant after multiple testing correction, there is enrichment for several pathways, including adipocytokine signaling, glucose transport, galactose metabolism, glycolysis and gluconeogenesis, and starch and sucrose metabolism pathways, all of which include both *G6PC2* and *G6PC*. While the *G6PC2* association with FG has previously been described (13), we note that *G6PC* mutations result in glycogen storage disorders (34).

Since *AKT2* lies in the insulin receptor signaling pathway and *AKT2* mutations are a known cause of both familial lipodystophy, severe insulin resistance and hypoglycemia (35-38) we next explored whether there was an enrichment of rare and low frequency variants in these gene sets ("Monogenic Genes," and "Insulin Receptor Signaling Genes") [**Supplementary Table 6A**]. First, we tested for global enrichment by aggregating all variants predicted to be deleterious using the annotation masks previously described for gene based testing (PTV-only, PTV+NS_{strict}, PTV+NS_{broad}, PTV+Missense). We found a significant enrichment of deleterious variants (protein truncating, splice site and non-synonymous) in the monogenic genes ($P = 2 \times 10^{-4}$) in exome array data [**Supplementary Table 6B**] but no such enrichment in an analysis of the exome sequencing data set (P = 0.87) [**Supplementary Table 6C**]. Conditional analyses demonstrated that in addition to *AKT2* p.Pro50Thr (P conditional on *AKT2* p.Pro50Thr = 0.0017), seven additional top ranked variants contribute to this signal (P conditional on *AKT2* p.Pro50Thr, *CFTR* p.Asp1270Asn, *INSR* p.Val1012Met, *ZMPSTE24* p.Arg178His, *ZFP57* p.Arg178His, *CFTR* splice donor variant rs78756941 and *PCNT* p.Glu1785Lys jointly = 0.0104) [**Supplementary Table S6D,E**]. No other novel associations were detected with the other gene sets and variant

masks, although when comparing the effects of the burden tests across the four variant aggregation categories, we observed a positive trend of effect as we examined the category containing the least predicted deleterious (PTV+missense) to the most predicted deleterious (PTV-only), although the confidence intervals widen as the number of included variants decrease [**Supplementary Fig. 6**].
To find specific genes harboring an enrichment of association with either FG or FI levels, we next focused on association results from the monogenic genes, testing each set for empirical enrichment. We found that a gene implicated in congenital generalized lipodystrophy, *CAV1* (39), showed enrichment of association with FG levels when considering the set of glucose-specific monogenic genes from the exome sequencing analysis (enrichment P = 0.03; *CAV1* P = 1.9×10^{-4} with protein truncating and low-frequency missense variants and P = 7.0×10^{-4} with protein truncating and predicted deleterious variants). Mutations in *CAV1* are characterized by extreme insulin resistance and lipodystrophy (39) but in our data no association of *CAV1* variants with FI levels was observed. We also observed a borderline enrichment for fasting insulin level with a gene-based burden test in the insulin receptor signaling pathway (enrichment P = 0.06; (*PTGS2* burden P = 1.1×10^{-4} with protein truncating and low-frequency missense variants; [**Supplementary Fig. 7**, **Supplementary Table S7A,B**].
We further examined the association of three quantitative traits related to insulin resistance: fasting adiponectin level, and 2 hour glucose and 2 hour insulin levels after an oral glucose tolerance test. Besides a nominally significance Other than the *AKT2* p.Pro50Thr allele association with 2 hour insulin level (Effect = 26% increase, 95% confidence interval = 16% - 38%, P = 7.86×10^{-8}), no other associations were observed [**Supplementary Fig. 7C**].

SUPPLEMENTARY NOTE 4: EXPRESSION PROFILE OF *AKT2*

GTEx

We compared the expression pattern of *AKT2* to the two other members of the *AKT* gene family, *AKT1* and *AKT3,* using multi-tissue RNA sequencing (RNA-seq) data from the pilot phase of the GTEx project. Detailed procedures for sample collection, RNA extraction, RNA-seq, and gene and transcript quantifications have been previously described (40). Briefly, in the pilot phase, a total of 9,365 tissue samples targeting more than 30 distinct human tissues were collected from 237 post-mortem donors. RNA was extracted, and 1,749 unique samples that passed QC (RIN value of 6.0 or higher and at least 1μg of total RNA), were selected for RNA-seq. Non strand-specific RNA sequencing after poly-A selection was performed using Illumina TruSeq RNA Sample Preparation protocol on the Illumina HiSeq 2000, and aligned with Tophat (v 1.4.1) (41) to UCSC hg19. Gencode (v 12) (42) was used as a transcriptome model for the alignment, and gene and isoform quantifications. Gene and exon level expression was quantified using RNA-SeQC (43) and the Flux Capacitor (v 1.2.3, http://flux.sammeth.net) was used in the quantification of the expression of several transcriptional elements including gene transcript, splice junctions and introns. In total, 44 tissues had data from more than one individual and were used in the analyses.

Genotyping and imputation: Samples were genotyped on the Illumina HumanOmni5-4v1_B SNP array and imputed to the 1,000 Genomes Phase 1 reference (an updated data freeze version from 19 April 2012, release v3) using IMPUTE2 (44; 45) as described (40).

Age and BMI associations: We studied BMI and age associations using a linear mixed model as implemented in the lmer function in the lme4 R package (46). Sex, age, BMI, and three PCs were included in the model as fixed covariates and the date of sequencing and the date of nucleic acid isolation as random covariates. The gene expression RPKM values were inverse variance rank normalized for these analyses.

eQTL analysis: The cis-eQTL for AKT2 in subcutaneous adipose tissue was extracted from the eQTL data generated during the pilot phase of the GTEx project. The methods have been previously described in detail (47). Briefly, the association of common (MAF ≥ 5%) SNPs with gene expression levels was studied using a linear model in MatrixEQTL (48) including sex, three genotyping PCs, and 15 expression PEER factors (49) as covariates. The cis-window was defined as one megabase (Mb) up- and down-stream of the transcription start site of each transcript. Prior to the eQTL analysis the RPKM values were inverse normalized across genes within each tissue and transformed into a standard normal based on rank.

EuroBATs

EuroBATs RNA-seq samples: Samples from photo protected subcutaneous adipose tissue from 766 twins were extracted (131 monozygotic twin pairs, 187 dizygotic twin pairs and 130 unrelated individuals) and processed as previously described (50; 51). In brief, samples were prepared for sequencing with the Illumina TruSeq sample preparation kit (Illumina, San Diego, CA) according to manufacturer's instructions and were sequenced on a HiSeq2000 machine. Afterwards, the 49-bp sequenced paired-end reads were mapped to the GRCh37reference genome (52) with BWA v0.5.9 (53). We use genes defined in the GENCODE 10 annotation (42), removing genes with more than 10% zero read count. RPKM values were root mean transformed.

Genotyping and imputation: Samples were genotyped on a combination of the HumanHap300, HumanHap610Q, 1M-Duo, and 1.2MDuo 1M Illumnia arrays, as described in Grundberg *et. al* (54). Samples were imputed into the 1000 Genomes Phase 1 reference panel (data freeze, 10/11/2010) (6) using IMPUTE2 (44; 45) and filtered (removing variants with MAF<1%, IMPUTE info value<0.8). Samples with both genotypes and expression values (N=720) were used in the subsequent analyses.

Gene-age, gene-BMI, and insulin associations: We used inverse normalized RPKM values to assess the effects of age and BMI on gene expression. We fit linear mixed models using R (55) with the lmer function in the lme4 package (46). Confounding factors in all

models included fixed effects (primer insert size, GC content mean) and random effects (primer index, date of sequencing, family relationship and zygosity). In addition to the adjusting for these fixed and random covariates, the analysis of age also adjusted for BMI and the analysis of BMI was adjusted for age. The P values to assess significance for age and BMI effects were calculated from the Chi-square distribution with 1 degree of freedom using likelihood ratio as the test statistic. FI was measured at the same time point as the fat biopsies, following a previously described protocol (56). Natural log transformed FI were adjusted for age or for age and BMI and the residuals were inverse rank normalized. FI-SNP and FI-*AKT2* association was tested with a linear model using the lm function in R.

eQTL analysis: We ran the eQTL analysis on residuals from a mixed model including the first 20 PCs as fixed effects and family relationship and zygosity as random effects. SNP-expression association was performed with a t-test statistic using the NP-GWAS software. We assessed statistical significance through 100,000 permutations.

METSIM

METSIM RNA samples: Subcutaneous fat biopsy samples were obtained from a sample of the participants of the baseline METSIM study. Total RNA was isolated from these samples using Qiagen miRNeasy Kit according to the manufacturer's instructions. RNA integrity number values were assessed with the Agilent Bioanalyzer 2100. High-quality samples (RNA integrity number>7.0) were used for transcriptional profiling with the Affymetrix Human Genome U219 Array. Genome Studio software (2010.v3) was used to obtain fluorescent intensities.

eQTL analysis and gene-age, gene-BMI and insulin associations: The SNP-gene associations were studied for all SNP within 1 Mb of a given gene. The RNA normalized expression data were adjusted for 35 PEER factors and inverse normal transformed PEER processed residuals were for used eQTL mapping (57). Linear mixed model EMMAX (58) accounts for sample relatedness and was implemented in EPACTS (http://genome.sph.umich.edu/wiki/EPACTS). The sample size for eQTL-mapping was N=770. BMI and age associations, as well as FI associations (with and without adjustment for BMI) were studied using the mixed linear model implemented in lme4 (46) in R. The fixed covariates including age and BMI were used as random covariates. Association between the SNPs associated with *AKT2* expression (eSNPs) and FI was tested with a linear model using the lm() function in R. The natural log transformed FI levels were adjusted for age and BMI and the residuals were inverse rank normalized. All analyses using expression data were conducted in 770 METSIM individuals, while for the tests of eSNP and FI association the sample size for analysis was 10,081.

Expression Profile of AKT2

To gain further insights into the tissues relevant for AKT2 function we explored gene and transcript expression patterns of *AKT2* (ENSG00000105221) from multiple (N = 44) human tissues using RNA sequencing (RNA-seq) data from the Genotype Tissue Expression (GTEx) Project (47).

In the GTEx data *AKT2* is ubiquitously expressed **[Supplementary Fig. 13A,B]**; the gene is present in all the available tissues (median expression across individuals RPKM(59) (reads per kb per million reads) > 7 in all tissues, [**Supplementary Table 8**] and in all individuals, in agreement with previous studies examining *AKT2* expression via RT-PCR, Western blot, and Northern Blot analysis (60-63), and documented essential role of AKT isoforms in biological processes throughout the body (64). No enrichment of *AKT2* expression is present in insulin sensitive tissues (i.e. pancreas, skeletal muscle, adipose tissue (both subcutaneous and visceral), liver and kidney cortex) via RNA sequencing as proposed in mouse and rat models, however, this is consistent with previous examination of *AKT2* mRNA in human tissues (61-63; 65). This GTEx RNA sequencing data does not address insulin-sensitive tissue enrichment seen at the level of AKT2 protein, yet in general mRNA levels correlate with protein abundance (66-68).

AKT2 has multiple alternatively spliced transcripts, yet little is known of their specific roles, and therefore we investigated which of the transcripts are the most abundant and which tissues these are active in Gencode version 12 used in the gene and transcript annotations lists 28 *AKT2* transcripts and 17 of these transcripts are expressed (mean RPKM > 1) in at least one of the studied tissues [**Supplementary Fig. 13C,D**]. However, majority of the expression appears to be due to three *AKT2* transcripts: *AKT2-004* (processed transcript) and *AKT2-001* (protein-coding) that span the full length of the gene, and *AKT2-008* (protein-coding), which does not include the downstream exons. Together these three transcripts constitute on average 44% (range 18-65%) of *AKT2* expression in the GTEx tissues. The two longer *AKT2* transcripts, *AKT2-004* and *AKT2-001*, follow similar expression pattern to the gene, while the shorter one, *AKT2-008*, shows more specific pattern of expression being most expressed in uterus, kidney cortex and esophagus mucosa.

The exon containing the p.Pro50Thr variant is included in 14 out of 28 expressed transcripts (all the 28 *AKT2* transcripts are expressed at a detectable level in at least one individual in at least one tissue), including in all the three most highly expressed transcripts [**Supplementary Fig. 13D**]. The expression profile of the exon containing p.Pro50Thr is similar to the whole *AKT2* gene with the tissues showing highest *AKT2* expression generally having the higher levels of expression of the exon containing p.Pro50Thr [**Supplementary Fig. 13B**]**.** Notably, the exon is expressed in all tissues and all individuals, further suggesting that the exon likely encodes part of the protein integral for its function.

Similarly to *AKT2*, the two other members of the *AKT* gene family, *AKT1* and *AKT3*, are expressed in all the tissues available in the GTEx data with the exception of rather low expression of *AKT3* in liver and whole blood. Of the three genes, *AKT1* is generally the most and *AKT3* the least abundant in all tissues. *AKT2* is the most highly expressed of the three homologs ($P < 0.05$ for all comparisons using one-sided paired Student's t-test and log2 transformed expression values) only in skeletal muscle, pituitary and cerebellum/cerebellar hemisphere, with the higher *AKT2* expression being most pronounced in skeletal muscle [**Supplementary Fig. 14]**.

AKT2 expression in adipose tissue and association with FI

To assess whether Pro50Thr was associated with *AKT2* expression, we tested for gene expression quantitative trait loci (eQTL) in available adipose tissue data. We found an eQTL in the 5'UTR of *AKT2* (rs11880261; MAF=35%) with the common allele associated with lower *AKT2* expression levels (**Supplementary Figure 15**; **Supplementary Table 9**). For Pro50Thr, we found the rare allele was associated with lower *AKT2* expression in adipose tissue (METSIM effect=-1.0 SD; $P=8.9\times10^{-4}$, EAF=0.8%). The rare Pro50Thr coding allele (T) sits on the same haplotype as the common allele of rs11880261 (C, $r^2=0.002$, D'=0.5 in the 1000 Genomes Finnish sample) that is associated with lower *AKT2* expression. A reciprocal conditional analysis showed that these are independent signals (Pro50Thr: $P_{conditional}=8.4\times10^{-3}$; eQTL: $P_{conditional}=1.9\times10^{-13}$). No association was detected between rs11880261 and FI levels (METSIM $P=0.30$, N=10,081; EuroBATS $P=0.80$, N=710), suggesting that the common variant eQTL does not drive the initial FI association.

Mendelian randomization analysis

To elaborate the potential causality behind the association between *AKT2* expression and fasting insulin association, we applied a Mendelian randomization based approach using the discovered eQTL SNPs as instrumental variables (IV) following a similar procedure as described recently (69). The association data for the SNP-gene, gene-FI, and SNP-FI analyses from EuroBATS and METSIM were first combined in a fixed-effects inverse-variance-weighted meta-analysis. We derived the IV estimator by taking the ratio of the regression coefficients from the SNP-FI and SNP-*AKT2* analyses, estimating standard error using the delta method. We used a Z test to determine the significance of the IV estimator and the difference between the IV estimator and the observational estimator. Power for this analysis was calculated using an online MR calculator (http://cnsgenomics.com/shiny/mRnd/) with the following values as input: sample size = 2091, alpha = 0.05, beta_xy =[0.01-0.1], beta_OLS = 0.05, R2_xz = 0.025, sigma_x = sigma_y = 1 (70).

Mendelian randomization with rs11880261 as an instrumental variable for AKT2 expression failed to show a causal relationship between AKT2 expression and FI (P=0.41) (Supplementary Table 10). However, power for the Mendelian randomization analysis is not sufficient to conclude there is no effect. Our instrument (rs11880261) explains about 2.5% of the variance in AKT2, but the observational association between AKT2 expression and FI is also weak. Depending on the estimate of the causal effect of AKT2 expression to FI, the power with the sample size of 2,091 can be as low as 5%.

Supplementary References

1. Prokopenko I, Langenberg C, Florez JC, Saxena R, Soranzo N, Thorleifsson G, Loos RJ, Manning AK, Jackson AU, Aulchenko Y, Potter SC, Erdos MR, Sanna S, Hottenga JJ, Wheeler E, Kaakinen M, Lyssenko V, Chen WM, Ahmadi K, Beckmann JS, Bergman RN, Bochud M, Bonnycastle LL, Buchanan TA, Cao A, Cervino A, Coin L, Collins FS, Crisponi L, de Geus EJ, Dehghan A, Deloukas P, Doney AS, Elliott P, Freimer N, Gateva V, Herder C, Hofman A, Hughes TE, Hunt S, Illig T, Inouye M, Isomaa B, Johnson T, Kong A, Krestyaninova M, Kuusisto J, Laakso M, Lim N, Lindblad U, Lindgren CM, McCann OT, Mohlke KL, Morris AD, Naitza S, Orru M, Palmer CN, Pouta A, Randall J, Rathmann W, Saramies J, Scheet P, Scott LJ, Scuteri A, Sharp S, Sijbrands E, Smit JH, Song K, Steinthorsdottir V, Stringham HM, Tuomi T, Tuomilehto J, Uitterlinden AG, Voight BF, Waterworth D, Wichmann HE, Willemsen G, Witteman JC, Yuan X, Zhao JH, Zeggini E, Schlessinger D, Sandhu M, Boomsma DI, Uda M, Spector TD, Penninx BW, Altshuler D, Vollenweider P, Jarvelin MR, Lakatta E, Waeber G, Fox CS, Peltonen L, Groop LC, Mooser V, Cupples LA, Thorsteinsdottir U, Boehnke M, Barroso I, Van Duijn C, Dupuis J, Watanabe RM, Stefansson K, McCarthy MI, Wareham NJ, Meigs JB, Abecasis GR: Variants in MTNR1B influence fasting glucose levels. Nat Genet 2009;41:77-81

2. Chambers JC, Zhang W, Zabaneh D, Sehmi J, Jain P, McCarthy MI, Froguel P, Ruokonen A, Balding D, Jarvelin MR, Scott J, Elliott P, Kooner JS: Common genetic variation near melatonin receptor MTNR1B contributes to raised plasma glucose and increased risk of type 2 diabetes among Indian Asians and European Caucasians. Diabetes 2009;58:2703-2708

3. Xing C, Cohen JC, Boerwinkle E: A weighted false discovery rate control procedure reveals alleles at FOXA2 that influence fasting glucose levels. Am J Hum Genet 2010;86:440-446

4. Dupuis J, Langenberg C, Prokopenko I, Saxena R, Soranzo N, Jackson AU, Wheeler E, Glazer NL, Bouatia-Naji N, Gloyn AL, Lindgren CM, Magi R, Morris AP, Randall J, Johnson T, Elliott P, Rybin D, Thorleifsson G, Steinthorsdottir V, Henneman P, Grallert H, Dehghan A, Hottenga JJ, Franklin CS, Navarro P, Song K, Goel A, Perry JR, Egan JM, Lajunen T, Grarup N, Sparso T, Doney A, Voight BF, Stringham HM, Li M, Kanoni S, Shrader P, Cavalcanti-Proenca C, Kumari M, Qi L, Timpson NJ, Gieger C, Zabena C, Rocheleau G, Ingelsson E, An P, O'Connell J, Luan J, Elliott A, McCarroll SA, Payne F, Roccasecca RM, Pattou F, Sethupathy P, Ardlie K, Ariyurek Y, Balkau B, Barter P, Beilby JP, Ben-Shlomo Y, Benediktsson R, Bennett AJ, Bergmann S, Bochud M, Boerwinkle E, Bonnefond A, Bonnycastle LL, Borch-Johnsen K, Bottcher Y, Brunner E, Bumpstead SJ, Charpentier G, Chen YD, Chines P, Clarke R, Coin LJ, Cooper MN, Cornelis M, Crawford G, Crisponi L, Day IN, de Geus EJ, Delplanque J, Dina C, Erdos MR, Fedson AC, Fischer-Rosinsky A, Forouhi NG, Fox CS, Frants R, Franzosi MG, Galan P, Goodarzi MO, Graessler J, Groves CJ, Grundy S, Gwilliam R, Gyllensten U, Hadjadj S, Hallmans G, Hammond N, Han X, Hartikainen AL, Hassanali N, Hayward C, Heath SC, Hercberg S, Herder C, Hicks AA, Hillman DR, Hingorani AD, Hofman A, Hui J, Hung J, Isomaa B, Johnson PR, Jorgensen T, Jula A, Kaakinen M, Kaprio J, Kesaniemi YA, Kivimaki M, Knight B, Koskinen S, Kovacs P, Kyvik KO, Lathrop GM, Lawlor DA, Le Bacquer O, Lecoeur C, Li Y, Lyssenko V, Mahley R, Mangino M, Manning AK, Martinez-Larrad MT, McAteer JB, McCulloch LJ, McPherson R, Meisinger C, Melzer D, Meyre D, Mitchell BD, Morken MA, Mukherjee S, Naitza S, Narisu N, Neville MJ, Oostra BA, Orru M, Pakyz R, Palmer CN, Paolisso G, Pattaro C, Pearson D, Peden JF, Pedersen NL, Perola M, Pfeiffer AF, Pichler I, Polasek O, Posthuma D, Potter SC, Pouta A, Province MA, Psaty BM, Rathmann W, Rayner NW, Rice K, Ripatti S, Rivadeneira F, Roden M, Rolandsson O, Sandbaek A, Sandhu M, Sanna S, Sayer AA, Scheet P, Scott LJ, Seedorf U, Sharp SJ, Shields B,

Sigurethsson G, Sijbrands EJ, Silveira A, Simpson L, Singleton A, Smith NL, Sovio U, Swift A, Syddall H, Syvanen AC, Tanaka T, Thorand B, Tichet J, Tonjes A, Tuomi T, Uitterlinden AG, van Dijk KW, van Hoek M, Varma D, Visvikis-Siest S, Vitart V, Vogelzangs N, Waeber G, Wagner PJ, Walley A, Walters GB, Ward KL, Watkins H, Weedon MN, Wild SH, Willemsen G, Witteman JC, Yarnell JW, Zeggini E, Zelenika D, Zethelius B, Zhai G, Zhao JH, Zillikens MC, Consortium D, Consortium G, Global BC, Borecki IB, Loos RJ, Meneton P, Magnusson PK, Nathan DM, Williams GH, Hattersley AT, Silander K, Salomaa V, Smith GD, Bornstein SR, Schwarz P, Spranger J, Karpe F, Shuldiner AR, Cooper C, Dedoussis GV, Serrano-Rios M, Morris AD, Lind L, Palmer LJ, Hu FB, Franks PW, Ebrahim S, Marmot M, Kao WH, Pankow JS, Sampson MJ, Kuusisto J, Laakso M, Hansen T, Pedersen O, Pramstaller PP, Wichmann HE, Illig T, Rudan I, Wright AF, Stumvoll M, Campbell H, Wilson JF, Anders Hamsten on behalf of Procardis C, investigators M, Bergman RN, Buchanan TA, Collins FS, Mohlke KL, Tuomilehto J, Valle TT, Altshuler D, Rotter JI, Siscovick DS, Penninx BW, Boomsma DI, Deloukas P, Spector TD, Frayling TM, Ferrucci L, Kong A, Thorsteinsdottir U, Stefansson K, van Duijn CM, Aulchenko YS, Cao A, Scuteri A, Schlessinger D, Uda M, Ruokonen A, Jarvelin MR, Waterworth DM, Vollenweider P, Peltonen L, Mooser V, Abecasis GR, Wareham NJ, Sladek R, Froguel P, Watanabe RM, Meigs JB, Groop L, Boehnke M, McCarthy MI, Florez JC, Barroso I: New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. Nat Genet 2010;42:105-116

5. Kim YJ, Go MJ, Hu C, Hong CB, Kim YK, Lee JY, Hwang JY, Oh JH, Kim DJ, Kim NH, Kim S, Hong EJ, Kim JH, Min H, Kim Y, Zhang R, Jia W, Okada Y, Takahashi A, Kubo M, Tanaka T, Kamatani N, Matsuda K, consortium M, Park T, Oh B, Kimm K, Kang D, Shin C, Cho NH, Kim HL, Han BG, Lee JY, Cho YS: Large-scale genome-wide association studies in East Asians identify new genetic loci influencing metabolic traits. Nat Genet 2011;43:990-995

6. Manning AK, Hivert MF, Scott RA, Grimsby JL, Bouatia-Naji N, Chen H, Rybin D, Liu CT, Bielak LF, Prokopenko I, Amin N, Barnes D, Cadby G, Hottenga JJ, Ingelsson E, Jackson AU, Johnson T, Kanoni S, Ladenvall C, Lagou V, Lahti J, Lecoeur C, Liu Y, Martinez-Larrad MT, Montasser ME, Navarro P, Perry JR, Rasmussen-Torvik LJ, Salo P, Sattar N, Shungin D, Strawbridge RJ, Tanaka T, van Duijn CM, An P, de Andrade M, Andrews JS, Aspelund T, Atalay M, Aulchenko Y, Balkau B, Bandinelli S, Beckmann JS, Beilby JP, Bellis C, Bergman RN, Blangero J, Boban M, Boehnke M, Boerwinkle E, Bonnycastle LL, Boomsma DI, Borecki IB, Bottcher Y, Bouchard C, Brunner E, Budimir D, Campbell H, Carlson O, Chines PS, Clarke R, Collins FS, Corbaton-Anchuelo A, Couper D, de Faire U, Dedoussis GV, Deloukas P, Dimitriou M, Egan JM, Eiriksdottir G, Erdos MR, Eriksson JG, Eury E, Ferrucci L, Ford I, Forouhi NG, Fox CS, Franzosi MG, Franks PW, Frayling TM, Froguel P, Galan P, de Geus E, Gigante B, Glazer NL, Goel A, Groop L, Gudnason V, Hallmans G, Hamsten A, Hansson O, Harris TB, Hayward C, Heath S, Hercberg S, Hicks AA, Hingorani A, Hofman A, Hui J, Hung J, Jarvelin MR, Jhun MA, Johnson PC, Jukema JW, Jula A, Kao WH, Kaprio J, Kardia SL, Keinanen-Kiukaanniemi S, Kivimaki M, Kolcic I, Kovacs P, Kumari M, Kuusisto J, Kyvik KO, Laakso M, Lakka T, Lannfelt L, Lathrop GM, Launer LJ, Leander K, Li G, Lind L, Lindstrom J, Lobbens S, Loos RJ, Luan J, Lyssenko V, Magi R, Magnusson PK, Marmot M, Meneton P, Mohlke KL, Mooser V, Morken MA, Miljkovic I, Narisu N, O'Connell J, Ong KK, Oostra BA, Palmer LJ, Palotie A, Pankow JS, Peden JF, Pedersen NL, Pehlic M, Peltonen L, Penninx B, Pericic M, Perola M, Perusse L, Peyser PA, Polasek O, Pramstaller PP, Province MA, Raikkonen K, Rauramaa R, Rehnberg E, Rice K, Rotter JI, Rudan I, Ruokonen A, Saaristo T, Sabater-Lleal M, Salomaa V, Savage DB, Saxena R, Schwarz P, Seedorf U, Sennblad B, Serrano-Rios M, Shuldiner AR, Sijbrands EJ, Siscovick DS, Smit JH, Small KS, Smith NL, Smith AV, Stancakova A, Stirrups K, Stumvoll M, Sun YV, Swift AJ, Tonjes A, Tuomilehto J, Trompet S, Uitterlinden AG, Uusitupa M, Vikstrom M, Vitart V, Vohl MC, Voight BF, Vollenweider P, Waeber G, Waterworth DM, Watkins H, Wheeler E, Widen E, Wild SH, Willems SM, Willemsen G, Wilson JF, Witteman JC, Wright AF, Yaghootkar H, Zelenika D, Zemunik T, Zgaga L, Replication DIG, Meta-analysis C, Multiple Tissue Human Expression Resource C, Wareham NJ, McCarthy MI, Barroso I, Watanabe RM, Florez JC, Dupuis J, Meigs JB, Langenberg C: A genome-wide approach accounting for body mass index identifies genetic variants influencing fasting glycemic traits and insulin resistance. Nat Genet 2012;44:659-669

7. Scott RA, Lagou V, Welch RP, Wheeler E, Montasser ME, Luan J, Magi R, Strawbridge RJ, Rehnberg E, Gustafsson S, Kanoni S, Rasmussen-Torvik LJ, Yengo L, Lecoeur C, Shungin D, Sanna S, Sidore C, Johnson PC, Jukema JW, Johnson T, Mahajan A, Verweij N, Thorleifsson G, Hottenga JJ, Shah S, Smith AV, Sennblad B, Gieger C, Salo P, Perola M, Timpson NJ, Evans DM, Pourcain BS, Wu Y, Andrews JS, Hui J, Bielak LF, Zhao W, Horikoshi M, Navarro P, Isaacs A, O'Connell JR, Stirrups K, Vitart V, Hayward C, Esko T, Mihailov E, Fraser RM, Fall T, Voight BF, Raychaudhuri S, Chen H, Lindgren CM, Morris AP, Rayner NW, Robertson N, Rybin D, Liu CT, Beckmann JS, Willems SM, Chines PS, Jackson AU, Kang HM, Stringham HM, Song K, Tanaka T, Peden JF, Goel A, Hicks AA, An P, Muller-Nurasyid M, Franco-Cereceda A, Folkersen L, Marullo L, Jansen H, Oldehinkel AJ, Bruinenberg M, Pankow JS, North KE, Forouhi NG, Loos RJ, Edkins S, Varga TV, Hallmans G, Oksa H, Antonella M, Nagaraja R, Trompet S, Ford I, Bakker SJ, Kong A, Kumari M, Gigante B, Herder C, Munroe PB, Caulfield M, Antti J, Mangino M, Small K, Miljkovic I, Liu Y, Atalay M, Kiess W, James AL, Rivadeneira F, Uitterlinden AG, Palmer CN, Doney AS, Willemsen G, Smit JH, Campbell S, Polasek O, Bonnycastle LL, Hercberg S, Dimitriou M, Bolton JL, Fowkes GR, Kovacs P, Lindstrom J, Zemunik T, Bandinelli S, Wild SH, Basart HV, Rathmann W, Grallert H, Replication DIG, Meta-analysis C, Maerz W, Kleber ME, Boehm BO, Peters A, Pramstaller PP, Province MA, Borecki IB, Hastie ND, Rudan I, Campbell H, Watkins H, Farrall M, Stumvoll M, Ferrucci L, Waterworth DM, Bergman RN, Collins FS, Tuomilehto J, Watanabe RM, de Geus EJ, Penninx BW, Hofman A, Oostra BA, Psaty BM, Vollenweider P, Wilson JF, Wright AF, Hovingh GK, Metspalu A, Uusitupa M, Magnusson PK, Kyvik KO, Kaprio J, Price JF, Dedoussis GV, Deloukas P, Meneton P, Lind L, Boehnke M, Shuldiner AR, van Duijn CM, Morris AD, Toenjes A, Peyser PA, Beilby JP, Korner A, Kuusisto J, Laakso M, Bornstein SR, Schwarz PE, Lakka TA, Rauramaa R, Adair LS, Smith GD, Spector TD, Illig T, de Faire U, Hamsten A, Gudnason V, Kivimaki M, Hingorani A, Keinanen-Kiukaanniemi SM, Saaristo TE, Boomsma DI, Stefansson K, van der Harst P, Dupuis J, Pedersen NL, Sattar N, Harris TB, Cucca F, Ripatti S, Salomaa V, Mohlke KL, Balkau B, Froguel P, Pouta A, Jarvelin MR, Wareham NJ, Bouatia-Naji N, McCarthy MI, Franks PW, Meigs JB, Teslovich TM, Florez JC, Langenberg C, Ingelsson

E, Prokopenko I, Barroso I: Large-scale association analyses identify new loci influencing glycemic traits and provide insight into the underlying biological pathways. Nat Genet 2012;44:991-1005
8. Chen G, Bentley A, Adeyemo A, Shriner D, Zhou J, Doumatey A, Huang H, Ramos E, Erdos M, Gerry N, Herbert A, Christman M, Rotimi C: Genome-wide association study identifies novel loci association with fasting insulin and insulin resistance in African Americans. Hum Mol Genet 2012;21:4530-4536
9. Comuzzie AG, Cole SA, Laston SL, Voruganti VS, Haack K, Gibbs RA, Butte NF: Novel genetic loci identified for the pathophysiology of childhood obesity in the Hispanic population. PLoS One 2012;7:e51954
10. Kristiansson K, Perola M, Tikkanen E, Kettunen J, Surakka I, Havulinna AS, Stancakova A, Barnes C, Widen E, Kajantie E, Eriksson JG, Viikari J, Kahonen M, Lehtimaki T, Raitakari OT, Hartikainen AL, Ruokonen A, Pouta A, Jula A, Kangas AJ, Soininen P, Ala-Korpela M, Mannisto S, Jousilahti P, Bonnycastle LL, Jarvelin MR, Kuusisto J, Collins FS, Laakso M, Hurles ME, Palotie A, Peltonen L, Ripatti S, Salomaa V: Genome-wide screen for metabolic syndrome susceptibility loci reveals strong lipid gene contribution but no evidence for common genetic basis for clustering of metabolic syndrome traits. Circ Cardiovasc Genet 2012;5:242-249
11. Go MJ, Hwang JY, Kim YJ, Hee Oh J, Kim YJ, Heon Kwak S, Soo Park K, Lee J, Kim BJ, Han BG, Cho MC, Cho YS, Lee JY: New susceptibility loci in MYL2, C12orf51 and OAS1 associated with 1-h plasma glucose as predisposing risk factors for type 2 diabetes in the Korean population. J Hum Genet 2013;58:362-365
12. Wessel J, Chu AY, Willems SM, Wang S, Yaghootkar H, Brody JA, Dauriz M, Hivert MF, Raghavan S, Lipovich L, Hidalgo B, Fox K, Huffman JE, An P, Lu Y, Rasmussen-Torvik LJ, Grarup N, Ehm MG, Li L, Baldridge AS, Stancakova A, Abrol R, Besse C, Boland A, Bork-Jensen J, Fornage M, Freitag DF, Garcia ME, Guo X, Hara K, Isaacs A, Jakobsdottir J, Lange LA, Layton JC, Li M, Hua Zhao J, Meidtner K, Morrison AC, Nalls MA, Peters MJ, Sabater-Lleal M, Schurmann C, Silveira A, Smith AV, Southam L, Stoiber MH, Strawbridge RJ, Taylor KD, Varga TV, Allin KH, Amin N, Aponte JL, Aung T, Barbieri C, Bihlmeyer NA, Boehnke M, Bombieri C, Bowden DW, Burns SM, Chen Y, Chen YD, Cheng CY, Correa A, Czajkowski J, Dehghan A, Ehret GB, Eiriksdottir G, Escher SA, Farmaki AE, Franberg M, Gambaro G, Giulianini F, Goddard WA, 3rd, Goel A, Gottesman O, Grove ML, Gustafsson S, Hai Y, Hallmans G, Heo J, Hoffmann P, Ikram MK, Jensen RA, Jorgensen ME, Jorgensen T, Karaleftheri M, Khor CC, Kirkpatrick A, Kraja AT, Kuusisto J, Lange EM, Lee IT, Lee WJ, Leong A, Liao J, Liu C, Liu Y, Lindgren CM, Linneberg A, Malerba G, Mamakou V, Marouli E, Maruthur NM, Matchan A, McKean-Cowdin R, McLeod O, Metcalf GA, Mohlke KL, Muzny DM, Ntalla I, Palmer ND, Pasko D, Peter A, Rayner NW, Renstrom F, Rice K, Sala CF, Sennblad B, Serafetinidis I, Smith JA, Soranzo N, Speliotes EK, Stahl EA, Stirrups K, Tentolouris N, Thanopoulou A, Torres M, Traglia M, Tsafantakis E, Javad S, Yanek LR, Zengini E, Becker DM, Bis JC, Brown JB, Cupples LA, Hansen T, Ingelsson E, Karter AJ, Lorenzo C, Mathias RA, Norris JM, Peloso GM, Sheu WH, Toniolo D, Vaidya D, Varma R, Wagenknecht LE, Boeing H, Bottinger EP, Dedoussis G, Deloukas P, Ferrannini E, Franco OH, Franks PW, Gibbs RA, Gudnason V, Hamsten A, Harris TB, Hattersley AT, Hayward C, Hofman A, Jansson JH, Langenberg C, Launer LJ, Levy D, Oostra BA, O'Donnell CJ, O'Rahilly S, Padmanabhan S, Pankow JS, Polasek O, Province MA, Rich SS, Ridker PM, Rudan I, Schulze MB, Smith BH, Uitterlinden AG, Walker M, Watkins H, Wong TY, Zeggini E, Consortium EP-I, Laakso M, Borecki IB, Chasman DI, Pedersen O, Psaty BM, Tai ES, van Duijn CM, Wareham NJ, Waterworth DM, Boerwinkle E, Kao WH, Florez JC, Loos RJ, Wilson JG, Frayling TM, Siscovick DS, Dupuis J, Rotter JI, Meigs JB, Scott RA, Goodarzi MO: Low-frequency and rare exome chip variants associate with fasting glucose and type 2 diabetes susceptibility. Nat Commun 2015;6:5897
13. Mahajan A, Sim X, Ng HJ, Manning A, Rivas MA, Highland HM, Locke AE, Grarup N, Im HK, Cingolani P, Flannick J, Fontanillas P, Fuchsberger C, Gaulton KJ, Teslovich TM, Rayner NW, Robertson NR, Beer NL, Rundle JK, Bork-Jensen J, Ladenvall C, Blancher C, Buck D, Buck G, Burtt NP, Gabriel S, Gjesing AP, Groves CJ, Hollensted M, Huyghe JR, Jackson AU, Jun G, Justesen JM, Mangino M, Murphy J, Neville M, Onofrio R, Small KS, Stringham HM, Syvanen AC, Trakalo J, Abecasis G, Bell GI, Blangero J, Cox NJ, Duggirala R, Hanis CL, Seielstad M, Wilson JG, Christensen C, Brandslund I, Rauramaa R, Surdulescu GL, Doney AS, Lannfelt L, Linneberg A, Isomaa B, Tuomi T, Jorgensen ME, Jorgensen T, Kuusisto J, Uusitupa M, Salomaa V, Spector TD, Morris AD, Palmer CN, Collins FS, Mohlke KL, Bergman RN, Ingelsson E, Lind L, Tuomilehto J, Hansen T, Watanabe RM, Prokopenko I, Dupuis J, Karpe F, Groop L, Laakso M, Pedersen O, Florez JC, Morris AP, Altshuler D, Meigs JB, Boehnke M, McCarthy MI, Lindgren CM, Gloyn AL, consortium TDG, Go TDc: Identification and functional characterization of G6PC2 coding variants influencing glycemic traits define an effector transcript at the G6PC2-ABCB11 locus. PLoS Genet 2015;11:e1004876
14. Johnson AD, Handsaker RE, Pulit SL, Nizzari MM, O'Donnell CJ, de Bakker PI: SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. Bioinformatics 2008;24:2938-2939
15. Eriksson JG: Epidemiology, genes and the environment: lessons learned from the Helsinki Birth Cohort Study. J Intern Med 2007;261:418-425
16. Perttila J, Merikanto K, Naukkarinen J, Surakka I, Martin NW, Tanhuanpaa K, Grimard V, Taskinen MR, Thiele C, Salomaa V, Jula A, Perola M, Virtanen I, Peltonen L, Olkkonen VM: OSBPL10, a novel candidate gene for high triglyceride trait in dyslipidemic Finnish subjects, regulates cellular lipid metabolism. J Mol Med (Berl) 2009;87:825-835
17. Raitakari OT, Juonala M, Ronnemaa T, Keltikangas-Jarvinen L, Rasanen L, Pietikainen M, Hutri-Kahonen N, Taittonen L, Jokinen E, Marniemi J, Jula A, Telama R, Kahonen M, Lehtimaki T, Akerblom HK, Viikari JS: Cohort profile: the cardiovascular risk in Young Finns Study. Int J Epidemiol 2008;37:1220-1226
18. Vartiainen E, Laatikainen T, Peltonen M, Juolevi A, Mannisto S, Sundvall J, Jousilahti P, Salomaa V, Valsta L, Puska P: Thirty-five-year trends in cardiovascular risk factors in Finland. Int J Epidemiol 2010;39:504-518
19. Dunning CJ, McKenzie M, Sugiana C, Lazarou M, Silke J, Connelly A, Fletcher JM, Kirby DM, Thorburn DR, Ryan MT: Human CIA30 is involved in the early assembly of mitochondrial complex I and mutations in its gene cause disease. EMBO J 2007;26:3227-3237

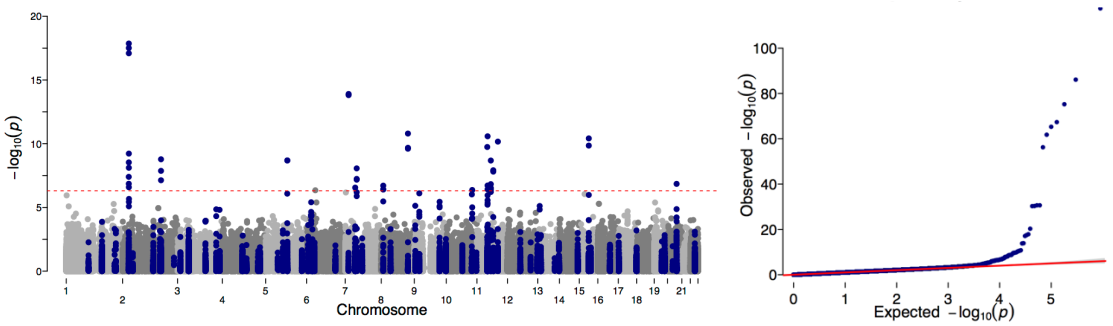
20. Fassone E, Taanman JW, Hargreaves IP, Sebire NJ, Cleary MA, Burch M, Rahman S: Mutations in the mitochondrial complex I assembly factor NDUFAF1 cause fatal infantile hypertrophic cardiomyopathy. J Med Genet 2011;48:691-697
21. Krucken J, Schroetel RM, Muller IU, Saidani N, Marinovski P, Benten WP, Stamm O, Wunderlich F: Comparative analysis of the human gimap gene cluster encoding a novel GTPase family. Gene 2004;341:291-304
22. Delaneau O, Marchini J, Zagury JF: A linear complexity phasing method for thousands of genomes. Nat Methods 2012;9:179-181
23. Excoffier L, Lischer HE: Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Resour 2010;10:564-567
24. Vilella AJ, Severin J, Ureta-Vidal A, Heng L, Durbin R, Birney E: EnsemblCompara GeneTrees: Complete, duplication-aware phylogenetic trees in vertebrates. Genome Res 2009;19:327-335
25. Yang Z: PAML 4: phylogenetic analysis by maximum likelihood. Mol Biol Evol 2007;24:1586-1591
26. McDonald JH, Kreitman M: Adaptive protein evolution at the Adh locus in Drosophila. Nature 1991;351:652-654
27. Palo JU, Ulmanen I, Lukka M, Ellonen P, Sajantila A: Genetic markers and population history: Finland revisited. Eur J Hum Genet 2009;17:1336-1346
28. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP: Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 2005;102:15545-15550
29. Pers TH, Karjalainen JM, Chan Y, Westra HJ, Wood AR, Yang J, Lui JC, Vedantam S, Gustafsson S, Esko T, Frayling T, Speliotes EK, Genetic Investigation of ATC, Boehnke M, Raychaudhuri S, Fehrmann RS, Hirschhorn JN, Franke L: Biological interpretation of genome-wide association studies using predicted gene functions. Nat Commun 2015;6:5890
30. Storey JD, Tibshirani R: Statistical significance for genomewide studies. Proc Natl Acad Sci U S A 2003;100:9440-9445
31. Feng S, Liu D, Zhan X, Wing MK, Abecasis GR: RAREMETAL: fast and powerful meta-analysis for rare variants. Bioinformatics 2014;30:2828-2829
32. Liu DJ, Peloso GM, Zhan X, Holmen OL, Zawistowski M, Feng S, Nikpay M, Auer PL, Goel A, Zhang H, Peters U, Farrall M, Orho-Melander M, Kooperberg C, McPherson R, Watkins H, Willer CJ, Hveem K, Melander O, Kathiresan S, Abecasis GR: Meta-analysis of gene-level tests for rare variant association. Nat Genet 2014;46:200-204
33. Mootha VK, Lindgren CM, Eriksson KF, Subramanian A, Sihag S, Lehar J, Puigserver P, Carlsson E, Ridderstrale M, Laurila E, Houstis N, Daly MJ, Patterson N, Mesirov JP, Golub TR, Tamayo P, Spiegelman B, Lander ES, Hirschhorn JN, Altshuler D, Groop LC: PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. Nat Genet 2003;34:267-273
34. Lei KJ, Shelly LL, Pan CJ, Sidbury JB, Chou JY: Mutations in the glucose-6-phosphatase gene that cause glycogen storage disease type 1a. Science 1993;262:580-583
35. Arya VB, Flanagan SE, Schober E, Rami-Merhar B, Ellard S, Hussain K: Activating AKT2 mutation: hypoinsulinemic hypoketotic hypoglycemia. J Clin Endocrinol Metab 2014;99:391-394
36. George S, Rochford JJ, Wolfrum C, Gray SL, Schinner S, Wilson JC, Soos MA, Murgatroyd PR, Williams RM, Acerini CL, Dunger DB, Barford D, Umpleby AM, Wareham NJ, Davies HA, Schafer AJ, Stoffel M, O'Rahilly S, Barroso I: A family with severe insulin resistance and diabetes due to a mutation in AKT2. Science 2004;304:1325-1328
37. Hussain K, Challis B, Rocha N, Payne F, Minic M, Thompson A, Daly A, Scott C, Harris J, Smillie BJ, Savage DB, Ramaswami U, De Lonlay P, O'Rahilly S, Barroso I, Semple RK: An activating mutation of AKT2 and human hypoglycemia. Science 2011;334:474
38. Tan K, Kimber WA, Luan J, Soos MA, Semple RK, Wareham NJ, O'Rahilly S, Barroso I: Analysis of genetic variation in Akt2/PKB-beta in severe insulin resistance, lipodystrophy, type 2 diabetes, and related metabolic phenotypes. Diabetes 2007;56:714-719
39. Cao H, Alston L, Ruschman J, Hegele RA: Heterozygous CAV1 frameshift mutations (MIM 601047) in patients with atypical partial lipodystrophy and hypertriglyceridemia. Lipids Health Dis 2008;7:3
40. GTEx-Consortium: Multi-tissue transcriptome analysis in a population sample: the Genotype-Tissue Expression (GTEx) pilot study. submitted 2015;
41. Trapnell C, Pachter L, Salzberg SL: TopHat: discovering splice junctions with RNA-Seq. Bioinformatics 2009;25:1105-1111
42. Harrow J, Frankish A, Gonzalez JM, Tapanari E, Diekhans M, Kokocinski F, Aken BL, Barrell D, Zadissa A, Searle S, Barnes I, Bignell A, Boychenko V, Hunt T, Kay M, Mukherjee G, Rajan J, Despacio-Reyes G, Saunders G, Steward C, Harte R, Lin M, Howald C, Tanzer A, Derrien T, Chrast J, Walters N, Balasubramanian S, Pei B, Tress M, Rodriguez JM, Ezkurdia I, van Baren J, Brent M, Haussler D, Kellis M, Valencia A, Reymond A, Gerstein M, Guigo R, Hubbard TJ: GENCODE: the reference human genome annotation for The ENCODE Project. Genome Res 2012;22:1760-1774
43. DeLuca DS, Levin JZ, Sivachenko A, Fennell T, Nazaire MD, Williams C, Reich M, Winckler W, Getz G: RNA-SeQC: RNA-seq metrics for quality control and process optimization. Bioinformatics 2012;28:1530-1532
44. Howie BN, Donnelly P, Marchini J: A Flexible and Accurate Genotype Imputation Method for the Next Generation of Genome-Wide Association Studies. PLoS Genet 2009;5:e1000529
45. Marchini J, Howie B, Myers S, McVean G, Donnelly P: A new multipoint method for genome-wide association studies by imputation of genotypes. Nat Genet 2007;39:906-913
46. Bates D, Maechler M, Bolker B: lme4: Linear mixed-effects models using S4 classes. R package version 0.999375-41. 2011;
47. Consortium GT: The Genotype-Tissue Expression (GTEx) project. Nat Genet 2013;45:580-585
48. Shabalin AA: Matrix eQTL: ultra fast eQTL analysis via large matrix operations. Bioinformatics 2012;28:1353-1358

49. Stegle O, Parts L, Piipari M, Winn J, Durbin R: Using probabilistic estimation of expression residuals (PEER) to obtain increased power and interpretability of gene expression analyses. Nat Protoc 2012;7:500-507
50. Buil A, Brown AA, Lappalainen T, Vinuela A, Davies MN, Zheng H-F, Richards JB, Glass D, Small KS, Durbin R, Spector TD, Dermitzakis ET: Gene-gene and gene-environment interactions detected by transcriptome sequence analysis in twins. Nat Genet 2015;47:88-91
51. Brown AA, Buil A, Viñuela A, Lappalainen T, Zheng H-F, Richards JB, Small KS, Spector TD, Dermitzakis ET, Durbin R: *Genetic interactions affecting human gene expression identified by variance association mapping*. Khaitovich P, Ed., 2014
52. The International Human Genome Sequencing Consortium: Initial sequencing and analysis of the human genome. Nature 2001;409:860-921
53. Li H, Durbin R: Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 2009;25:1754-1760
54. Grundberg E, Small KS, Hedman AK, Nica AC, Buil A, Keildson S, Bell JT, Yang T-P, Meduri E, Barrett A, Nisbett J, Sekowska M, Wilk A, Shin S-Y, Glass D, Travers M, Min JL, Ring S, Ho K, Thorleifsson G, Kong A, Thorsteindottir U, Ainali C, Dimas AS, Hassanali N, Ingle C, Knowles D, Krestyaninova M, Lowe CE, Di Meglio P, Montgomery SB, Parts L, Potter S, Surdulescu G, Tsaprouni L, Tsoka S, Bataille V, Durbin R, Nestle FO, O'Rahilly S, Soranzo N, Lindgren CM, Zondervan KT, Ahmadi KR, Schadt EE, Stefansson K, Smith GD, McCarthy MI, Deloukas P, Dermitzakis ET, Spector TD: Mapping cis- and trans-regulatory effects across multiple tissues in twins. Nat Genet 2012;44:1084-1089
55. The R Project for Statistical Computing [article online], Available from http://www.r-project.org/.
56. Falchi M, Wilson SG, Paximadas D, Swaminathan R, Spector TD: Quantitative Linkage Analysis for Pancreatic B-cell Function and Insulin Resistance in a Large Twin Cohort. Diabetes 2008;57:1120-1124
57. Stegle O, Parts L, Durbin R, Winn J: A Bayesian framework to account for complex non-genetic factors in gene expression levels greatly increases power in eQTL studies. PLoS Comput Biol 2010;6:e1000770
58. Kang HM, Sul JH, Service SK, Zaitlen NA, Kong SY, Freimer NB, Sabatti C, Eskin E: Variance component model to account for sample structure in genome-wide association studies. Nat Genet 2010;42:348-354
59. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B: Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Methods 2008;5:621-628
60. Jones PF, Jakubowicz T, Hemmings BA: Molecular cloning of a second form of rac protein kinase. Cell Regul 1991;2:1001-1009
61. Konishi H, Shinomura T, Kuroda S, Ono Y, Kikkawa U: Molecular cloning of rat RAC protein kinase alpha and beta and their association with protein kinase C zeta. Biochem Biophys Res Commun 1994;205:817-825
62. Yang ZZ, Tschopp O, Di-Poi N, Bruder E, Baudry A, Dummler B, Wahli W, Hemmings BA: Dosage-dependent effects of Akt1/protein kinase Balpha (PKBalpha) and Akt3/PKBgamma on thymus, skin, and cardiovascular and nervous system development in mice. Mol Cell Biol 2005;25:10407-10418
63. Zinda MJ, Johnson MA, Paul JD, Horn C, Konicek BW, Lu ZH, Sandusky G, Thomas JE, Neubauer BL, Lai MT, Graff JR: AKT-1, -2, and -3 are expressed in both normal and tumor tissues of the lung, breast, prostate, and colon. Clin Cancer Res 2001;7:2475-2479
64. Peng XD, Xu PZ, Chen ML, Hahn-Windgassen A, Skeen J, Jacobs J, Sundararajan D, Chen WS, Crawford SE, Coleman KG, Hay N: Dwarfism, impaired skin development, skeletal muscle atrophy, delayed bone development, and impeded adipogenesis in mice lacking Akt1 and Akt2. Genes Dev 2003;17:1352-1365
65. Altomare DA, Lyons GE, Mitsuuchi Y, Cheng JQ, Testa JR: Akt2 mRNA is highly expressed in embryonic brown fat and the AKT2 kinase is activated by insulin. Oncogene 1998;16:2407-2411
66. Greenbaum D, Colangelo C, Williams K, Gerstein M: Comparing protein abundance and mRNA expression levels on a genomic scale. Genome Biol 2003;4:117
67. Ning K, Fermin D, Nesvizhskii AI: Comparative analysis of different label-free mass spectrometry based protein abundance estimates and their correlation with RNA-Seq gene expression data. J Proteome Res 2012;11:2261-2271
68. Schwanhausser B, Busse D, Li N, Dittmar G, Schuchhardt J, Wolf J, Chen W, Selbach M: Global quantification of mammalian gene expression control. Nature 2011;473:337-342
69. Fall T, Hagg S, Magi R, Ploner A, Fischer K, Horikoshi M, Sarin AP, Thorleifsson G, Ladenvall C, Kals M, Kuningas M, Draisma HH, Ried JS, van Zuydam NR, Huikari V, Mangino M, Sonestedt E, Benyamin B, Nelson CP, Rivera NV, Kristiansson K, Shen HY, Havulinna AS, Dehghan A, Donnelly LA, Kaakinen M, Nuotio ML, Robertson N, de Bruijn RF, Ikram MA, Amin N, Balmforth AJ, Braund PS, Doney AS, Doring A, Elliott P, Esko T, Franco OH, Gretarsdottir S, Hartikainen AL, Heikkila K, Herzig KH, Holm H, Hottenga JJ, Hypponen E, Illig T, Isaacs A, Isomaa B, Karssen LC, Kettunen J, Koenig W, Kuulasmaa K, Laatikainen T, Laitinen J, Lindgren C, Lyssenko V, Laara E, Rayner NW, Mannisto S, Pouta A, Rathmann W, Rivadeneira F, Ruokonen A, Savolainen MJ, Sijbrands EJ, Small KS, Smit JH, Steinthorsdottir V, Syvanen AC, Taanila A, Tobin MD, Uitterlinden AG, Willems SM, Willemsen G, Witteman J, Perola M, Evans A, Ferrieres J, Virtamo J, Kee F, Tregouet DA, Arveiler D, Amouyel P, Ferrario MM, Brambilla P, Hall AS, Heath AC, Madden PA, Martin NG, Montgomery GW, Whitfield JB, Jula A, Knekt P, Oostra B, van Duijn CM, Penninx BW, Davey Smith G, Kaprio J, Samani NJ, Gieger C, Peters A, Wichmann HE, Boomsma DI, de Geus EJ, Tuomi T, Power C, Hammond CJ, Spector TD, Lind L, Orho-Melander M, Palmer CN, Morris AD, Groop L, Jarvelin MR, Salomaa V, Vartiainen E, Hofman A, Ripatti S, Metspalu A, Thorsteinsdottir U, Stefansson K, Pedersen NL, McCarthy MI, Ingelsson E, Prokopenko I, European Network for G, Genomic Epidemiology c: The role of adiposity in cardiometabolic traits: a Mendelian randomization analysis. PLoS Med 2013;10:e1001474
70. Brion MJ, Shakhbazov K, Visscher PM: Calculating statistical power in Mendelian randomization studies. Int J Epidemiol 2013;42:1497-1501

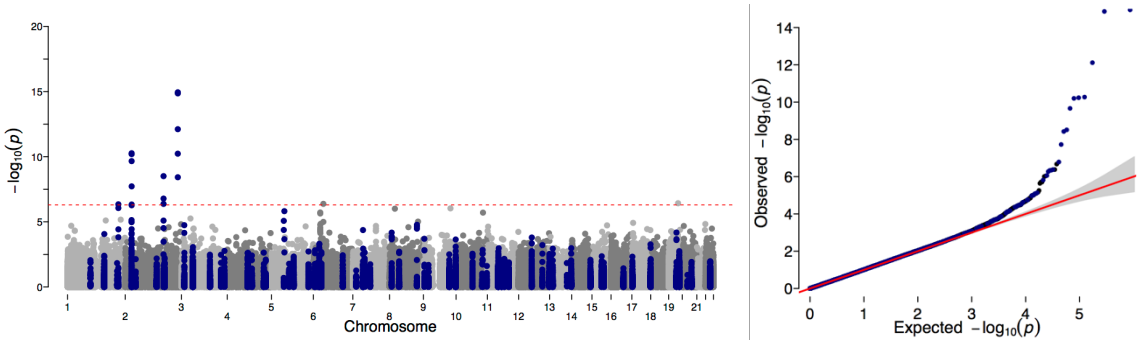
Supplementary Figures

SUPPLEMENTARY FIGURE S1

A. Fasting Plasma Glucose *

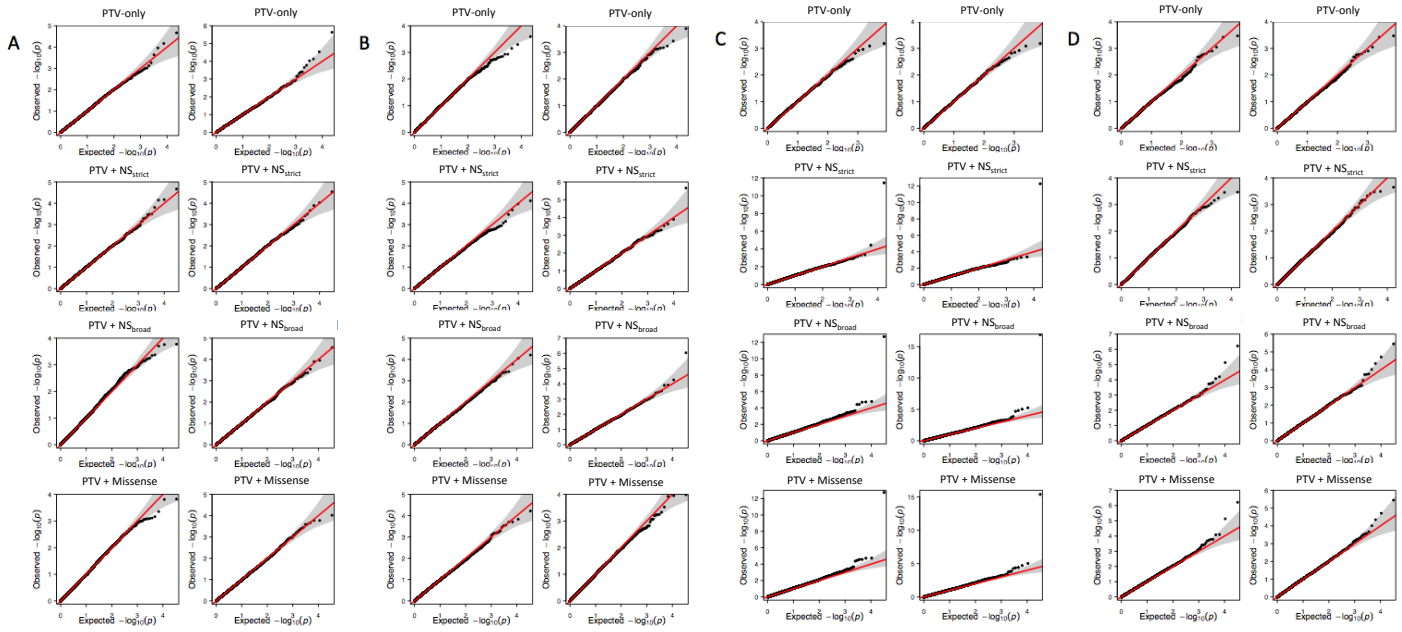


B. Fasting Insulin



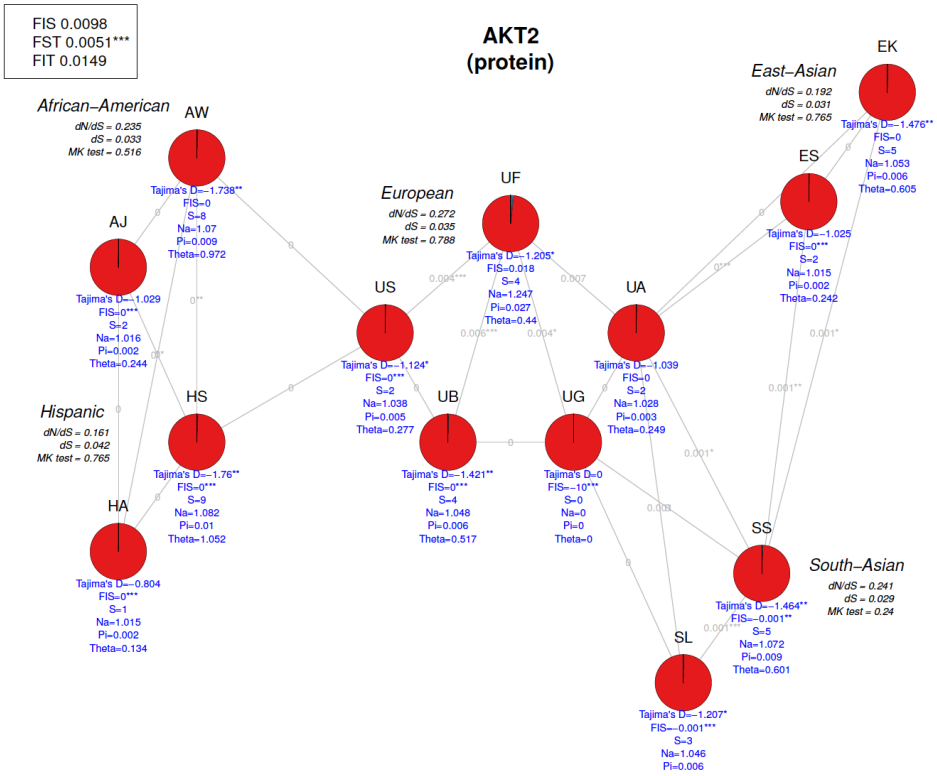
Manhattan and QQ plots for exome-wide association analysis with FG (A) and FI levels (B). On the Manhattan plots, variants within regions of known association are colored in dark blue, and variants outside those regions are colored in gray. The red horizontal line represents the exome-wide significance threshold for single variant associations ($P<2.5\times10^{-7}$). * For readability, the FG Manhattan plot is truncated at –log10(P) = 20, although variants in the *G6PC2* region on chromosome 2 have –log10(P values) > 20.

SUPPLEMENTARY FIGURE S2



QQ plots from the gene based association tests for FI and FG. Two tests were applied, SKAT (left column) and Burden (right column) to four annotation masks (PTV, PTV+NS_{Broad}, PTV+NS_{Strict}, PTV+Missense). **A.** FI with variants in exome sequencing data set. **B.** FG with variants in exome sequencing data set. **C.** FI with variants in exome chip data set. The point deviating from the diagonal is the association test for *AKT2*; see **Supplementary Table 2A** for association details. **D.** FG with variants in exome chip data set.

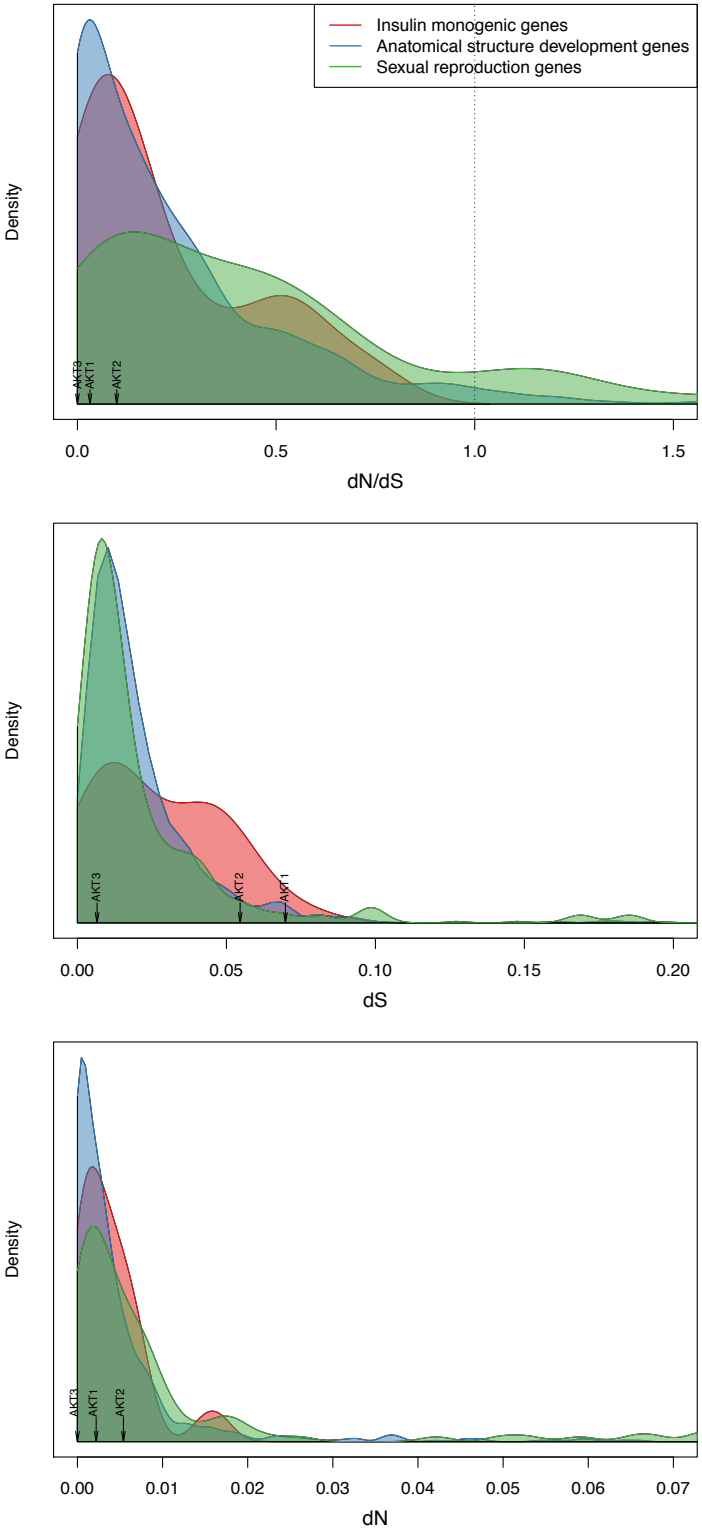
SUPPLEMENTARY FIGURE S3



Population structure and diversity indices of AKT2 protein in the exome sequencing data set. Each pie represents the frequency of different haplotypes, estimated from phased exome sequencing data in the five continental ancestries (grouped by study or country of origin). Significance of Tajima's D and F-statistics (global F_{ST}, F_{IS}, F_{IT}, and pairwise F_{ST} (gray line), and within population F_{IS}) are indicated with asterisk: * P-value < 0.05; ** P-value < 0.01; *** P-value < 0.001.

S: Number of segregating sites; Na: expected number of alleles; Pi (π): Mean number of pairwise differences; Theta (θ): Watterson's θ estimate; dN/ds: ratio of non-synonymous nucleotide substitutions per non-synonymous site (dN) and number of synonymous nucleotide substitutions per synonymous site (dS); MK: McDonald-Kreitman test.

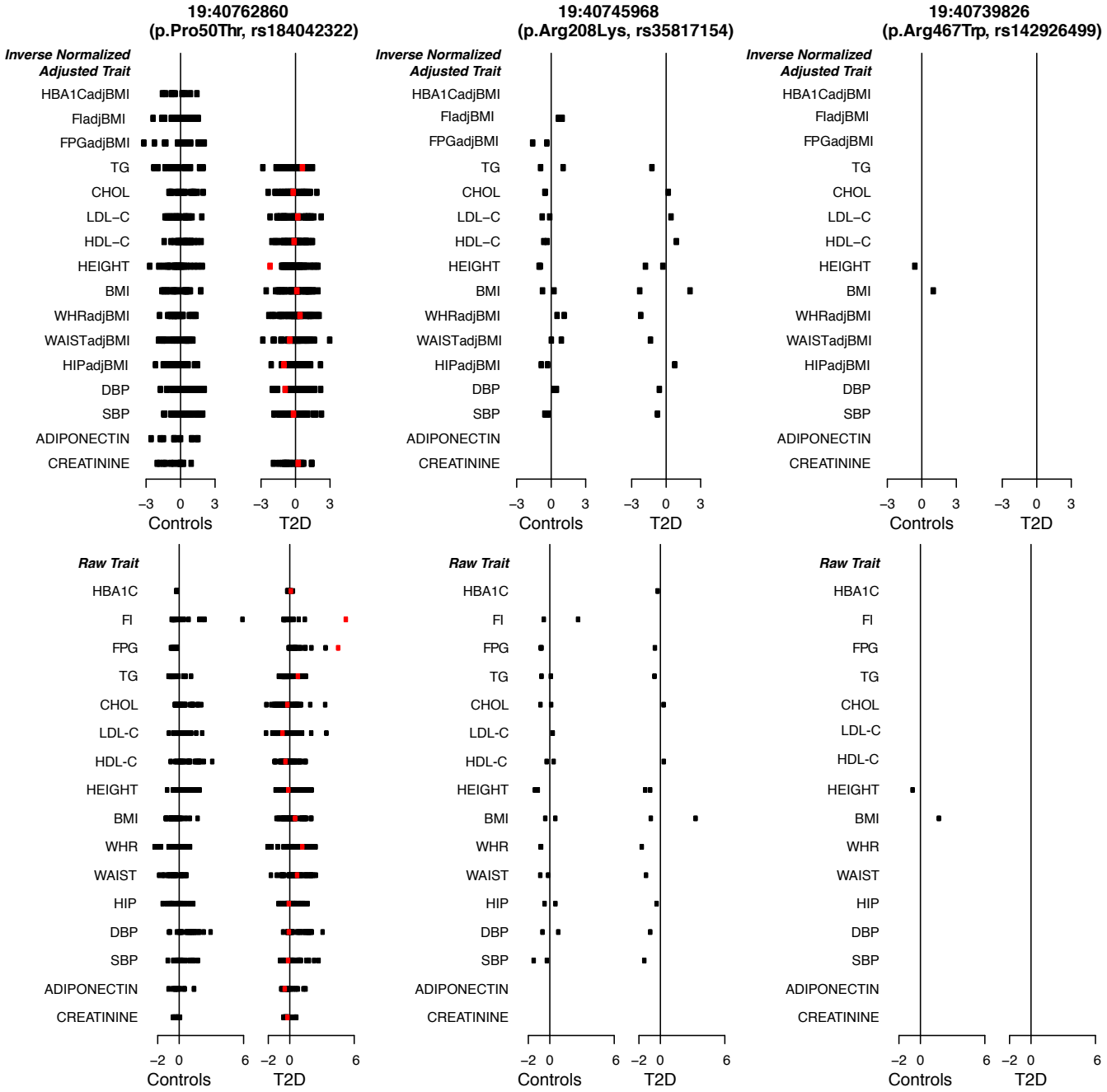
African-American: AJ – Jackson Heart Study, AW – Wake Forest School of Medicine Study; **East-Asian**: EK – Korea Association Research Project, ES – Singapore Diabetes Cohort Study and Singapore Prospective Study Program; **European**: UA – Ashkenazi (US, Israel), UB – UKT2D Consortium (UK), UF (Finland) – Metabolic Syndrome in Men Study (METSIM), Finland-United States Investigation of NIDDM Genetics (FUSION) Study, Malmo-Botnia Study, UG (Germany) – KORA-gen (Germany), US (Sweden) – Malmo-Botnia Study; **Hispanic**: HA – San Antonio Family Heart Study, San Antonio Family Diabetes/ Gallbladder Study, Veterans Administration Genetic Epidemiology Study, and the Investigation of Nephropathy and Diabetes Study family component, HS – Starr County, Texas; **South-Asian**: SL – London Life Sciences Population Study, SS – Singapore Indian Eye Study.



***AKT* family conservation compared to other genes.** The dN/dS ratio is calculated by comparing homologous coding sequences between human and chimpanzee. It shows the degree to which selection is acting on a gene: ratio<1 points to negative selection/purifying selection, i.e. evolutionary pressure to conserve the sequence in ancestral state, ratio>1 to positive selection, and ratio=1 to neutral evolution. Three *AKT* homologs are highly conserved when compared to the set of “Insulin monogenic” genes (37 genes), to which *AKT2* belongs, and two other gene sets: 1,002 anatomical structure development genes (“conserved”), and 132 sexual reproduction genes (“fast evolving”).

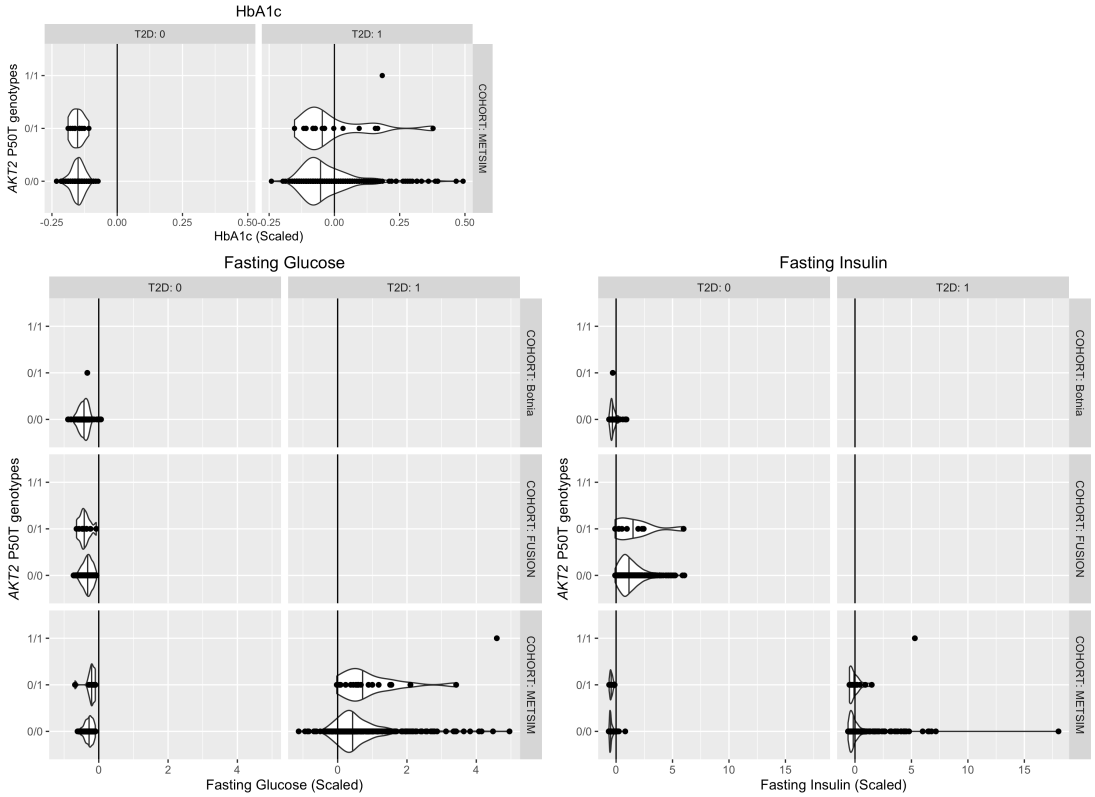
SUPPLEMENTARY FIGURE S5A

Diabetes



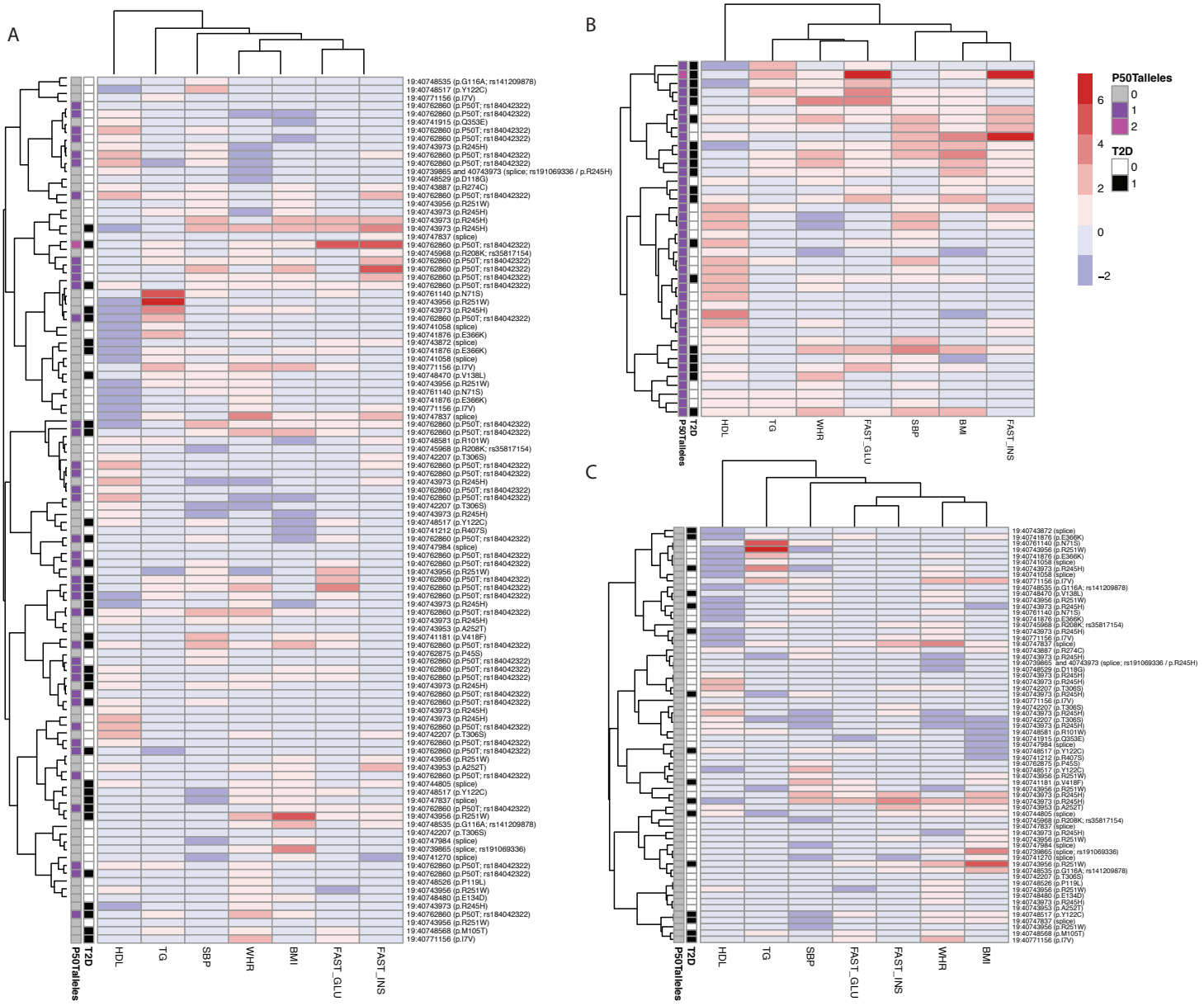
Trait values among *AKT2* variant carriers. Profile of the inverse normalized, adjusted metabolic trait values (top plot) and scaled (normalized by overall mean and standard deviation) raw trait values (bottom plot) of carriers of three *AKT2* variants: *AKT2* p.Pro50Thr, *AKT2* p.Arg208Lys and *AKT2* p.Arg467Trp from the T2D-GENES whole exome sequencing data set. Points on the graph are observed trait values for heterozygous (black) and homozygous (red) carriers of the variants, split by type 2 diabetes status. Trait abbreviations: HBA1C- glycated hemoglobin, FAST_INS- fasting insulin, FAST_GLU- fasting plasma glucose, TG- triglycerides, CHOL- total cholesterol, LDL-C, low-density lipoprotein cholesterol, HDL-C- high-density lipoprotein cholesterol, BMI- body mass index, WHR- waist to hip ratio, WASITC- waist circumference, HIPC- hip circumference, DBP- diastolic blood pressure, SBP- systolic blood pressure. adjBMI- trait adjusted for BMI

SUPPLEMENTARY FIGURE S5B



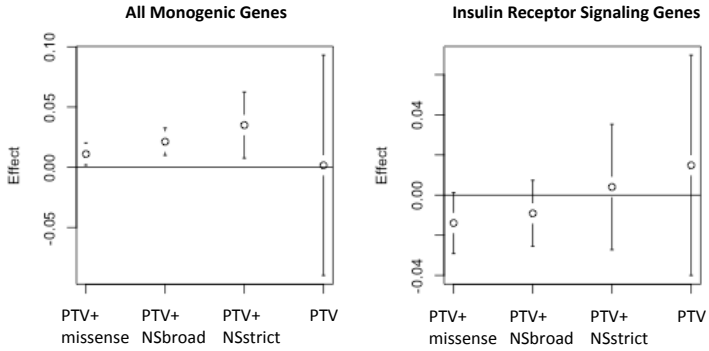
HbA1c, Fasting Glucose and Fasting Insulin distributions in T2D-GENES exome sequence data subset of Finnish cohorts (Botnia, FUSION, and METSIM). Scaled (normalized by overall mean and standard deviation) trait distributions are displayed by genotype group and type 2 diabetes status.

SUPPLEMENTARY FIGURE S5C



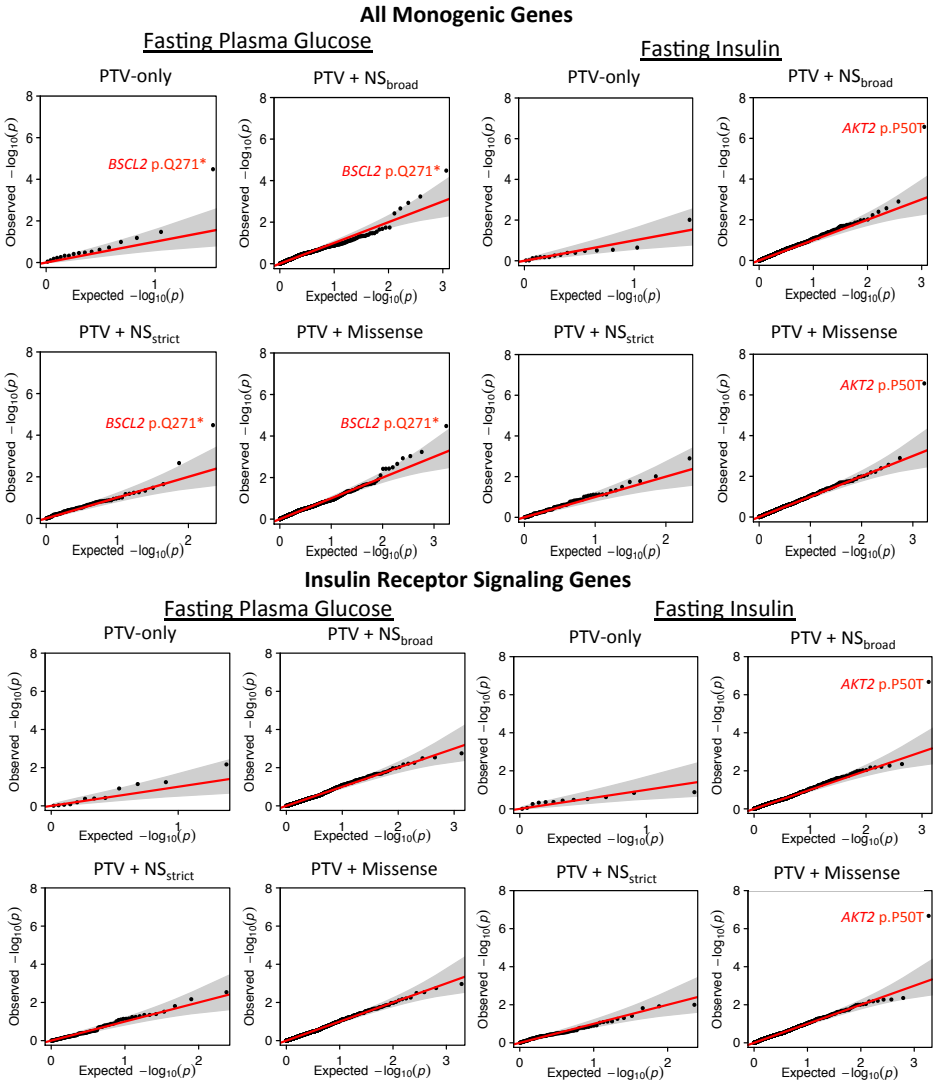
Phenotype clustering of *AKT2* missense variant carriers in the T2D-GENES whole exome sequencing dataset on seven metabolic traits: all missense carriers (**A**), carriers of *AKT2* p.Pro50Ala variant (**B**), and carriers of the other variants (**C**), (see **Supplementary Table 3**). The row labels indicate the variant carried by an individual. P50Talleles: the number of Ala alleles carried; T2D: 0 for controls and 1 for individuals with type 2 diabetes.

SUPPLEMENTARY FIGURE S6



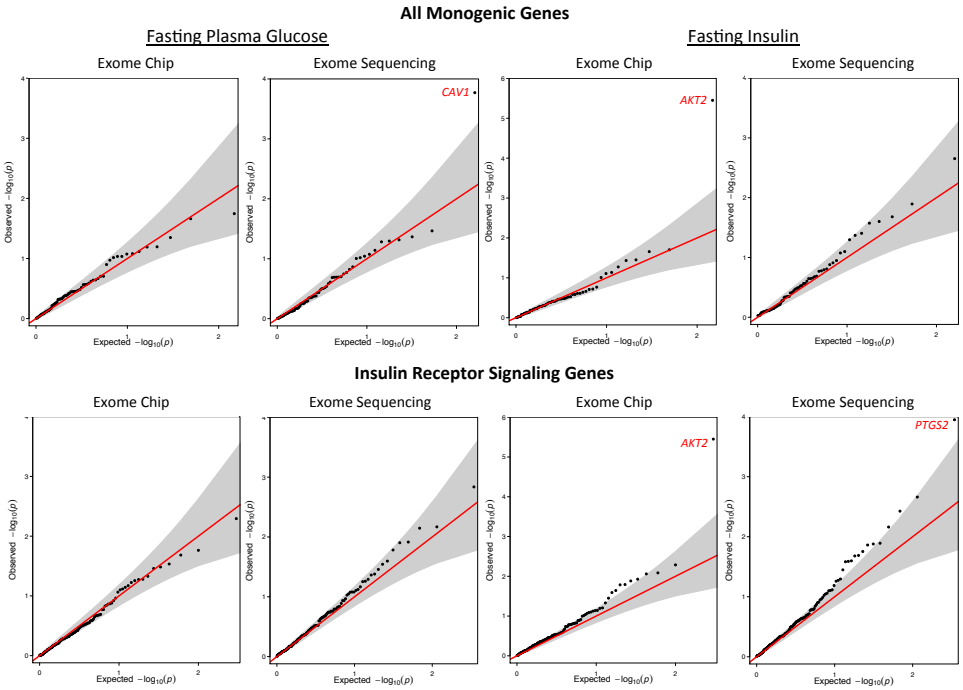
The trend in the estimate of the effect size of the global gene burden test for the four variant aggregation categories. The effect estimates (and 95% confidence interval) were provided as output of the burden test result in the RareMETALS package in R.

Supplementary Figure S7A



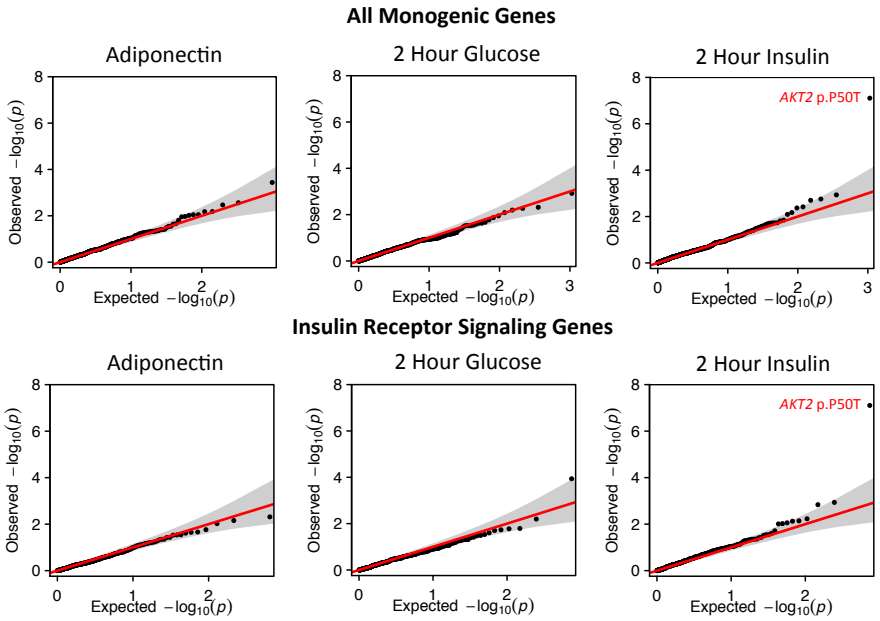
Monogenic enrichment in single variant association tests. Single variant association results from the FG and FI association analysis for variants in the four masks in the monogenic gene sets (top) and the insulin receptor signaling genes (bottom).

SUPPLEMENTARY FIGURE S7B



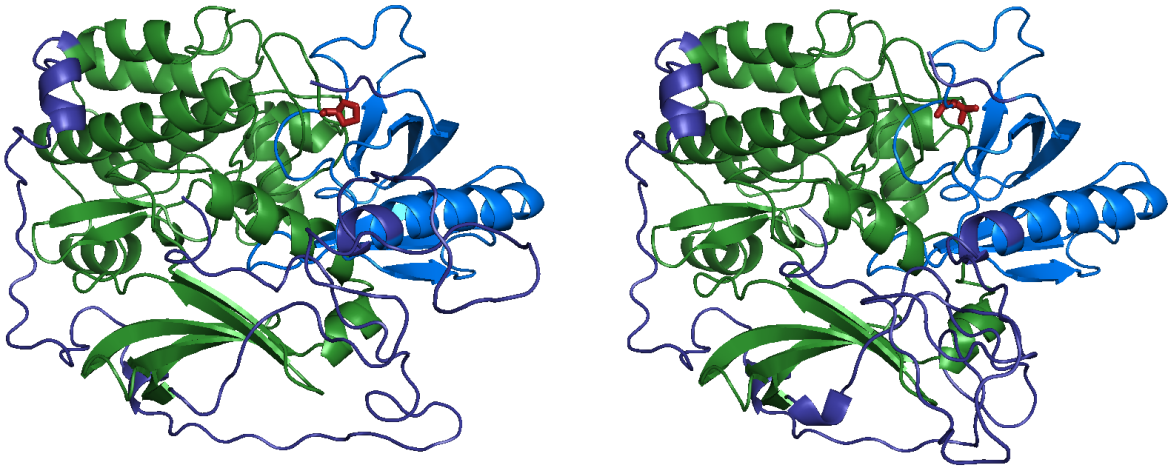
Pathway enrichment in gene-based tests. Gene burden association results from the fasting glucose and fasting insulin analysis for variants in the PTV+Missense mask in the monogenic gene sets (top) and the insulin receptor signaling genes (bottom).

SUPPLEMENTARY FIGURE S7C



Pathway associations in traits related to insulin resistance. Single variant association results for three traits related to insulin resistance: fasting adiponectin levels, 2 hour glucose level and 2 hour insulin level after an oral glucose tolerance test. The variants in these plots are in the PTV+Missense annotation category, with results from variants in the monogenic gene sets (top) and the insulin receptor signaling genes (bottom).

SUPPLEMENTARY FIGURE S8



Predicted structure change in AKT2 due to *AKT2* p.Pro50Thr. The left plot shows the predicted structure of wild-type AKT2. The right plot shows the predicted structure of AKT2.Thr50.

SUPPLEMENTARY FIGURE S9

A. General linear analysis

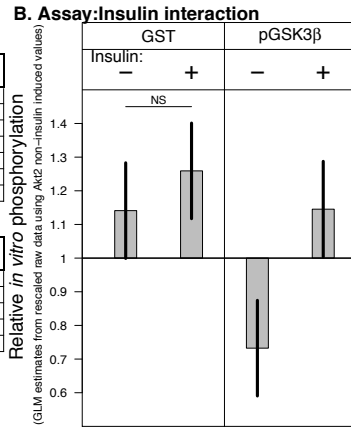
"Round" model:

Variables	*DF*	*Variance explained (%)*	*F*	*Pr(>F)*
Round	2	2.73%	1.228	0.300
Assay	1	8.42%	7.572	**0.008**
Insulin induction	1	12.38%	11.125	**0.001**
Round:Assay	2	1.60%	0.718	0.492
Round:Insulin	2	4.52%	2.033	0.140
Assay:Insulin	1	3.34%	2.999	0.088
Round:Assay:Insulin	2	0.27%	0.121	0.887

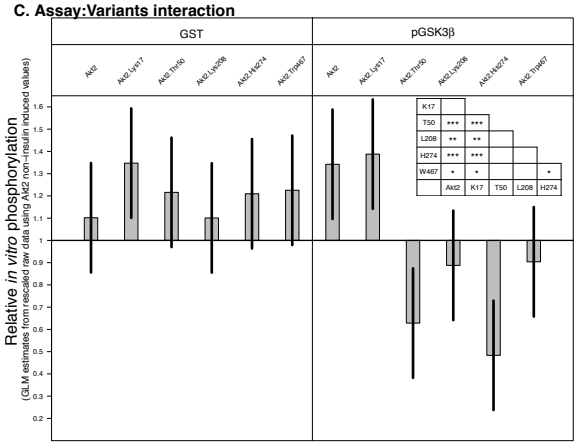
Full model:

Variables	*DF*	*Variance explained (%)*	*F*	*Pr(>F)*
Assay	1	8.42%	14.71	**3.12E-04**
Insulin induction	1	12.38%	21.61	**1.98E-05**
Variants	5	23.52%	8.21	**6.49E-06**
Assay:Insulin	1	3.34%	5.83	**1.90E-02**
Assay:Variant	5	19.13%	6.68	**5.64E-05**

B. Assay:Insulin interaction

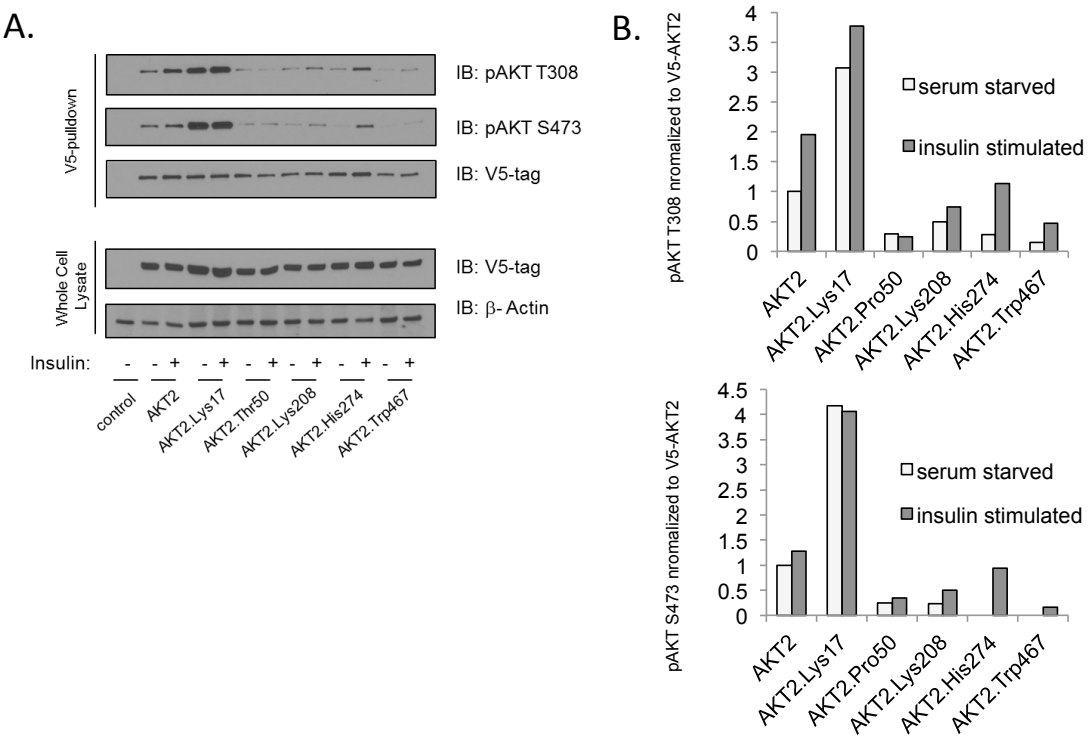


C. Assay:Variants interaction



In vitro kinase (IVK) assay. **A.** Results of a generalized linear model (GLM) applied on rescaled raw data. The relative substrate phosphorylation values were generated by dividing each value in each round of analysis with the value for non-stimulated, serum-starved AKT2. A first GLM ("Round" model) was analyzed including the Round as variable; the three independent rounds were not significant: we used them as replicate in the Full model. The plots represent the GLM estimates (and 95% CI) in the Full model for the two significant interactions: **B**. Assay:Insulin. **C**. Assay:Variants. For the Glycogen Synthase Kinase 3 β (GSK3β), the different AKT2 variants show significant relative phosphorylation (pairwise comparison p-values from contrast analysis reported in inset table). For GST-GSK3 peptide, none of the AKT2 variants showed different relative phosphorylation values. * P < 0.05, ** P < 0.01, *** P < 0.001. DF: degrees of freedom, F: statistic testing the importance of the grouping term, Pr(>F): P value of the F statistic.

SUPPLEMENTARY FIGURE S10



C.

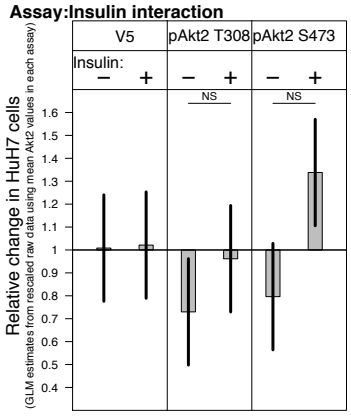
General linear analysis

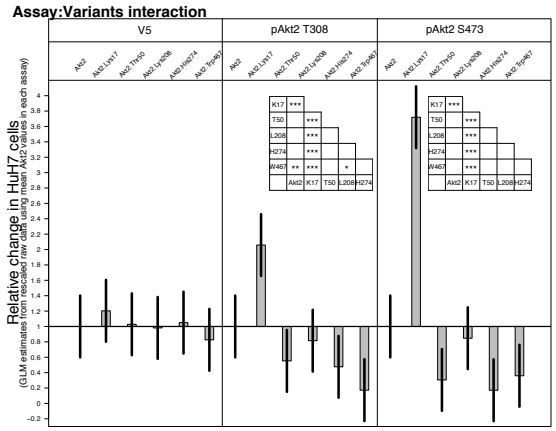
"Round" model:

Variables	*df*	*Variance explained (%)*	*F*	*Pr(>F)*
Round	2	1.86%	0.903	0.409
Assay	2	1.04%	0.504	0.606
Insulin induction	1	2.00%	1.941	0.167
Round:Assay	4	0.20%	0.049	0.995
Round:Insulin	2	0.11%	0.055	0.946
Assay:Insulin	2	1.37%	0.664	0.517
Round:Assay:Insulin	4	0.63%	0.152	0.962

Full model:

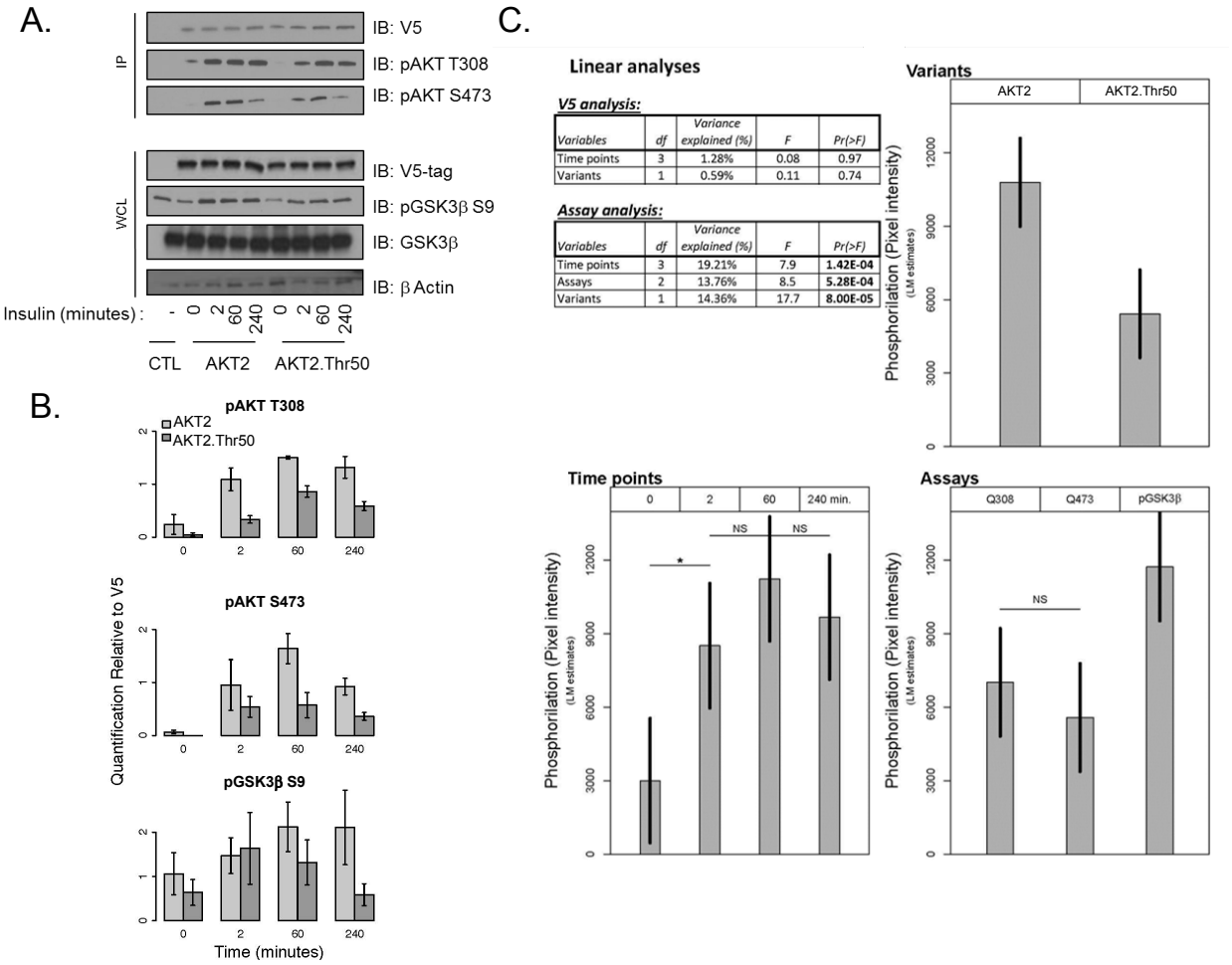
Variables	*df*	*Variance explained (%)*	*F*	*Pr(>F)*
Assay	2	1.04%	1.96	**1.47E-01**
Variants	5	46.52%	35.13	**2.20E-16**
Insulin induction	1	2.00%	7.56	**7.28E-03**
Assay:Variant	10	26.02%	9.83	**8.39E-11**
Assay:Insulin	2	1.37%	2.59	8.11E-02





Phosphorylation of AKT2 activation sites in HuH7 liver cells (A) HuH7 cells cells were infected with lentiviral control, V5-AKT2, V5-AKT2-Lys17, V5-AKT2-Thr50, V5-AKT2-Lys208, V5-AKT2-His274, V5-AKT2-Trp467, blasticidin selected and starved for 18 hr (white bar), and stimulated for 20 min with 100nm insulin (grey bar). V5-tagged AKT2 was isolated from cell lysates with anti-V5 agarose beads and immunoblots (IB) were probed with indicated antibodies. (B) Phosphorylated AKT2 Thr308 and Ser473 were quantified and normalized to total by V5-AKT2. (C) Linear model for the statistical analysis of quantified pAKT2. The "Round" model tests for significant differences between the three rounds of analysis. The Full model examines significance of assay (V5, pAKT2 T308 and pAKT2 S473) and variants (AKT2, AKT2.Lys17, AKT2.Thr50, AKT2.Lys208, AKT2.His274 and AKT2.Trp467) and their interactions. * P < 0.05, ** P < 0.01, *** P < 0.001. DF: degrees of freedom, F: statistic testing the importance of the grouping term, Pr(>F): P value of the F statistic.

SUPPLEMENTARY FIGURE S11

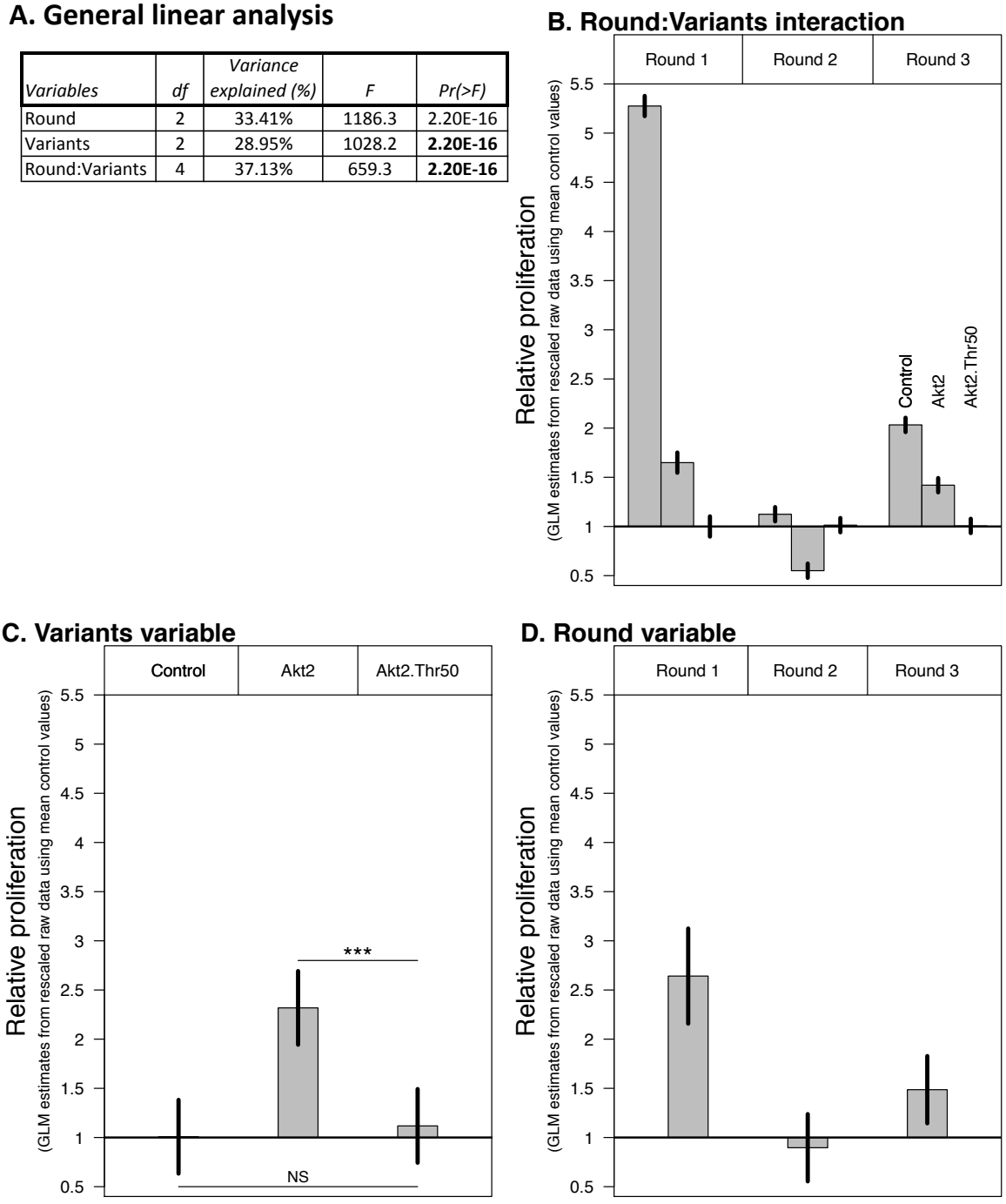


Time-course analysis of AKT2 phosphorylation (A) HeLa cells were infected with lentiviral V5-AKT2, V5-AKT2-Thr50, or control pLX304, blasticidin selected and starved for 18 hours and then stimulated for 0, 2, 60, and 240 minutes with 100nm insulin. V5-tagged AKT2 was isolated from cell lysates with anti-V5 agarose beads. Immunoprecipitated (IP) V5-AKT2 and whole cell lysates (WCL) were immunoblotted (IB) with the indicated antibodies. Immunoblots are representative of three independent replicates. (B) Quantification of the three replicates of indicated immunoblots relative to total V5-AKT2. (C) Linear Model (LM) statistical analysis across all three independent replicates. Error bars represent the standard deviation (SD). * P < 0.05, ** P < 0.01, *** P < 0.001.

SUPPLEMENTARY FIGURE S12

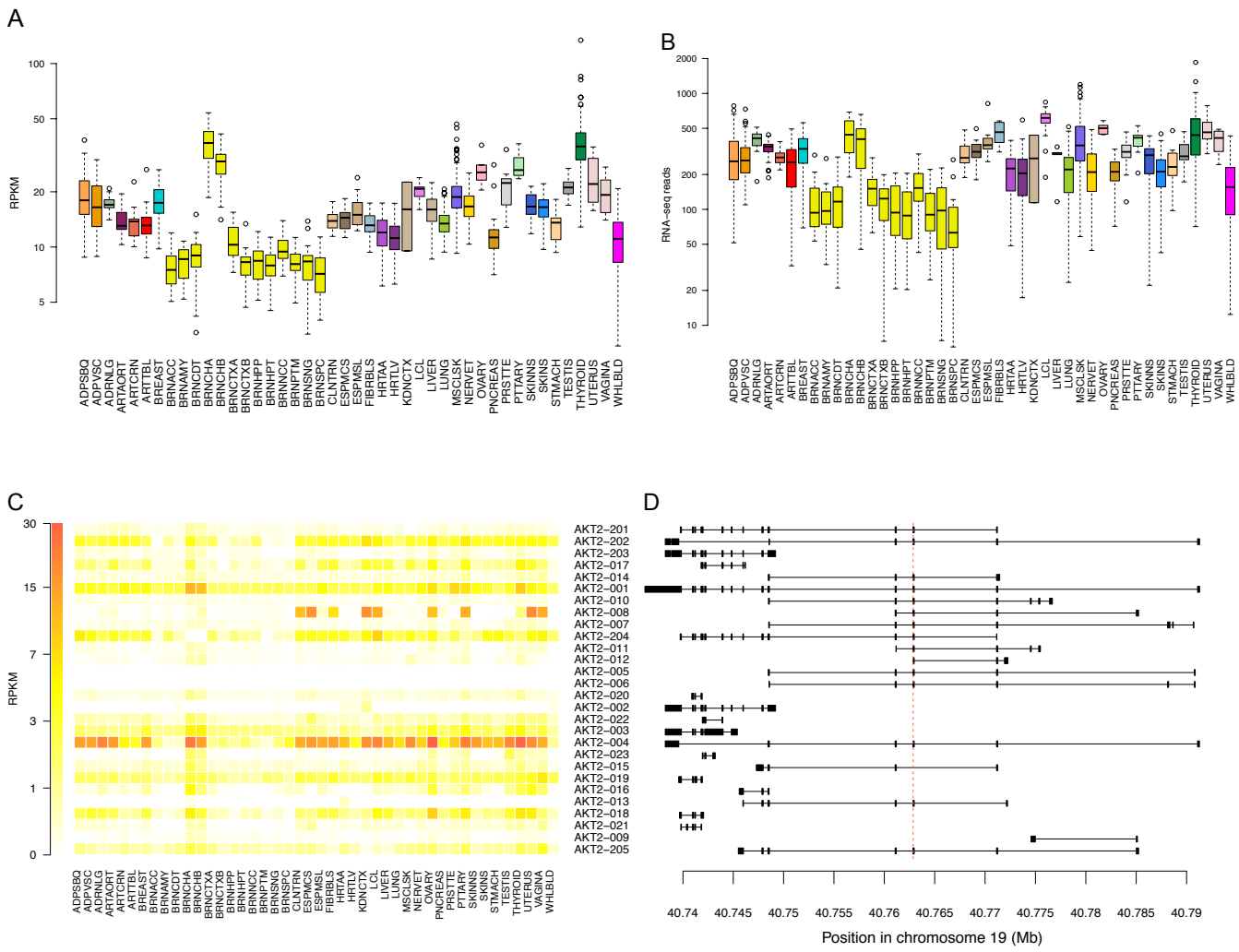
A. General linear analysis

Variables	*df*	*Variance explained (%)*	*F*	*Pr(>F)*
Round	2	33.41%	1186.3	2.20E-16
Variants	2	28.95%	1028.2	**2.20E-16**
Round:Variants	4	37.13%	659.3	**2.20E-16**



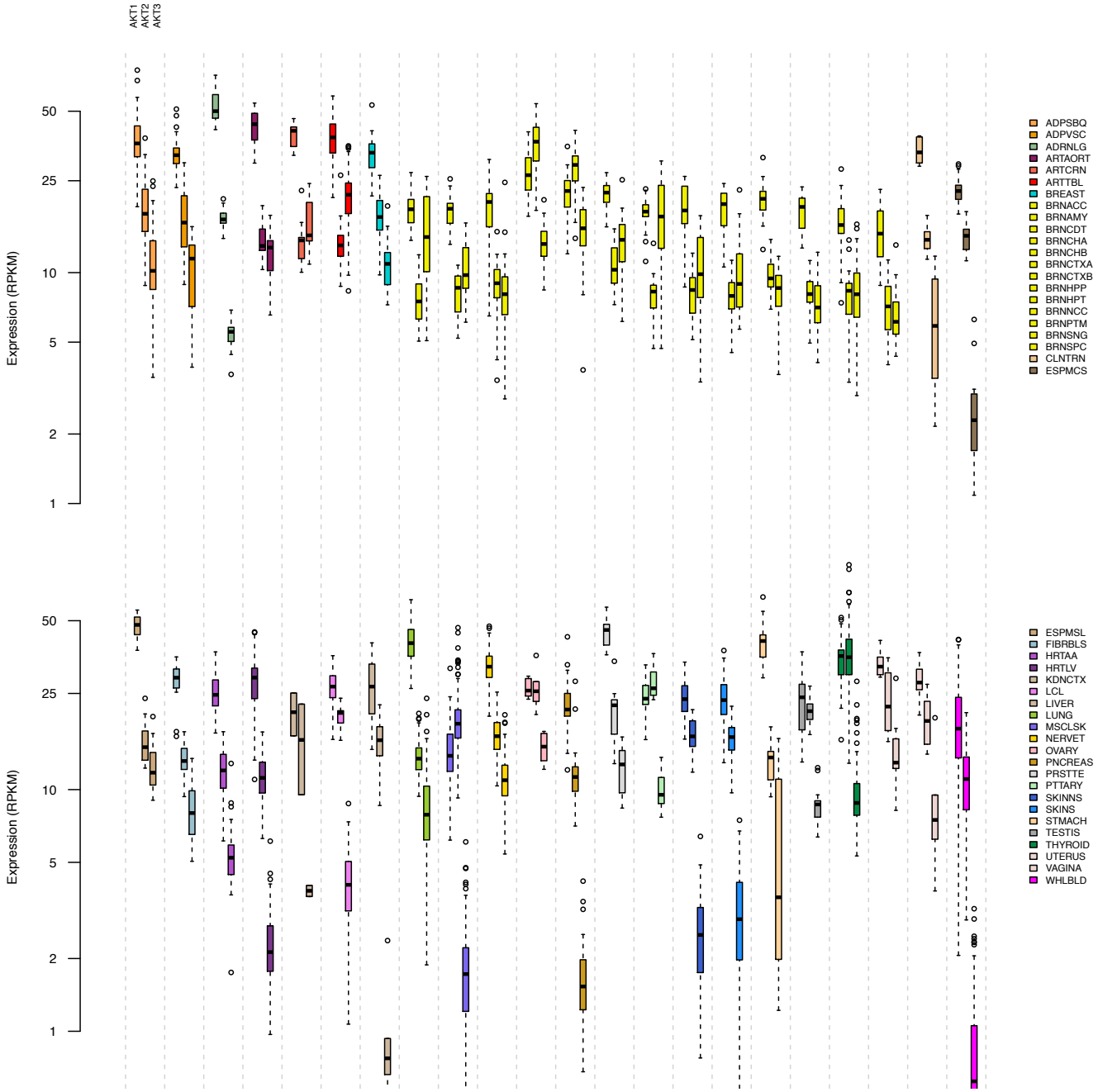
Proliferation assay. A. Results of a generalized linear model (GLM) applied on rescaled raw data (absorbance value) to test for significant difference in proliferation between the three rounds of analysis, the three variants and an interaction between round and variants. The rescaling was performed by dividing all the values in each round by the average absorbance in controls. The plots represent the GLM estimates (and 95% CI) for the **B.** Round:Variant interaction and individual variables: **C.** Round and **D.** Variants. * P < 0.05, ** P < 0.01, *** P < 0.001. DF: degrees of freedom, F: statistic testing the importance of the grouping term, Pr(>F): P value of the F statistic.

SUPPLEMENTARY FIGURE S13



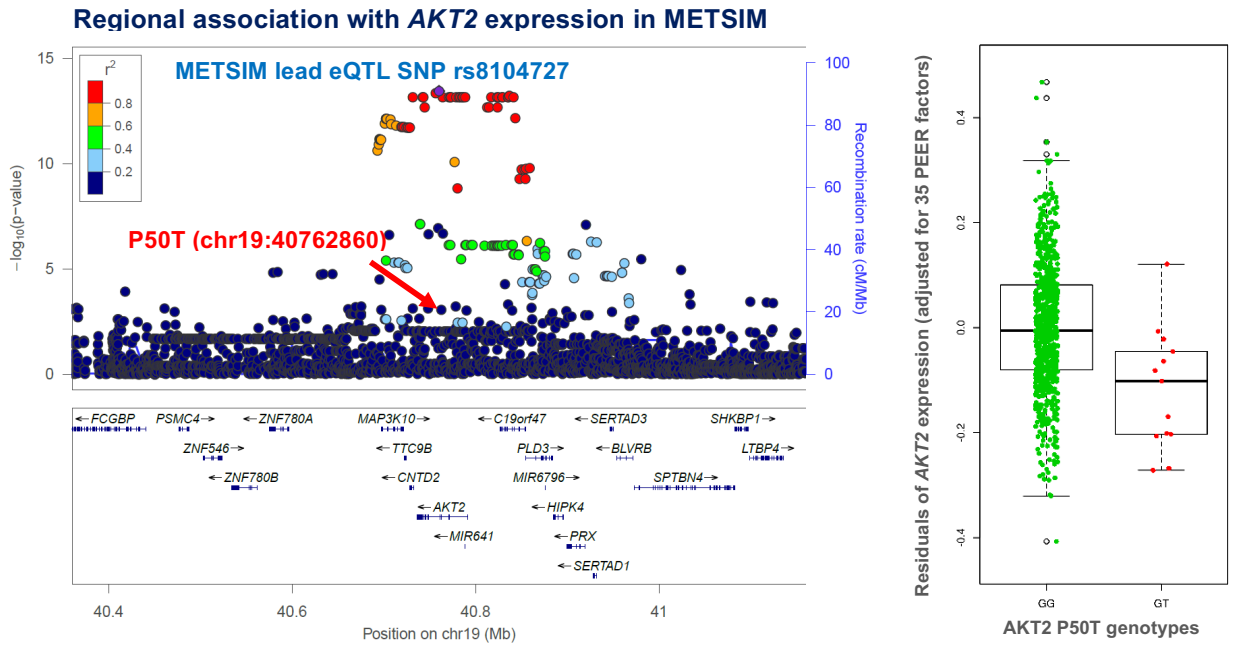
***AKT2* expression in human tissues. A**. Boxplot displaying the level and distribution of *AKT2* gene expression (in reads per kilobase per million mapped reads, RPKM) in 44 human tissues available in the GTEx RNA-seq data. **B**. Box plot of the expression (in RNA-seq reads) of the *AKT2* exon of affected by the p.Pro50Thr variant. Read counts are not normalized by the total number of reads per sample, resulting in larger variance in the expression within each tissue. **C**. Heat map of expression patterns of the 28 *AKT2* transcripts in the GTEx tissues, as annotated in Gencode version 12. Intensity of color in each cell represents the expression of the transcript in that tissue; white indicating no expression, and red indicating higher expression. **D**. Visualization of the transcript structure of *AKT2* (Gencode v12). The affected exon, highlighted with the red dashed line, is included in the majority of the *AKT2* transcripts and in all the three most highly expressed transcripts. The tissues are presented in the same order across panels A-C, and colored similarly in panels A and B. Tissue abbreviations are listed in **Supplementary Table 8**.

SUPPLEMENTARY FIGURE S14



Expression of the *AKT* gene family across human tissues. Each cluster of three boxplots represents the expression of *AKT1* (**left**), *AKT2* (**middle**) and *AKT3* (**right**) in each tissue. *AKT2* is the isoform with the highest expression (P-value < 0.05) in BRNCHA (Brain – Cerebellum), BRNCHB (Brain - Cerebellar Hemisphere), MSCLSK (Muscle – Skeletal) and PTTARY (Pituitary). Tissue abbreviations are listed in **Supplementary Table 8**.

SUPPLEMENTARY FIGURE S15



	Increasing allele / decreasing alleles	Frequency of decreasing allele	Initial Effect of decreasing allele	P	Conditional Effect of decreasing allele	Conditional P
AKT2 Pro50Thr	G/T	0.0083	-0.980	8.9E-04	-0.754	8.4E-03
Lead eSNP rs8104727	T/C	0.647	-0.403	3.6E-14	-0.391	1.9E-13

Expression analysis with common eQTL SNP and *AKT2* p.Pro50Thr. Top left plot: The regional association plot of variants in the *AKT2* region testing association with *AKT2* expression. The SNP showing the most significant signal in this plot, rs8104727, is a proxy for rs11880261 (r^2 = 1, D' = 1 in the 1000 Genomes phase 3 Finnish sample). Top right plot: observed AKT2 expression levels for the two *AKT2* p.Pro50Thr genotypes observed in the METSIM cohort. Bottom table: eQTL statistics and reciprocal conditional analysis with the two SNPs: rs8104727and *AKT2* p.Pro50Thr. The "Beta conditional" and "P conditional" columns highlight the associations with *AKT2* expression after conditioning on the other SNP.

Supplementary Tables

SUPPLEMENTARY TABLE 1 .

Details and characteristics of studies included in the analysis.

Supplementary Table 1A: Study details including references, ascertainment, sample QC, variant QC and association covariates.

Stage	Ancestry	Study	Citation(s)	PubMed ID(s)	Sample Ascertainment	Genotyping array	Call rate	Exclusion criteria	Call rate	Filtering criteria	Calling algorithm	Association covariates
Discovery [ExomeChip]	European [Finnish]	FIN-D2D 2007	Kotronen, A. et al. Non-alcoholic and alcoholic fatty liver disease - two diseases of affluence associated with the metabolic syndrome and type 2 diabetes: the FIN-D2D survey. BMC Public Health. 2010 May 10;10:237.	20459722	- Population-based survey - Glucose tolerance classified according to WHO 1999 criteria - T1D and T2D (fasting plasma glucose concentration ≥7.0 mmol/l or 2-h plasma glucose concentration ≥11.1 mmol/l) cases excluded - Further excluded individuals with HbA1c ≥6.5% according to ADA 2012 criteria for T2D	Illumina HumanExome-12v1-1_A	>99%	- call rate ≤99% - heterozygosity >median + 3*IQR - technical duplicates with lower call rate - Non-European population outliers - sex discrepancy - contamination score >10%	≥95%	- exclude 101 indels with different allele mapping across the two sites - exclude variants that had chromosome or position or allele mismatch - exclude duplicated variants, keeping the one with higher call rate. If the call rate is the same, arbitrarily take the one with that comes first in the file - call rate <95% - exact HWE $<10^{-5}$	Illumina GenCall using standard Illumina cluster files + Zcall	- age, age2, sex, BMI for EMMAX-analysis - age, age2, sex, BMI, PC1, PC2, PC3, PC4 for rvtest analysis
Discovery [ExomeChip]	European [Finnish]	The Finnish Diabetes Prevention Study (DPS)	Tuomilehto, J. et al. Prevention of type 2 diabetes mellitus by changes in lifestyle among subjects with impaired glucose tolerance. N Engl J Med. 2001 May 3;344(18):1343-50.	11333990	- Randomised controlled trial - All subjects were impaired glucose tolerant at baseline, from mean of two OGTTs using WHO 1985 criteria - Excluded individuals with fasting plasma glucose ≥7.0 mmol/l or 2-h plasma glucose concentration ≥11.1 mmol/l or HbA1c ≥6.5% according to ADA 2012 criteria for T2D	Illumina HumanExome-12v1-1_A	>99%	- call rate ≤99% - heterozygosity >median + 3*IQR - technical duplicates with lower call rate - Non-European population outliers - sex discrepancy - contamination score >10%	≥95%	- exclude 101 indels with different allele mapping across the two sites - exclude variants that had chromosome or position or allele mismatch - exclude duplicated variants, keeping the one with higher call rate. If the call rate is the same, arbitrarily take the one with that comes first in the file - call rate <95% - exact HWE $<10^{-5}$	Genotype calls generated on cluster boundaries trained on using study samples + manual review of clusterplots	- age, age2, sex, BMI for EMMAX-analysis - age, age2, sex, BMI, PC1, PC2, PC3, PC4 for rvtest analysis
Discovery [ExomeChip]	European [Finnish]	The Dose Responses to Exercise Training (DR's EXTRA) Study	Kouki, R. et al. Diet, fitness and metabolic syndrome--the DR's EXTRA study. Nutr Metab Cardiovasc Dis. 2012 Jul;22(7):553-60.	21186108	- Randomised controlled trial - Glucose tolerance classified according to WHO 1999 criteria - T1D and T2D (fasting plasma glucose concentration ≥7.0 mmol/l or 2-h plasma glucose concentration ≥11.1 mmol/l or physician diagnosed) cases excluded	HumanExome-12v1-1_A	>99%	- call rate ≤99% - heterozygosity >median + 3*IQR - technical duplicates with lower call rate - Non-European population outliers - sex discrepancy - contamination score >10%	≥95%	- exclude 101 indels with different allele mapping across the two sites - exclude variants that had chromosome or position or allele mismatch - exclude duplicated variants, keeping the one with higher call rate. If the call rate is the same, arbitrarily take the one with that comes first in the file - call rate <95% - exact HWE $<10^{-5}$	Illumina GenCall using standard Illumina cluster files + Zcall	- age, age2, sex, BMI for EMMAX-analysis - age, age2, sex, BMI, PC1, PC2, PC3, PC4 for rvtest analysis
Discovery [ExomeChip]	European [Finnish]	National FINRISK 2007 Study (FINRISK 2007)	Vartiainen, E. et al. Thirty-five-year trends in cardiovascular risk factors in Finland. Int J Epidemiol. 2010 Apr;39(2):504-18.	19959603	- T2D case control study - Glucose tolerance classified according to WHO 1999 criteria - T1D and T2D (fasting plasma glucose concentration ≥7.0 mmol/l or 2-h plasma glucose concentration ≥11.1 mmol/l) cases excluded	HumanExome-12v1-1_A	>99%	- call rate ≤99% - heterozygosity >median + 3*IQR - technical duplicates with lower call rate - Non-European population outliers - sex discrepancy - contamination score >10%	≥95%	- exclude 101 indels with different allele mapping across the two sites - exclude variants that had chromosome or position or allele mismatch - exclude duplicated variants, keeping the one with higher call rate. If the call rate is the same, arbitrarily take the one with that comes first in the file - call rate <95% - exact HWE $<10^{-5}$	Illumina GenCall using standard Illumina cluster files + Zcall	- age, age2, sex, BMI for EMMAX-analysis - age, age2, sex, BMI, PC1, PC2, PC3, PC4 for rvtest analysis
Discovery [ExomeChip]	European [Finnish]	Finland-United States Investigation of NIDDM Genetics (FUSION) Study	Valle, T. et al. Mapping genes for NIDDM. Design of the Finland-United States Investigation of NIDDM Genetics (FUSION) Study. Diabetes Care. 1998 Jun;21(6):949-58.; Scott, L.J., et al. A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. Science. 2007 Jun 1;316(5829):1341-5.	9614613; 17463248	- T2D case control study - Glucose tolerance classified according to WHO 1999 criteria - T2D (fasting plasma glucose concentration ≥7.0 mmol/l or 2-h plasma glucose concentration ≥11.1 mmol/l, by report of diabetes medication use, or based on medical record review), and known or probable T1D among their first degree relatives were excluded.	HumanExome-12v1-1_A	>99%	- call rate ≤99% - heterozygosity >median + 3*IQR - technical duplicates with lower call rate - Non-European population outliers - sex discrepancy - contamination score >10%	≥95%	- exclude 101 indels with different allele mapping across the two sites - exclude variants that had chromosome or position or allele mismatch - exclude duplicated variants, keeping the one with higher call rate. If the call rate is the same, arbitrarily take the one with that comes first in the file - call rate <95% - exact HWE $<10^{-5}$	Illumina GenCall using standard Illumina cluster files + Zcall	- age, age2, sex, BMI, study origin for EMMAX-analysis - age, age2, sex, BMI, study origin, PC1, PC2, PC3, PC4 for rvtest analysis
Discovery [ExomeChip]	European [Finnish]	Metabolic Syndrome in Men Study (METSIM)	Stancáková, A. et al. Changes in insulin sensitivity and insulin release in relation to glycemia and glucose tolerance in 6,414 Finnish men. Diabetes. 2009 May;58(5):1212-21.	19223598	- Population-based cross-sectional study - Glucose tolerance classified according to WHO 1997 criteria - T1D and T2D (fasting plasma glucose concentration ≥7.0 mmol/l or 2-h plasma glucose concentration ≥11.1 mmol/l) cases excluded - Further excluded individuals with HbA1c ≥6.5% according to ADA 2012 criteria for T2D	HumanExome-12v1_A	>99%	- call rate ≤99% - heterozygosity >median + 3*IQR - technical duplicates with lower call rate - Non-European population outliers - sex discrepancy - contamination score >10%	≥95%	- exclude 101 indels with different allele mapping across the two sites - exclude variants that had chromosome or position or allele mismatch - exclude duplicated variants, keeping the one with higher call rate. If the call rate is the same, arbitrarily take the one with that comes first in the file - call rate <95% - exact HWE $<10^{-5}$	Genotype calls generated on cluster boundaries trained on using study samples + manual review of clusterplots	- age, age2, BMI for EMMAX-analysis - age, age2, BMI, PC1, PC2, PC3, PC4 for rvtest analysis
Discovery [ExomeChip]	European [Danish]	Health2006	Thuesen, B.H. et al. Cohort Profile: The Health2006 cohort, Research Centre for Prevention and Health. Int J Epidemiol. 2013 Apr 24.	23615486	- Population-based cohort - Glucose tolerance classified according to WHO 1999 criteria - T1D and T2D (fasting plasma glucose concentration ≥7.0 mmol/l) cases excluded	Illumina HumanExome-12v1	≥98%	- call rate <98% - heterozygosity - sex discrepancy - discordance with previous genotypes	≥95%	- exclude duplicated variants, keeping the one with higher call rate. - call rate <95% - HWE $<10^{-4}$ - cluster separation score 0.4	Illumina GenCall using standard Illumina cluster files + Zcall	- age, age^2, BMI for EMMAX-analysis - age, age^2, BMI, PC1-10 for RareMetalWorker analysis
Discovery [ExomeChip]	European [Danish]	Inter99	Jørgensen, T. et al. A randomized non-pharmacological intervention study for prevention of ischaemic heart disease: baseline results Inter99. Eur J Cardiovasc Prev Rehabil. 2003 Oct;10(5):377-86.	14663300	- Population-based cohort - Glucose tolerance classified according to WHO 1999 criteria - T1D and T2D (fasting plasma glucose concentration ≥7.0 mmol/l or 2-h plasma glucose concentration ≥11.1 mmol/l) cases excluded	Illumina HumanExome-12v1	≥98%	- call rate <98% - heterozygosity - sex discrepancy - discordance with previous genotypes	≥95%	- exclude duplicated variants, keeping the one with higher call rate. - call rate <95% - HWE $<10^{-4}$ - cluster separation score 0.4	Illumina GenCall using standard Illumina cluster files + Zcall	- age, age^2, BMI for EMMAX-analysis - age, age^2, BMI, PC1-10 for RareMetalWorker analysis
Discovery [ExomeChip]	European [Danish]	Vejle Biobank	Albrechtsen, A. et al. Exome sequencing-driven discovery of coding polymorphisms associated with common metabolic phenotypes. Diabetologia. 2013 Feb;56(2):298-310.	23160641	- Controls from T2D case-control - Glucose tolerance classified according to WHO 1999 criteria - T1D and T2D (fasting plasma glucose concentration ≥7.0 mmol/l or 2-h plasma glucose concentration ≥11.1 mmol/l) cases excluded	Illumina HumanExome-12v1	≥98%	- call rate <98% - heterozygosity - sex discrepancy - discordance with previous genotypes	≥95%	- exclude duplicated variants, keeping the one with higher call rate. - call rate <95% - HWE $<10^{-4}$ - cluster separation score 0.4	Illumina GenCall using standard Illumina cluster files + Zcall	- age, age^2, BMI for EMMAX-analysis - age, age^2, BMI, PC1-10 for RareMetalWorker analysis
Discovery [ExomeChip]	European [UK]	Genetics of Diabetes Audit and Research Tayside (GoDARTS)	Morris, A.D. et al. The diabetes audit and research in Tayside Scotland (DARTS) study: electronic record linkage to create a diabetes register. DARTS/MEMO Collaboration. BMJ. 1997 Aug 30;315(7107):524-8.	9329309	- Population-based cohort - T2D cases, sample with fasting plasma glucose concentration ≥7.0 mmol/l and pregnant women were excluded	Illumina HumanExome-12v1_A	>99%	- call rate ≤99% - heterozygosity 4SD of mean - technical duplicates with lower call rate - Non-European population outliers, or non-European reported ancestry - sex discrepancy	>99%	- exclude variants that had chromosome or position or allele mismatch - exclude duplicated variants, keeping the one with higher call rate. - call rate <98% for GenCall and <99% for zCall - exact HWE $<10^{-4}$ - GenTrain score <0.6 and Cluster separation score <0.4	Illumina GenCall using standard Illumina cluster files + Zcall	- age, age^2, sex, and BMI for EMMAX-analysis - age, age^2, sex, BMI, PC1, and PC2 for RareMetalWorker analysis
Discovery [ExomeChip]	European [UK]	Twins UK	Moayyeri A, Hammond CJ, Hart DJ, Spector TD. The UK Adult Twin Registry (TwinsUK Resource). Twin Res Hum Genet. 2013 Feb;16(1):144-9.	23088889	- Unrelated samples selected as controls from the Twins UK study - T1D and T2D cases and samples with recorded family history of diabetes, or if either twin was ever recorded as impaired glucose tolerant (defined as fasting plasma glucose concentration >6.1mmol/L in any reading), non-fasting were excluded.	Illumina HumanExome-12v1_A	>99%	- call rate ≤99% - heterozygosity 4SD of mean - technical duplicates with lower call rate - Non-European population outliers, or non-European reported ancestry - sex discrepancy	>99%	- exclude variants that had chromosome or position or allele mismatch - exclude duplicated variants, keeping the one with higher call rate. - call rate <98% for GenCall and <99% for zCall - exact HWE $<10^{-4}$ - GenTrain score <0.6 and Cluster separation score <0.4	Illumina GenCall using standard Illumina cluster files + Zcall	- age, age^2, sex, and BMI for EMMAX-analysis - age, age^2, sex, BMI, PC1, and PC2 for RareMetalWorker analysis

Stage	Ancestry	Study	Citation(s)	PubMed ID(s)	Sample Ascertainment	Genotyping array	Call rate	Exclusion criteria	Call rate	Filtering criteria	Calling algorithm	Association covariates
Discovery [ExomeChip]	European [UK]	Oxford BioBank (OBB)	http://www.oxfordbiobank.org.uk/	NA	- T2D cases (on diabetic treatment or fasting glucose ≥7 mmol/l) were excluded.	Illumina HumanExome-12v1_A	>99%	- call rate ≤99% - heterozygosity 4SD of mean - technical duplicates with lower call rate - Non-European population outliers, or non-European reported ancestry - sex discrepancy	>99%	- exclude variants that had chromosome or position or allele mismatch - exclude duplicated variants, keeping the one with higher call rate. - call rate <98% for GenCall and <99% for zCall - exact HWE $<10^{-4}$ - GenTrain score <0.6 and Cluster separation score <0.4	Illumina GenCall using standard Illumina cluster files + Zcall	- age, age^2, sex, and BMI for EMMAX-analysis - age, age^2, sex, BMI, PC1, and PC2 for RareMetalWorker analysis
Discovery [ExomeChip]	European [Swedish]	Prospective Investigation of the Vasculature in Uppsala Seniors (PIVUS)	Lind, L. et al. A comparison of three different methods to evaluate endothelium-dependent vasodilation in the elderly: the Prospective Investigation of the Vasculature in Uppsala Seniors (PIVUS) study. Arterioscler Thromb Vasc Biol. 2005 Nov;25(11):2368-75.	16141402	- Population-based cohort - T1D, T2D cases or fasting plasma glucose concentration ≥7 mmol/l, pregnant individuals, and samples with non-fasting blood excluded	Illumina HumanExome-12v1_A	>99%	- call rate ≤99% - heterozygosity 4SD of mean - technical duplicates with lower call rate - Non-European population outliers - sex discrepancy	>99%	- exclude variants that had chromosome or position or allele mismatch - exclude duplicated variants, keeping the one with higher call rate. - call rate <98% for GenCall and <99% for zCall - exact HWE $<10^{-4}$ - GenTrain score <0.6 and Cluster separation score <0.4	Illumina GenCall using standard Illumina cluster files + Zcall	- age, age^2, sex, and BMI for EMMAX-analysis - age, age^2, sex, BMI, PC1, and PC2 for RareMetalWorker analysis
Discovery [ExomeChip]	European [Swedish]	Uppsala Longitudinal Study of Adult Men (ULSAM)	Hedstrand, H. A study of middle-aged men with particular reference to risk factors for cardiovascular disease. Ups J Med Sci Suppl. 1975;19:1-61.	1216390	- Population-based cohort - T1D, T2D cases or fasting plasma glucose concentration ≥7 mmol/l, and samples with non-fasting blood excluded	Illumina HumanExome-12v1_A	>99%	- call rate ≤99% - heterozygosity 4SD of mean - technical duplicates with lower call rate - Non-European population outliers - sex discrepancy	>99%	- exclude variants that had chromosome or position or allele mismatch - exclude duplicated variants, keeping the one with higher call rate. - call rate <98% for GenCall and <99% for zCall - exact HWE $<10^{-4}$ - GenTrain score <0.6 and Cluster separation score <0.4	Illumina GenCall using standard Illumina cluster files + Zcall	- age, age^2, and BMI for EMMAX-analysis - age, age^2, sex, BMI, PC1, and PC2 for RareMetalWorker analysis
Discovery [ExomeChip]	European [Finnish]	Prevalence, Prediction and Prevention of Diabetes (PPP)-Botnia study	Isomaa, B. et al. A family history of diabetes is associated with reduced physical fitness in the Prevalence, Prediction and Prevention of Diabetes (PPP)-Botnia study. Diabetologia. 2010 Aug;53(8):1709-13.	20454776	- Population-based cohort - T1D, T2D cases or fasting plasma glucose concentration ≥7 mmol/l, pregnant individuals, and samples with non-fasting blood excluded	Illumina HumanExome-12v1.1	>99%	-call rate ≤99% - heterozygosity 4SD of mean - gender discordance - GWAS discordance - genotyping platform fingerprint discordance - population outliers	>99%	- genotyping cluster checks within batches, outliers removed. - exact HWE $<10^{-4}$	Birdseed with cluster filter	- age, age2, and BMI for EMMAX-analysis - age, age2, sex, BMI, PC1, PC2, PC3, and PC4 for RareMetalWorker analysis
Discovery [ExomeSeq]	African American	Jackson Heart Study (AJ)	Taylor, H. A. et al. Toward resolution of cardiovascular health disparities in African Americans: design and methods of the Jackson Heart Study. Ethn Dis 15, S6–4 (2005)	16320381	- No T2D by ADA 2004 definition, fasting plasma glucose <100 mg/dl, and HbA1c <6% at each of two exams - Controls were matched to cases in a two-stage approach: 1. Strong matches (greedy algorithm): age > 50, sex match, BMI within 1 unit, and age within 5 years (N=457 matched pairs) 2. Closest available matches: sex match and BMI > 25; for females, BMI within 5 units and age within 20 years; for males, BMI within 8 units and age within 25 years (N=117 matched pairs)	- Agilent Truseq capture reagents, and individually -barcoded samples sequenced on Illumina HiSeq2000	- poor quality samples removed on the basis of multiple metrics: array genotype concordance (where available), mean heterozygosity and homozygosity, high singleton counts for samples, Variant Quality Score Recalibration (VSQR) for SNVs, and hard filtering for INDELs after genotype calling with GATK - using autosomal variants that passed extended QC and with MAF>1% in all ancestry groups for trans-ethnic kinship analyses, compute identity-by-state (IBS) between each pair of samples on the basis of independent variants (trans-ethnic r2<0.05) and constructed axes of genetic variation through princip'al components analysis implemented in EIGENSTRAT to identify ethnic outliers - identified duplicates on the basis of IBS, and excluded the sample from each pair with lowest call rate and/or mismatch with external information.	- sequence reads from all exome sequenced samples processed jointly and aligned to the reference genome (hg19) with Picard (http://picard.sourceforge.net) - polymorphic sites and genotypes called with GATK (https://www.broadinstitute.org/gatk/) - poor quality samples and variants removed on the basis of multiple metrics: array genotype concordance (where available), mean heterozygosity and homozygosity, high singleton counts for samples, Variant Quality Score Recalibration (VSQR) for SNVs, and hard filtering for INDELs. - within each ancestry group (African-American, East-Asian, European, Hispanic, and South-Asian), extended QC further excluded variants on the basis of call rate (<90% in any study in ancestry group), deviation from Hardy-Weinberg equilibrium (exact p<10-6, females only for X chromosome, in any study in ancestry group) or differential call rate between T2D cases and controls (p<10-4, all studies combined across ancestry group)	'- within each study, age, sex, BMI, and other study-specific covariates for EMMAX-analysis	
Discovery [ExomeSeq]	African American	Wake Forest School of Medicine Study (AW)	Palmer, N. D. et al. A genome-wide association search for type 2 diabetes genes in African Americans. PLoS One 7:e29202. (2012)	22238593	- No current diagnosis of diabetes or renal disease - Individuals recruited from community and internal medicine clinics							
Discovery [ExomeSeq]	East Asian [Korean]	Korea Association Research Project (EK)	Cho, Y. S. et al. A large-scale genome-wide association study of Asian populations uncovers genetic factors influencing eight quantitative traits. Nat. Genet. 41, 527–534 (2009)	19396169	- No past history of diabetes - No anti-diabetic medication - Fasting plasma glucose <5.6 mmol/l and plasma glucose 2 hours after ingestion of 75g oral glucose load <7.8 mmol/l at both baseline and follow up timepoints - Older subjects with normal glucose prioritized							
Discovery [ExomeSeq]	East Asian [Singapore Chinese]	Singapore Diabetes Cohort Study and Singapore Prospective Study Program (ES)	Sim, X. et al. Transferability of type 2 diabetes implicated loci in multi-ethnic cohorts from Southeast Asia. PLoS Genet. 7(4), e1001363 (2011)	21490949	- Fasting blood glucose <6 mmol/l - No personal history of diabetes - No anti-diabetic medication - Older controls preferentially selected							
Discovery [ExomeSeq]	European [Ashkenazim]	Ashkenazi (UA)	Atzmon, G. et al. Lipoprotein genotype and conserved pathway for exceptional longevity in humans. PLoS Biol. 4(4), e113 (2006); Atzmon, G. et al. Evolution in health and medicine Sackler colloquium: Genetic variation in human telomerase is associated with telomere length in Ashkenazi centenarians. Proc Natl Acad Sci U S A. 107 (Suppl 1), 1710-1717 (2010); Permutt, M.A. et al. A genome scan for type 2 diabetes susceptibility loci in a genetically isolated population. Diabetes 50(3), 681-685 (2001); Blech et al. Predicting diabetic nephropathy using a multifactorial genetic model. PLoS One 6(4), e18743 (2011)	16602826; 19915151; 11246891; 21533139	- Fasting blood glucose <7 mmol/l - No personal history of diabetes - No anti-diabetic medications							
Discovery [ExomeSeq]	European [Finnish]	Metabolic Syndrome in Men Study (METSIM)	Stancakova, A. et al. Changes in insulin sensitivity and insulin release in relation to glycemia and glucose tolerance in 6,414 Finnish men. Diabetes 58, 1212–1221 (2009)	19223598	- Normal glucose tolerance at baseline and follow-up visits - Prioritized samples with no family history of diabetes and meeting strict NGT criteria: fasting glucose <5.6 mmol/l and 2 hour post-challenge glucose <7.8 mmol/l - Additional samples selected with fasting glucose <6.1 mmol/l and 2 hour post-challenge glucose <7.8 mmol/l - Unrelated samples - Older controls preferentially selected							
Discovery [ExomeSeq]	European [Finnish]	Finland-United States Investigation of NIDDM Genetics (FUSION) Study	Valle, T. et al. Mapping genes for NIDDM. Design of the Finland-United States Investigation of NIDDM Genetics (FUSION) Study. Diabetes Care 21(6), 949-958 (1998); Scott, L. et al. A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. Science 316(5829), 1341-1345 (2007)	9614613; 17463248	- Unrelated controls with normal glucose tolerance (NGT) based on WHO (1999) definitions: fasting plasma glucose <6.1 mM and 2 hour postload glucose during an OGTT <7.8 mM - Frequency matched to cases by birth province; BMI ≥18.5kg/m^2; age ≤80 - Within each birth province, prioritized samples from stage 2 replication with highest values for age + 2*BMI							
Discovery [ExomeSeq]	European [German]	KORA-gen	Wichmann, H. E., Gieger, C. and Illig, T. KORA-gen–resource for population genetics, controls and a broad spectrum of disease phenotypes. Gesundheitswesen 67 Suppl 1, 26–30 (2005)	16032514	- Controls selected from KORA F4 - All controls are normal glucose tolerant: fasting glucose level <6.1 mmol/l and two hour glucose level after oral glucose tolerance test <7.8 mmol/l - Controls are either >60 years of age with BMI >32 or over 65 years of age with BMI >31							
Discovery [ExomeSeq]	European [UK]	UKT2D Consortium	Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 447, 661–78 (2007); Voight, B.F. et al. Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. Nat. Genet. 42, 579–589 (2010); Spector, T.D. and Williams, F.M. The UK Adult Twin Registry (TwinsUK). Twin Res. Hum. Genet. 9, 899–906 (2006)	17554300; 20581827; 23088889	- Unrelated samples selected as controls from the Twins UK study - A twin pair was considered for selection if there was no recorded family history of diabetes, neither twin was ever recorded as impaired glucose tolerant (defined as fasting glucose >6.1mmol/L in any reading), there were available quantitative trait and genetic (GWAs) data, and no evidence of admixture in MDS analysis of GWAs data - From set of qualifying twin pairs, the best control twin was selected from each pair with the lowest ratio of fasting glucose level to BMI across all readings, and further prioritization of the qualifying unrelated samples involved selecting samples that had the lowest fasting glucose to (BMI * age) ratios - Top two principal components were used to perform pairwise sample matching between cases and possible controls, and the best control for each case was selected							

Stage	Ancestry	Study	Citation(s)	PubMed ID(s)	Sample Ascertainment	Genotyping array	Call rate	Exclusion criteria	Call rate	Filtering criteria	Calling algorithm	Association covariates
Discovery [ExomeSeq]	European [Finnish, Swedish]	Malmo-Botnia Study	Groop, L. et al. Metabolic consequences of a family history of NIDDM (the Botnia study): evidence for sex-specific parental effects. Diabetes 45, 1585–93 (1996); Lindholm, E., Agardh, E., Tuomi, T., Groop, L. & Agardh, C. D. Classifying diabetes according to the new WHO clinical stages. Eur. J. of Epid. 17, 983–9 (2001); Parker, A. et al. A gene conferring susceptibility to type 2 diabetes in conjunction with obesity is located on chromosome 18p11. Diabetes 50, 675–80 (2001); Berglund G. et al. The Malmo Diet and Cancer Study. Design and feasibility. J Intern Med. Jan;233(1):45-51 (1993). Berglund, G. et al. Long-term outcome of the Malmo Preventive Project: Mortality and cardiovascular morbidity. J. of Intern. Med. 247, 19–29 (2000); Lyssenko, V. et al. Clinical risk factors, DNA variants, and the development of type 2 diabetes. NEJM 359, 2220–32 (2008); Isomaa, B. et al. A family history of diabetes is associated with reduced physical fitness in the Prevalence, Prediction and Prevention of Diabetes (PPP)-Botnia study. Diabetologia. Aug;53(8):1709-13 (2010); Bøg-Hansen, E. et al. Risk factor clustering in patients with hypertension and non-insulin-dependent diabetes mellitus. The Skaraborg Hypertension Project. J Intern Med. Mar;243(3):223-32 (1998).	8866565; 12380709; 11246890; 8429286; 10672127; 19020324; 20454776; 9627160	- Controls selected from the extreme of a liability score distribution, based upon gender, age and BMI at last follow-up visit; only BMI and gender used to construct scores for Malmö study - Eligible controls limited to individuals above 35 years of age at follow-up and with a BMI between 20 and 40 - To match for ethnicity, equal numbers of controls were selected from the Botnia and Malmö studies						
Discovery [ExomeSeq]	Hispanic	San Antonio Family Heart Study, San Antonio Family Diabetes/ Gallbladder Study, Veterans Administration Genetic Epidemiology Study, and the Investigation of Nephropathy and Diabetes Study family component (HA)	Mitchell, B. D. et al. Genetic and environmental contributions to cardiovascular risk factors in Mexican Americans. The San Antonio Family Heart Study. Circulation 94, 2159–2170 (1996); Hunt, K. J. et al. Genome-wide linkage analyses of type 2 diabetes in Mexican Americans: the San Antonio Family Diabetes/Gallbladder Study. Diabetes 54, 2655–2662 (2005); Coletta, D. K. et al. Genome-wide linkage scan for genes influencing plasma triglyceride levels in the Veterans Administration Genetic Epidemiology Study. Diabetes 58, 279–284 (2009); Knowler, W. C. et al. The Family Investigation of Nephropathy and Diabetes (FIND): design and methods. J. Diabetes Complicat. 19, 1–9 (2005)	8901667; 16123354; 18931038; 15642484	- Fasting glucose <126 mg/dl at each visit - If OGTT performed, 2 hour glucose must be <200mg/dl - No self-reported antidiabetic therapy at any visit, including oral agents or insulin prescribed as a result of physician-diagnosed diabetes - Prioritize samples with strict NGT with no family history first, then NGT in two visits, followed by oldest age						
Discovery [ExomeSeq]	Hispanic	Starr County, Texas (HS)	Hanis, C. L. et al. Diabetes among Mexican Americans in Starr County, Texas. Am. J. Epidemiol. 118, 659–672 (1983); Below JE, et al. Genome-wide association and meta-analysis in populations from Starr County, Texas and Mexico City identify type 2 diabetes susceptibility loci and enrichment for eQTLs in top signals. Diabetologia 54, 2047-2055 (2011)	6637993 21573907	- Controls ascertained from epidemiologically represented sample of individuals in Starr County, TX - Individuals with known diagnosis of diabetes excluded - Impaired glucose tolerant and impaired fasting glucose controls retained due to the age difference between cases and controls (controls are younger on average) and to allow sufficient sample size						
Discovery [ExomeSeq]	South Asian [UK Indian Asians]	London Life Sciences Population Study (SL)	Chambers, J.C. et al. Genome-wide association study identifies variants in TMPRSS6 associated with hemoglobin levels. Nat. Genet. 41, 1170-1172 (2009); Chambers, J.C. et al. Common genetic variation near melatonin receptor MTNR1B contributes to raised plasma glucose and increased risk of type 2 diabetes among Indian Asians and European Caucasians. Diabetes 58, 2703-2708 (2009); van der Harst, P. et al. Seventy-five genetic loci influencing the human red blood cell. Nature 492, 369-375 (2012)	19820698; 19651812; 23222517	- No previous history of diabetes - No anti-diabetic medication - Fasting plasma glucose <6.0 mmol/L						
Discovery [ExomeSeq]	South Asian [Singapore Indians]	Singapore Indian Eye Study (SS)	Sim, X. et al. Transferability of type 2 diabetes implicated loci in multi-ethnic cohorts from Southeast Asia. PLoS Genet. 7(4), e1001363 (2011)	21490949	- HbA1c <6% - No personal history of diabetes - Not taking antidiabetes medication - Older controls preferentially selected							
Replication [Array]	European [Finnish]	The Cardiovascular Risk in Young Finns Study (YFS)	Raitakari, O.T. et al. Cohort profile: the cardiovascular risk in Young Finns Study. Int. J. Epidemiol., 37, 1220–1226 (2008)	18263651	- Population-based survey - T1D and T2D (fasting plasma glucose concentration ≥7.0 mmol/l or on diabetes medication) cases excluded - Further excluded pregnant individuals	Custom generated Illumina 670K array	≥95%	- excessive heterozygosity - closely related individuals - sex discrepancy	≥95%	- call rate <95%	See methods	- age, sex, BMI, PCs 1-10
Replication [Array]	European [Finnish]	Helsinki Birth Cohort Study (HBCS)	Eriksson, J.G. Epidemiology, genes and the environment: lessons learned from the Helsinki Birth Cohort Study. J. Intern. Med., 261, 418–425 (2007)	17444881	- Birth cohort - T1D and T2D (fasting plasma glucose concentration ≥7.0 mmol/l) cases excluded	Custom generated Illumina 670K array	≥95%	- excessive heterozygosity - closely related individuals - sex discrepancy	≥95%	- call rate <95%	See methods	- age, sex, BMI, PCs 1-10
Replication [Array]	European [Finnish]	The Health 2000 GenMets Study (GenMets)	Perttila, J. et al. OSBPL10, a novel candidate gene for high triglyceride trait in dyslipidemic Finnish subjects, regulates cellular lipid metabolism. J. Mol. Med., 87, 825–835 (2009)	19554302	- Population-based survey - T1D and T2D (fasting plasma glucose concentration ≥7.0 mmol/l or on diabetes medication) cases excluded	Illumina 610K array	≥95%	- excessive heterozygosity - closely related individuals - sex discrepancy	≥95%	- call rate <95%	See methods	- age, sex, BMI, PCs 1-10
Replication [Array]	European [Finnish]	The National FINRISK Study 1997 and 2002 (FINRISK 1997 and 2002)	Vartiainen E. et al. Thirty-five-year trends in cardiovascular risk factors in Finland. Int J Epidemiol., 39, 504-518 (2010)	19959603	- Population-based survey - T1D and T2D (fasting plasma glucose concentration ≥7.0 mmol/l or on diabetes medication) cases excluded - Non-fasting individuals excluded	Illumina HumanCoreExome-12v1-0	≥95%	- excessive heterozygosity - closely related individuals - sex discrepancy	≥95%	- call rate <95%	See methods	- age, sex, BMI, PCs 1-10

Supplementary Table 1B: Sample characteristics of the studies contributing to the analysis.

	Ancestry	European [Finnish]	European [Finnish]	European [Finnish]	European [Finnish]	European [Finnish]	European [Finnish]	European [Danish]	European [Danish]	European [Danish]	European [UK]	European [UK]	European [UK]	European [Swedish]	European [Finnish]
	Study	FIN-D2D 2007	The Finnish Diabetes Prevention Study (DPS)	The Dose Responses to Exercise Training (DR's EXTRA) Study	National FINRISK 2007 Study (FINRISK 2007)	Finland-United States Investigation of NIDDM Genetics (FUSION) Study	Metabolic Syndrome in Men (METSIM)	Health2006	Inter99	Vejle Biobank	Genetics of Diabetes Audit and Research Tayside (GoDARTS)	Twins UK	Oxford BioBank (OBB)	Pivus and Ulsam	Prevalence, Prediction and Prevention of Diabetes (PPP)-Botnia study
Single variant analysis (relateds)	# available	2132	328	622	549	1414	6599	3371	5546	439	802	701	4513	1859	4533
	Mean fasting glucose (SD), mmol/l	F: 5.79 (0.45) M: 6.04 (0.43)	F: 5.86 (0.59) M: 5.93 (0.54)	F: 5.45 (0.47) M: 5.68 (0.48)	F: 5.58 (0.30) M: 5.71 (0.27)	F: 5.22 (0.46) M: 5.37 (0.49)	F: 0 (0) M: 5.72 (0.48)	F: 5.31 (0.52) M: 5.58 (0.51)	F: 5.28 (0.47) M: 5.58 (0.48)	F: 5.07 (0.32) M: 5.19 (0.3)	F: 4.81 (0.45) M: 4.98 (0.48)	F: 4.71 (0.59) M: 4.88 (0.64)	F: 5.11 (0.41) M: 5.41 (0.45)	F: 4.96 (0.56) M: 5.26 (0.58)	F: 5.24 (0.54) M: 5.29 (0.56)
	# Females (%)	1154 (54.13)	224 (68.29)	448 (72.03)	328 (59.74)	653 (46.18)	0 (0.00)	1459 (43.3)	2626 (47.3)	147 (33.5)	369 (46.01)	582 (83.02)	2474 (54.82)	444 (23.88)	2448 (0.54)
	Mean age (SD), years	F: 58.76 (8.23) M: 59.49 (8.44)	F: 55.06 (7.19) M: 56.10 (7.16)	F: 66.29 (5.22) M: 67.90 (6.38)	F: 63.97 (7.53) M: 65.64 (6.49)	F: 62.65 (7.70) M: 60.40 (8.11)	F: 0 (0) M: 57.40 (7.24)	F: 47.89 (12.32) M: 48.77 (12.14)	F: 45.45 (7.91) M: 45.83 (7.82)	F: 55 (12.5) M: 62.36 (10.21)	F: 57.07 (11.09) M: 58.95 (10.80)	F: 44.13 (12.98) M: 45.18 (12.06)	F: 41.57 (5.98) M: 42.04 (5.62)	F: 70.26 (0.15) M: 70.72 (0.67)	F: 48.73 (15.67) M: 48.43 (15.56)
	Mean BMI (SD), kg/m^2	F: 27.00 (4.96) M: 26.81 (3.74)	F: 31.78 (4.88) M: 29.54 (3.46)	F: 27.23 (4.66) M: 27.25 (3.59)	F: 30.08 (4.91) M: 28.59 (3.85)	F: 26.52 (4.24) M: 26.56 (3.49)	F: 0 (0) M: 26.89 (3.83)	F: 25.11 (4.7) M: 26.37 (3.84)	F: 25.46 (4.77) M: 26.58 (3.86)	F: 21.75 (1.92) M: 23.05 (1.47)	F: 26.61 (4.78) M: 27.59 (4.05)	F: 23.75 (4.01) M: 23.81 (3.85)	F: 25.44 (4.80) M: 26.61 (4.00)	F: 26.92 (4.70) M: 26.17 (3.37)	F: 25.85 (4.66) M: 26.66 (3.75)
	# available (with BMI)	2128	328	622	549	1414	6597	3369	5542	439	801	701	4513	1857	4532
	Mean fasting glucose (SD), mmol/l	F: 5.79 (0.45) M: 6.04 (0.42)	F: 5.86 (0.59) M: 5.93 (0.54)	F: 5.45 (0.47) M: 5.68 (0.48)	F: 5.58 (0.30) M: 5.71 (0.27)	F: 5.22 (0.46) M: 5.37 (0.49)	F: 0 (0) M: 5.72 (0.48)	F: 5.31 (0.52) M: 5.58 (0.51)	F: 5.28 (0.47) M: 5.58 (0.48)	F: 5.07 (0.32) M: 5.19 (0.3)	F: 4.81 (0.45) M: 4.98 (0.48)	F: 4.71 (0.59) M: 4.88 (0.64)	F: 5.11 (0.41) M: 5.41 (0.45)	F: 4.96 (0.56) M: 5.26 (0.58)	F: 5.24 (0.54) M: 5.29 (0.56)
	# Females (%)	1152 (54.14)	224 (68.29)	448 (72.03)	328 (59.74)	653 (46.18)	0 (0.00)	1459 (43.3)	2625 (47.4)	147 (33.5)	368 (45.94)	582 (83.02)	2474 (54.82)	444 (23.91)	2447 (0.54)
	Mean age (SD), years	F: 58.77 (8.23) M: 59.47 (8.43)	F: 55.06 (7.19) M: 56.10 (7.16)	F: 66.29 (5.22) M: 67.90 (6.38)	F: 63.97 (7.53) M: 65.64 (6.49)	F: 62.65 (7.70) M: 60.40 (8.11)	F: 0 (0) M: 57.40 (7.24)	F: 47.9 (12.31) M: 48.77 (12.14)	F: 45.46 (7.91) M: 45.84 (7.82)	F: 55 (12.5) M: 62.36 (10.21)	F: 57.04 (11.09) M: 58.95 (10.80)	F: 44.13 (12.98) M: 45.18 (12.06)	F: 41.57 (5.98) M: 42.04 (5.62)	F: 70.26 (0.15) M: 70.72 (0.67)	F: 48.73 (15.67) M: 48.43 (15.56)
	Mean BMI (SD), kg/m^2	F: 27.00 (4.96) M: 26.81 (3.74)	F: 31.78 (4.88) M: 29.54 (3.46)	F: 27.23 (4.66) M: 27.25 (3.59)	F: 30.08 (4.91) M: 28.59 (3.85)	F: 26.52 (4.24) M: 26.56 (3.49)	F: 0 (0) M: 26.89 (3.83)	F: 25.11 (4.7) M: 26.37 (3.84)	F: 25.46 (4.77) M: 26.58 (3.86)	F: 21.75 (1.92) M: 23.05 (1.47)	F: 26.68 (4.78) M: 27.59 (4.05)	F: 23.75 (4.01) M: 23.81 (3.85)	F: 25.44 (4.80) M: 26.61 (4.00)	F: 26.92 (4.70) M: 26.21 (3.37)	F: 25.85 (4.66) M: 26.66 (3.75)
	# available	2111	306	657	548	1342	6596	3370	5337	0	244	0	4136	1853	4492
	Mean fasting insulin (SD), pmol/l	F: 40.90 (20.63) M: 43.84 (26.76)	F: 82.99 (41.26) M: 84.39 (42.06)	F: 41.66 (25.96) M: 44.82 (32.97)	F: 40.74 (21.64) M: 37.02 (19.73)	F: 52.18 (33.05) M: 49.27 (30.38)	F: 0 (0) M: 50.70 (36.47)	F: 37.98 (24.96) M: 40.78 (28.24)	F: 38.73 (25.22) M: 42.59 (27.82)	NA NA	F: 64.15 (31.44) M: 80.97 (55.91)	NA NA	F: 94.71 (39.61) M: 105.54 (51.97)	F: 58.00 (35.62) M: 77.45 (45.29)	F: 8.46 (6.6) M: 9.16 (7.5)
	# Females (%)	1141 (54.05)	214 (69.93)	472 (71.84)	327 (59.67)	632 (47.09)	0 (0.00)	1458 (43.3)	2538 (47.6)	NA	127 (52.05)	NA	2240 (54.16)	445 (24.02)	2426 (0.54)
	Mean age (SD), years	F: 58.75 (8.23) M: 59.49 (8.46)	F: 54.98 (7.09) M: 55.77 (7.27)	F: 66.33 (5.21) M: 67.89 (6.34)	F: 64.02 (7.48) M: 65.64 (6.49)	F: 62.71 (7.80) M: 60.70 (8.11)	F: 0 (0) M: 57.40 (7.24)	F: 47.89 (12.32) M: 48.78 (12.14)	F: 45.4 (7.92) M: 45.82 (7.8)	NA NA	F: 56.40 (10.82) M: 60.04 (9.25)	NA NA	F: 41.65 (5.99) M: 42.11 (5.63)	F: 70.26 (0.15) M: 70.72 (0.67)	F: 48.69 (15.68) M: 48.39 (15.53)
	Mean BMI (SD), kg/m^2	F: 27.01 (4.97) M: 26.79 (3.73)	F: 31.92 (4.89) M: 29.64 (3.47)	F: 27.19 (4.66) M: 27.10 (3.58)	F: 30.05 (4.89) M: 28.59 (3.85)	F: 26.52 (4.23) M: 26.51 (3.42)	F: 0 (0) M: 26.89 (3.83)	F: 25.11 (4.7) M: 26.37 (3.83)	F: 25.5 (4.79) M: 26.58 (3.87)	NA NA	F: 25.66 (4.29) M: 27.36 (3.02)	NA NA	F: 25.46 (4.75) M: 26.58 (3.97)	F: 26.92 (4.69) M: 26.18 (3.38)	F: 25.85 (4.66) M: 26.67 (3.759)
	# available (with BMI)	2107	306	657	548	1342	6594	3368	5333	0	244	NA	4136	1851	4491
	Mean fasting insulin (SD), pmol/l	F: 40.91 (20.65) M: 43.79 (26.74)	F: 82.99 (41.26) M: 84.39 (42.06)	F: 41.66 (25.96) M: 44.82 (32.97)	F: 40.74 (21.64) M: 37.02 (19.73)	F: 52.18 (33.05) M: 49.27 (30.38)	F: 0 (0) M: 50.70 (36.47)	F: 37.96 (24.96) M: 40.78 (28.24)	F: 38.7 (25.2) M: 42.6 (27.82)	NA NA	F: 64.15 (31.44) M: 80.97 (55.91)	NA NA	F: 94.71 (39.61) M: 105.54 (51.97)	F: 58.00 (35.62) M: 77.40 (45.19)	F: 8.46 (6.6) M: 9.16 (7.5)
	# Females (%)	1139 (54.06)	214 (69.93)	472 (71.84)	327 (59.67)	632 (47.09)	0 (0.00)	1458 (43.3)	2537 (47.6)	NA	127 (52.05)	NA	2240 (54.16)	445 (24.04)	2425 (0.54)
	Mean age (SD), years	F: 58.77 (8.23) M: 59.47 (8.45)	F: 54.98 (7.09) M: 55.77 (7.27)	F: 66.33 (5.21) M: 67.89 (6.34)	F: 64.02 (7.48) M: 65.64 (6.49)	F: 62.71 (7.80) M: 60.70 (8.11)	F: 0 (0) M: 57.40 (7.24)	F: 47.9 (12.31) M: 48.78 (12.14)	F: 45.41 (7.92) M: 45.82 (7.8)	NA NA	F: 56.40 (10.82) M: 60.04 (9.25)	NA NA	F: 41.65 (5.99) M: 42.11 (5.63)	F: 70.26 (0.15) M: 70.72 (0.67)	F: 48.69 (15.69) M: 48.39 (15.53)
	Mean BMI (SD), kg/m^2	F: 27.01 (4.97) M: 26.79 (3.73)	F: 31.92 (4.89) M: 29.64 (3.47)	F: 27.19 (4.66) M: 27.10 (3.58)	F: 30.05 (4.89) M: 28.59 (3.85)	F: 26.52 (4.23) M: 26.51 (3.42)	F: 0 (0) M: 26.89 (3.83)	F: 25.11 (4.7) M: 26.37 (3.83)	F: 25.5 (4.79) M: 26.58 (3.87)	NA NA	F: 25.66 (4.29) M: 27.36 (3.02)	NA NA	F: 25.46 (4.75) M: 26.58 (3.97)	F: 26.92 (4.69) M: 26.22 (3.38)	F: 25.85 (4.66) M: 26.67 (3.759)
Gene-level analysis (unrelateds)	# available	2132	328	622	549	1414	6599	3159	5481	431	801	697	4442	1804	4206
	Mean fasting glucose (SD), mmol/l	F: 5.79 (0.45) M: 6.04 (0.43)	F: 5.86 (0.59) M: 5.93 (0.54)	F: 5.45 (0.47) M: 5.68 (0.48)	F: 5.58 (0.30) M: 5.71 (0.27)	F: 5.22 (0.46) M: 5.37 (0.49)	F: 0 (0) M: 5.72 (0.48)	F: 5.31 (0.51) M: 5.58 (0.5)	F: 5.28 (0.47) M: 5.58 (0.48)	F: 5.08 (0.32) M: 5.19 (0.3)	F: 4.81 (0.45) M: 4.98 (0.48)	F: 4.71 (0.59) M: 4.87 (0.64)	F: 5.11 (0.41) M: 5.41 (0.45)	F: 4.96 (0.56) M: 5.26 (0.58)	F: 5.24 (0.55) M: 5.29 (0.56)
	# Females (%)	1154 (54.13)	224 (68.29)	448 (72.03)	328 (59.74)	653 (46.18)	0 (0.00)	1371 (43.4)	2593 (47.3)	145 (33.6)	369 (46.07)	580 (83.21)	2430 (54.71)	437 (24.22)	2302 (0.55)
	Mean age (SD), years	F: 58.76 (8.23) M: 59.49 (8.44)	F: 55.06 (7.19) M: 56.10 (7.16)	F: 66.29 (5.22) M: 67.90 (6.38)	F: 63.97 (7.53) M: 65.64 (6.49)	F: 62.65 (7.70) M: 60.40 (8.11)	F: 0 (0) M: 57.40 (7.24)	F: 48.03 (12.17) M: 48.87 (12.06)	F: 45.46 (7.91) M: 45.85 (7.81)	F: 55.05 (12.61) M: 62.29 (10.26)	F: 57.07 (11.09) M: 58.92 (10.80)	F: 44.19 (12.95) M: 45.08 (12.12)	F: 41.57 (5.99) M: 42.09 (5.61)	F: 70.26 (0.15) M: 70.72 (0.67)	F: 48.7 (15.53) M: 48.36 (15.58)
	Mean BMI (SD), kg/m^2	F: 27.00 (4.96) M: 26.81 (3.74)	F: 31.78 (4.88) M: 29.54 (3.46)	F: 27.23 (4.66) M: 27.25 (3.59)	F: 30.08 (4.91) M: 28.59 (3.85)	F: 26.52 (4.24) M: 26.56 (3.49)	F: 0 (0) M: 26.89 (3.83)	F: 25.13 (4.75) M: 26.36 (3.85)	F: 25.47 (4.77) M: 26.58 (3.84)	F: 21.73 (1.93) M: 23.03 (1.47)	F: 26.61 (4.78) M: 27.60 (4.05)	F: 23.76 (4.01) M: 23.79 (3.82)	F: 25.43 (4.81) M: 26.60 (4.00)	F: 26.93 (4.71) M: 26.18 (3.40)	F: 25.87 (4.64) M: 26.59 (3.749)
	# available (with BMI)	2128	328	622	549	1414	6597	3157	5477	431	800	697	4442	1802	4205
	Mean fasting glucose (SD), mmol/l	F: 5.79 (0.45) M: 6.04 (0.42)	F: 5.86 (0.59) M: 5.93 (0.54)	F: 5.45 (0.47) M: 5.68 (0.48)	F: 5.58 (0.30) M: 5.71 (0.27)	F: 5.22 (0.46) M: 5.37 (0.49)	F: 0 (0) M: 5.72 (0.48)	F: 5.31 (0.51) M: 5.58 (0.5)	F: 5.28 (0.47) M: 5.58 (0.48)	F: 5.08 (0.32) M: 5.19 (0.3)	F: 4.81 (0.45) M: 4.98 (0.48)	F: 4.71 (0.59) M: 4.87 (0.64)	F: 5.11 (0.41) M: 5.41 (0.45)	F: 4.96 (0.56) M: 5.26 (0.58)	F: 5.24 (0.55) M: 5.29 (0.56)
	# Females (%)	1152 (54.14)	224 (68.29)	448 (72.03)	328 (59.74)	653 (46.18)	0 (0.00)	1371 (43.4)	2592 (47.3)	145 (33.6)	368 (46.00)	580 (83.21)	2430 (54.71)	437 (24.25)	2301 (0.55)
	Mean age (SD), years	F: 58.77 (8.23) M: 59.47 (8.43)	F: 55.06 (7.19) M: 56.10 (7.16)	F: 66.29 (5.22) M: 67.90 (6.38)	F: 63.97 (7.53) M: 65.64 (6.49)	F: 62.65 (7.70) M: 60.40 (8.11)	F: 0 (0) M: 57.40 (7.24)	F: 48.04 (12.16) M: 48.87 (12.06)	F: 45.47 (7.91) M: 45.85 (7.81)	F: 55.05 (12.61) M: 62.29 (10.26)	F: 57.04 (11.09) M: 58.92 (10.80)	F: 44.19 (12.95) M: 45.08 (12.12)	F: 41.57 (5.99) M: 42.09 (5.61)	F: 70.26 (0.15) M: 70.72 (0.67)	F: 48.69 (15.53) M: 48.36 (15.58)
	Mean BMI (SD), kg/m^2	F: 27.00 (4.96) M: 26.81 (3.74)	F: 31.78 (4.88) M: 29.54 (3.46)	F: 27.23 (4.66) M: 27.25 (3.59)	F: 30.08 (4.91) M: 28.59 (3.85)	F: 26.52 (4.24) M: 26.56 (3.49)	F: 0 (0) M: 26.89 (3.83)	F: 25.13 (4.75) M: 26.36 (3.85)	F: 25.47 (4.77) M: 26.58 (3.84)	F: 21.73 (1.93) M: 23.03 (1.47)	F: 26.68 (4.78) M: 27.60 (4.05)	F: 23.76 (4.01) M: 23.79 (3.82)	F: 25.43 (4.81) M: 26.60 (4.00)	F: 26.93 (4.71) M: 26.22 (3.40)	F: 25.87 (4.64) M: 26.59 (3.749)
	# available	2111	306	657	548	1342	6596	3158	5272	0	244	0	4075	1799	4171
	Mean fasting insulin (SD), pmol/l	F: 40.90 (20.63) M: 43.84 (26.76)	F: 82.99 (41.26) M: 84.39 (42.06)	F: 41.66 (25.96) M: 44.82 (32.97)	F: 40.74 (21.64) M: 37.02 (19.73)	F: 52.18 (33.05) M: 49.27 (30.38)	F: 0 (0) M: 50.70 (36.47)	F: 37.99 (24.95) M: 40.9 (28.46)	F: 38.68 (25.22) M: 42.59 (27.79)	NA NA	F: 64.15 (31.44) M: 80.97 (55.91)	NA NA	F: 94.69 (39.65) M: 105.55 (51.87)	F: 58.06 (35.82) M: 77.50 (45.65)	F: 8.42 (6.1) M: 9.02 (7.1)
	# Females (%)	1141 (54.05)	214 (69.93)	472 (71.84)	327 (59.67)	632 (47.09)	0 (0.00)	1370 (43.4)	2505 (47.5)	NA	127 (52.05)	NA	2204 (54.09)	438 (24.35)	2283 (0.55)
	Mean age (SD), years	F: 58.75 (8.23) M: 59.49 (8.46)	F: 54.98 (7.09) M: 55.77 (7.27)	F: 66.33 (5.21) M: 67.89 (6.34)	F: 64.02 (7.48) M: 65.64 (6.49)	F: 62.71 (7.80) M: 60.70 (8.11)	F: 0 (0) M: 57.40 (7.24)	F: 48.03 (12.17) M: 40.9 (28.46)	F: 45.41 (7.92) M: 42.59 (27.79)	NA NA	F: 56.40 (10.82) M: 60.04 (9.25)	NA NA	F: 41.64 (6.00) M: 42.15 (5.63)	F: 70.26 (0.15) M: 70.72 (0.67)	F: 48.65 (15.54) M: 48.32 (15.55)
	Mean BMI (SD), kg/m^2	F: 27.01 (4.97) M: 26.79 (3.73)	F: 31.92 (4.89) M: 29.64 (3.47)	F: 27.19 (4.66) M: 27.10 (3.58)	F: 30.05 (4.89) M: 28.59 (3.85)	F: 26.52 (4.23) M: 26.51 (3.42)	F: 0 (0) M: 26.89 (3.83)	F: 25.13 (4.75) M: 26.37 (3.85)	F: 25.51 (4.79) M: 26.58 (3.85)	NA NA	F: 25.66 (4.29) M: 27.36 (3.02)	NA NA	F: 25.45 (4.77) M: 26.58 (3.97)	F: 26.94 (4.71) M: 26.19 (3.41)	F: 25.87 (4.64) M: 26.59 (3.758)
	# available (with BMI)	2107	306	657	548	1342	6594	3156	5268	0	244	NA	4075	1797	4170
	Mean fasting insulin (SD), pmol/l	F: 40.91 (20.65) M: 43.79 (26.74)	F: 82.99 (41.26) M: 84.39 (42.06)	F: 41.66 (25.96) M: 44.82 (32.97)	F: 40.74 (21.64) M: 37.02 (19.73)	F: 52.18 (33.05) M: 49.27 (30.38)	F: 0 (0) M: 50.70 (36.47)	F: 37.97 (24.96) M: 40.9 (28.46)	F: 38.65 (25.2) M: 42.6 (27.8)	NA NA	F: 64.15 (31.44) M: 80.97 (55.91)	NA NA	F: 94.69 (39.65) M: 105.55 (51.87)	F: 58.06 (35.82) M: 77.45 (45.55)	F: 8.42 (6.1) M: 9.02 (7.1)
	# Females (%)	1139 (54.06)	214 (69.93)	472 (71.84)	327 (59.67)	632 (47.09)	0 (0.00)	1370 (43.4)	2504 (47.5)	NA	127 (52.05)	NA	2204 (54.09)	438 (24.37)	2282 (0.55)
	Mean age (SD), years	F: 58.77 (8.23) M: 59.47 (8.45)	F: 54.98 (7.09) M: 55.77 (7.27)	F: 66.33 (5.21) M: 67.89 (6.34)	F: 64.02 (7.48) M: 65.64 (6.49)	F: 62.71 (7.80) M: 60.70 (8.11)	F: 0 (0) M: 57.40 (7.24)	F: 37.97 (24.96) M: 48.87 (12.06)	F: 38.65 (25.2) M: 45.83 (7.78)	NA NA	F: 56.40 (10.82) M: 60.04 (9.25)	NA NA	F: 41.64 (6.00) M: 42.15 (5.63)	F: 70.26 (0.15) M: 70.72 (0.67)	F: 48.65 (15.55) M: 48.32 (15.55)
	Mean BMI (SD), kg/m^2	F: 27.01 (4.97) M: 26.79 (3.73)	F: 31.92 (4.89) M: 29.64 (3.47)	F: 27.19 (4.66) M: 27.10 (3.58)	F: 30.05 (4.89) M: 28.59 (3.85)	F: 26.52 (4.23) M: 26.51 (3.42)	F: 0 (0) M: 26.89 (3.83)	F: 25.13 (4.75) M: 26.37 (3.85)	F: 25.51 (4.79) M: 26.58 (3.85)	NA NA	F: 25.66 (4.29) M: 27.36 (3.02)	NA NA	F: 25.45 (4.77) M: 26.58 (3.97)	F: 26.94 (4.71) M: 26.23 (3.41)	F: 25.87 (4.64) M: 26.59 (3.758)

	Ancestry	African American	African American	East Asian [Korean]	East Asian [Singapore Chinese]	European [Ashkenazim]	European [Finnish]	European [Finnish]	European [German]	European [UK]	European [Finnish, Swedish]	Hispanic	Hispanic	South Asian [UK Indian Asians]	South Asian [Singapore Indians]
	Study	Jackson Heart Study (AJ)	Wake Forest School of Medicine Study (AW)	Korea Association Research Project (EK)	Singapore Diabetes Cohort Study and Singapore Prospective Study Program (ES)	Ashkenazi (UA)	Metabolic Syndrome in Men Study (METSIM)	Finland-United States Investigation of NIDDM Genetics (FUSION) Study	KORA-gen	UKT2D Consortium	Malmo-Botnia Study	San Antonio Family Heart Study, San Antonio Family Diabetes/ Gallbladder Study, Veterans Administration Genetic Epidemiology Study, and the Investigation of Nephropathy and Diabetes Study family component (HA)	Starr County, Texas (HS)	London Life Sciences Population Study (SL)	Singapore Indian Eye Study (SS)
le variant analysis (relateds)	# available	508	NA	556	549	332	498	476	90	320	442	154	699	508	NA
	Mean fasting glucose (SD), mmol/l	F: 4.83 (0.35)	NA	F: 4.42 (0.37)	F: 4.73 (0.43)	F: 4.69 (0.73)	F: 0 (0)	F: 5.28 (0.39)	F: 5.97 (0.38)	F: 4.72 (0.47)	F: 5.10 (0.38)	F: 5.06 (0.56)	F: 4.63 (0.45)	F: 5.09 (0.42)	NA
	# Females (%)	M: 4.84 (0.32) 321 (63.19)	NA NA	M: 4.44 (0.41) 326 (58.63)	M: 4.95 (0.48) 335 (61.02)	M: 4.78 (0.75) 191 (57.53)	M: 5.49 (0.32) 0 (0.00)	M: 5.44 (0.36) 214 (44.96)	M: 6.16 (0.43) 57 (63.33)	M: 4.70 (0.53) 265 (82.81)	M: 5.18 (0.33) 194 (43.89)	M: 5.19 (0.58) 92 (59.74)	M: 4.82 (0.46) 502 (71.82)	M: 5.15 (0.38) 80 (15.75)	NA NA
	Mean age (SD), years	F: 55.81 (11.40) M: 56.49 (11.25)	NA NA	F: 62.91 (3.52) M: 63.72 (3.60)	F: 57.92 (6.45) M: 58.54 (7.61)	F: 80.29 (14.74) M: 76.96 (11.68)	F: 0 (0) M: 54.74 (4.54)	F: 63.78 (7.05) M: 62.19 (7.29)	F: 68.95 (5.43) M: 70.91 (5.79)	F: 60.93 (10.23) M: 59.80 (8.99)	F: 68.04 (8.04) M: 65.57 (8.08)	F: 51.89 (14.28) M: 49.42 (15.24)	F: 39.12 (9.40) M: 39.49 (11.10)	F: 63.51 (8.64) M: 63.14 (9.32)	NA NA
	Mean BMI (SD), kg/m^2	F: 32.98 (6.87) M: 30.01 (5.14)	NA NA	F: 24.19 (3.14) M: 23.10 (2.83)	F: 22.63 (3.41) M: 22.83 (3.20)	F: 24.17 (4.27) M: 26.24 (3.71)	F: 0 (0) M: 25.82 (3.15)	F: 28.49 (4.44) M: 27.48 (3.35)	F: 34.65 (3.51) M: 34.17 (3.42)	F: 31.04 (6.17) M: 28.41 (3.71)	F: 33.70 (4.11) M: 32.25 (3.85)	F: 31.60 (7.34) M: 28.40 (4.54)	F: 30.42 (6.52) M: 29.49 (5.32)	F: 28.26 (4.39) M: 26.96 (3.35)	NA NA
	# available (with BMI)	508	NA	556	548	323	498	476	90	315	442	145	699	508	NA
	Mean fasting glucose (SD), mmol/l	F: 4.83 (0.35)	NA	F: 4.42 (0.37)	F: 4.73 (0.43)	F: 4.70 (0.72)	F: 0 (0)	F: 5.28 (0.39)	F: 5.97 (0.38)	F: 4.73 (0.46)	F: 5.10 (0.38)	F: 5.05 (0.57)	F: 4.63 (0.45)	F: 5.09 (0.42)	NA
	# Females (%)	M: 4.84 (0.32) 321 (63.19)	NA NA	M: 4.44 (0.41) 326 (58.63)	M: 4.95 (0.48) 335 (61.13)	M: 4.79 (0.76) 185 (57.28)	M: 5.49 (0.32) 0 (0.00)	M: 5.44 (0.36) 214 (44.96)	M: 6.16 (0.43) 57 (63.33)	M: 4.70 (0.53) 260 (82.54)	M: 5.18 (0.33) 194 (43.89)	M: 5.15 (0.56) 86 (59.31)	M: 4.82 (0.46) 502 (71.82)	M: 5.15 (0.38) 80 (15.75)	NA NA
	Mean age (SD), years	F: 55.81 (11.40) M: 56.49 (11.25)	NA NA	F: 62.91 (3.52) M: 63.72 (3.60)	F: 57.92 (6.45) M: 58.63 (7.51)	F: 80.33 (14.66) M: 76.51 (11.40)	F: 0 (0) M: 54.74 (4.54)	F: 63.78 (7.05) M: 62.19 (7.29)	F: 68.95 (5.43) M: 70.91 (5.79)	F: 61.17 (10.10) M: 59.80 (8.99)	F: 68.04 (8.04) M: 65.57 (8.08)	F: 52.12 (14.33) M: 48.71 (14.82)	F: 39.12 (9.40) M: 39.49 (11.10)	F: 63.51 (8.64) M: 63.14 (9.32)	NA NA
	Mean BMI (SD), kg/m^2	F: 32.98 (6.87) M: 30.01 (5.14)	NA NA	F: 24.19 (3.14) M: 23.10 (2.83)	F: 22.63 (3.41) M: 22.83 (3.20)	F: 24.17 (4.27) M: 26.24 (3.71)	F: 0 (0) M: 25.82 (3.15)	F: 28.49 (4.44) M: 27.48 (3.35)	F: 34.65 (3.51) M: 34.17 (3.42)	F: 31.04 (6.17) M: 28.41 (3.71)	F: 33.70 (4.11) M: 32.25 (3.85)	F: 31.68 (7.39) M: 28.28 (4.55)	F: 30.42 (6.52) M: 29.49 (5.32)	F: 28.26 (4.39) M: 26.96 (3.35)	NA NA
	# available	507	NA	556	548	117	497	473	90	293	206	154	697	431	NA
	Mean fasting insulin (SD), pmol/l	F: 95.25 (52.59)	NA	F: 47.68 (35.27)	F: 42.55 (26.07)	F: 87.65 (57.64)	F: 0 (0)	F: 52.75 (26.44)	F: 48.41 (54.15)	F: 64.85 (47.30)	F: 62.98 (36.10)	F: 109.11 (103.46)	F: 40.03 (45.84)	F: 59.76 (34.45)	NA
	# Females (%)	M: 83.26 (50.58) 321 (63.31)	NA NA	M: 38.44 (29.40) 326 (58.63)	M: 40.54 (32.61) 334 (60.95)	M: 112.94 (51.51) 62 (52.99)	M: 38.32 (22.83) 0 (0.00)	M: 52.86 (29.22) 214 (45.24)	M: 63.71 (38.14) 57 (63.33)	M: 62.60 (85.24) 242 (82.59)	M: 69.28 (43.87) 96 (46.60)	M: 88.95 (52.39) 92 (59.74)	M: 48.72 (51.91) 500 (71.74)	M: 66.27 (41.37) 59 (13.69)	NA NA
	Mean age (SD), years	F: 55.81 (11.40) M: 56.47 (11.27)	NA NA	F: 62.91 (3.52) M: 63.72 (3.60)	F: 57.94 (6.45) M: 58.54 (7.61)	F: 81.21 (13.43) M: 76.51 (10.20)	F: 0 (0) M: 54.74 (4.55)	F: 63.78 (7.05) M: 62.21 (7.30)	F: 68.95 (5.43) M: 70.91 (5.79)	F: 61.39 (10.26) M: 59.98 (9.25)	F: 66.23 (9.38) M: 63.92 (10.16)	F: 51.89 (14.28) M: 49.42 (15.24)	F: 39.11 (9.41) M: 39.49 (11.10)	F: 62.50 (8.31) M: 62.47 (9.14)	NA NA
	Mean BMI (SD), kg/m^2	F: 32.98 (6.87) M: 30.02 (5.15)	NA NA	F: 24.19 (3.14) M: 23.10 (2.83)	F: 22.60 (3.38) M: 22.83 (3.20)	F: 24.45 (3.85) M: 27.18 (3.24)	F: 0 (0) M: 25.82 (3.15)	F: 28.49 (4.44) M: 27.47 (3.36)	F: 34.65 (3.51) M: 34.17 (3.42)	F: 31.25 (6.27) M: 28.62 (3.72)	F: 31.07 (3.74) M: 29.22 (3.58)	F: 31.60 (7.34) M: 28.40 (4.54)	F: 30.42 (6.53) M: 29.49 (5.32)	F: 28.14 (3.95) M: 26.89 (3.19)	NA NA
	# available (with BMI)	507	NA	556	547	114	497	473	90	293	206	145	697	431	NA
	Mean fasting insulin (SD), pmol/l	F: 95.25 (52.59)	NA	F: 47.68 (35.27)	F: 42.55 (26.07)	F: 86.40 (57.75)	F: 0 (0)	F: 52.75 (26.44)	F: 48.41 (54.15)	F: 64.85 (47.30)	F: 62.98 (36.10)	F: 108.08 (105.91)	F: 40.03 (45.84)	F: 59.76 (34.45)	NA
	# Females (%)	M: 83.26 (50.58) 321 (63.31)	NA NA	M: 38.44 (29.40) 326 (58.63)	M: 39.69 (30.22) 334 (61.06)	M: 112.07 (51.58) 60 (52.63)	M: 38.32 (22.83) 0 (0.00)	M: 52.86 (29.22) 214 (45.24)	M: 63.71 (38.14) 57 (63.33)	M: 62.60 (85.24) 242 (82.59)	M: 69.28 (43.87) 96 (46.60)	M: 89.46 (53.64) 86 (59.31)	M: 48.72 (51.91) 500 (71.74)	M: 66.27 (41.37) 59 (13.69)	NA NA
	Mean age (SD), years	F: 55.81 (11.40) M: 56.47 (11.27)	NA NA	F: 62.91 (3.52) M: 63.72 (3.60)	F: 57.94 (6.45) M: 58.63 (7.51)	F: 81.77 (13.22) M: 76.11 (9.85)	F: 0 (0) M: 54.74 (4.55)	F: 63.78 (7.05) M: 62.21 (7.30)	F: 68.95 (5.43) M: 70.91 (5.79)	F: 61.39 (10.26) M: 59.98 (9.25)	F: 66.23 (9.38) M: 63.92 (10.16)	F: 52.12 (14.33) M: 48.71 (14.82)	F: 39.11 (9.41) M: 39.49 (11.10)	F: 62.50 (8.31) M: 62.47 (9.14)	NA NA
	Mean BMI (SD), kg/m^2	F: 32.98 (6.87) M: 30.02 (5.15)	NA NA	F: 24.19 (3.14) M: 23.10 (2.83)	F: 22.60 (3.38) M: 22.83 (3.20)	F: 24.45 (3.85) M: 27.18 (3.24)	F: 0 (0) M: 25.82 (3.15)	F: 28.49 (4.44) M: 27.47 (3.36)	F: 34.65 (3.51) M: 34.17 (3.42)	F: 31.25 (6.27) M: 28.62 (3.72)	F: 31.07 (3.74) M: 29.22 (3.58)	F: 31.68 (7.39) M: 28.28 (4.55)	F: 30.42 (6.53) M: 29.49 (5.32)	F: 28.14 (3.95) M: 26.89 (3.19)	NA NA
ne-level analysis (unrelateds)	# available	483	NA	555	549	332	496	470	90	320	436	154	699	503	NA
	Mean fasting glucose (SD), mmol/l	F: 4.84 (0.35)	NA	F: 4.42 (0.37)	F: 4.73 (0.43)	F: 4.69 (0.73)	F: 0 (0)	F: 5.28 (0.39)	F: 5.97 (0.38)	F: 4.72 (0.47)	F: 5.10 (0.38)	F: 5.06 (0.56)	F: 4.63 (0.45)	F: 5.08 (0.42)	NA
	# Females (%)	M: 4.84 (0.33) 303 (62.73)	NA NA	M: 4.44 (0.41) 326 (58.74)	M: 4.95 (0.48) 335 (61.02)	M: 4.78 (0.75) 191 (57.53)	M: 5.50 (0.32) 0 (0.00)	M: 5.45 (0.35) 212 (45.11)	M: 6.16 (0.43) 57 (63.33)	M: 4.70 (0.53) 265 (82.81)	M: 5.18 (0.33) 191 (43.81)	M: 5.19 (0.58) 92 (59.74)	M: 4.82 (0.46) 502 (71.82)	M: 5.15 (0.38) 79 (15.71)	NA NA
	Mean age (SD), years	F: 55.64 (11.26) M: 56.63 (11.30)	NA NA	F: 62.91 (3.52) M: 63.70 (3.59)	F: 57.92 (6.45) M: 58.54 (7.61)	F: 80.29 (14.74) M: 76.96 (11.68)	F: 0 (0) M: 54.75 (4.55)	F: 63.75 (7.08) M: 62.14 (7.26)	F: 68.95 (5.43) M: 70.91 (5.79)	F: 60.93 (10.23) M: 59.80 (8.99)	F: 68.13 (7.81) M: 65.78 (7.85)	F: 51.89 (14.28) M: 49.42 (15.24)	F: 39.12 (9.40) M: 39.49 (11.10)	F: 63.38 (8.62) M: 63.14 (9.31)	NA NA
	Mean BMI (SD), kg/m^2	F: 33.03 (6.87) M: 30.06 (5.18)	NA NA	F: 24.19 (3.14) M: 23.11 (2.83)	F: 22.63 (3.41) M: 22.83 (3.20)	F: 24.17 (4.27) M: 26.24 (3.71)	F: 0 (0) M: 25.82 (3.15)	F: 28.51 (4.45) M: 27.44 (3.35)	F: 34.65 (3.51) M: 34.17 (3.42)	F: 31.04 (6.17) M: 28.41 (3.71)	F: 33.71 (4.11) M: 32.24 (3.86)	F: 31.60 (7.34) M: 28.40 (4.54)	F: 30.42 (6.52) M: 29.49 (5.32)	F: 28.23 (4.41) M: 26.96 (3.34)	NA NA
	# available (with BMI)	483	NA	555	548	323	496	470	90	315	436	145	699	503	NA
	Mean fasting glucose (SD), mmol/l	F: 4.84 (0.35)	NA	F: 4.42 (0.37)	F: 4.73 (0.43)	F: 4.70 (0.72)	F: 0 (0)	F: 5.28 (0.39)	F: 5.97 (0.38)	F: 4.73 (0.46)	F: 5.10 (0.38)	F: 5.05 (0.57)	F: 4.63 (0.45)	F: 5.08 (0.42)	NA
	# Females (%)	M: 4.84 (0.33) 303 (62.73)	NA NA	M: 4.44 (0.41) 326 (58.74)	M: 4.95 (0.48) 335 (61.13)	M: 4.79 (0.76) 185 (57.28)	M: 5.50 (0.32) 0 (0.00)	M: 5.45 (0.35) 212 (45.11)	M: 6.16 (0.43) 57 (63.33)	M: 4.70 (0.53) 260 (82.54)	M: 5.18 (0.33) 191 (43.81)	M: 5.15 (0.56) 86 (59.31)	M: 4.82 (0.46) 502 (71.82)	M: 5.15 (0.38) 79 (15.71)	NA NA
	Mean age (SD), years	F: 55.64 (11.26) M: 56.63 (11.30)	NA NA	F: 62.91 (3.52) M: 63.70 (3.59)	F: 57.92 (6.45) M: 58.63 (7.51)	F: 80.33 (14.66) M: 76.51 (11.40)	F: 0 (0) M: 54.75 (4.55)	F: 63.75 (7.08) M: 62.14 (7.26)	F: 68.95 (5.43) M: 70.91 (5.79)	F: 61.17 (10.10) M: 59.80 (8.99)	F: 68.13 (7.81) M: 65.78 (7.85)	F: 52.12 (14.33) M: 48.71 (14.82)	F: 39.12 (9.40) M: 39.49 (11.10)	F: 63.38 (8.62) M: 63.14 (9.31)	NA NA
	Mean BMI (SD), kg/m^2	F: 33.03 (6.87) M: 30.06 (5.18)	NA NA	F: 24.19 (3.14) M: 23.11 (2.83)	F: 22.63 (3.41) M: 22.83 (3.20)	F: 24.17 (4.27) M: 26.24 (3.71)	F: 0 (0) M: 25.82 (3.15)	F: 28.51 (4.45) M: 27.44 (3.35)	F: 34.65 (3.51) M: 34.17 (3.42)	F: 31.04 (6.17) M: 28.41 (3.71)	F: 33.71 (4.11) M: 32.24 (3.86)	F: 31.68 (7.39) M: 28.28 (4.55)	F: 30.42 (6.52) M: 29.49 (5.32)	F: 28.23 (4.41) M: 26.96 (3.34)	NA NA
	# available	482	NA	555	548	117	495	467	90	293	201	154	697	426	NA
	Mean fasting insulin (SD), pmol/l	F: 96.08 (53.19)	NA	F: 47.68 (35.27)	F: 42.55 (26.07)	F: 87.65 (57.64)	F: 0 (0)	F: 52.75 (26.51)	F: 48.41 (54.15)	F: 64.85 (47.30)	F: 61.64 (35.24)	F: 109.11 (103.46)	F: 40.03 (45.84)	F: 59.96 (34.72)	NA
	# Females (%)	M: 83.77 (51.28) 303 (62.86)	NA NA	M: 38.53 (29.43) 326 (58.74)	M: 40.54 (32.61) 334 (60.95)	M: 112.94 (51.51) 62 (52.99)	M: 38.31 (22.83) 0 (0.00)	M: 52.96 (29.30) 212 (45.40)	M: 63.71 (38.14) 57 (63.33)	M: 62.60 (85.24) 242 (82.59)	M: 68.76 (43.75) 94 (46.77)	M: 88.95 (52.39) 92 (59.74)	M: 48.72 (51.91) 500 (71.74)	M: 66.13 (41.51) 58 (13.62)	NA NA
	Mean age (SD), years	F: 55.64 (11.26) M: 56.61 (11.33)	NA NA	F: 62.91 (3.52) M: 63.70 (3.59)	F: 57.94 (6.45) M: 58.54 (7.61)	F: 81.21 (13.43) M: 76.51 (10.20)	F: 0 (0) M: 54.76 (4.55)	F: 63.75 (7.08) M: 62.16 (7.27)	F: 68.95 (5.43) M: 70.91 (5.79)	F: 61.39 (10.26) M: 59.98 (9.25)	F: 66.22 (9.09) M: 64.34 (9.86)	F: 51.89 (14.28) M: 49.42 (15.24)	F: 39.11 (9.41) M: 39.49 (11.10)	F: 62.31 (8.26) M: 62.47 (9.12)	NA NA
	Mean BMI (SD), kg/m^2	F: 33.03 (6.87) M: 30.07 (5.20)	NA NA	F: 24.19 (3.14) M: 23.11 (2.83)	F: 22.60 (3.38) M: 22.83 (3.20)	F: 24.45 (3.85) M: 27.18 (3.24)	F: 0 (0) M: 25.82 (3.15)	F: 28.51 (4.45) M: 27.43 (3.36)	F: 34.65 (3.51) M: 34.17 (3.42)	F: 31.25 (6.27) M: 28.62 (3.72)	F: 31.04 (3.72) M: 29.11 (3.55)	F: 31.60 (7.34) M: 28.40 (4.54)	F: 30.42 (6.53) M: 29.49 (5.32)	F: 28.09 (3.97) M: 26.88 (3.18)	NA NA
	# available (with BMI)	482	NA	555	547	114	495	467	90	293	201	145	697	426	NA
	Mean fasting insulin (SD), pmol/l	F: 96.08 (53.19)	NA	F: 47.68 (35.27)	F: 42.55 (26.07)	F: 86.40 (57.75)	F: 0 (0)	F: 52.75 (26.51)	F: 48.41 (54.15)	F: 64.85 (47.30)	F: 61.64 (35.24)	F: 108.08 (105.91)	F: 40.03 (45.84)	F: 59.96 (34.72)	NA
	# Females (%)	M: 83.77 (51.28) 303 (62.86)	NA NA	M: 38.53 (29.43) 326 (58.74)	M: 39.69 (30.22) 334 (61.06)	M: 112.07 (51.58) 60 (52.63)	M: 38.31 (22.83) 0 (0.00)	M: 52.96 (29.30) 212 (45.40)	M: 63.71 (38.14) 57 (63.33)	M: 62.60 (85.24) 242 (82.59)	M: 68.76 (43.75) 94 (46.77)	M: 89.46 (53.64) 86 (59.31)	M: 48.72 (51.91) 500 (71.74)	M: 66.13 (41.51) 58 (13.62)	NA NA
	Mean age (SD), years	F: 55.64 (11.26) M: 56.61 (11.33)	NA NA	F: 62.91 (3.52) M: 63.70 (3.59)	F: 57.94 (6.45) M: 58.63 (7.51)	F: 81.77 (13.22) M: 76.11 (9.85)	F: 0 (0) M: 54.76 (4.55)	F: 63.75 (7.08) M: 62.16 (7.27)	F: 68.95 (5.43) M: 70.91 (5.79)	F: 61.39 (10.26) M: 59.98 (9.25)	F: 66.22 (9.09) M: 64.34 (9.86)	F: 52.12 (14.33) M: 48.71 (14.82)	F: 39.11 (9.41) M: 39.49 (11.10)	F: 62.31 (8.26) M: 62.47 (9.12)	NA NA
	Mean BMI (SD), kg/m^2	F: 33.03 (6.87) M: 30.07 (5.20)	NA NA	F: 24.19 (3.14) M: 23.11 (2.83)	F: 22.60 (3.38) M: 22.83 (3.20)	F: 24.45 (3.85) M: 27.18 (3.24)	F: 0 (0) M: 25.82 (3.15)	F: 28.51 (4.45) M: 27.43 (3.36)	F: 34.65 (3.51) M: 34.17 (3.42)	F: 31.25 (6.27) M: 28.62 (3.72)	F: 31.04 (3.72) M: 29.11 (3.55)	F: 31.68 (7.39) M: 28.28 (4.55)	F: 30.42 (6.53) M: 29.49 (5.32)	F: 28.09 (3.97) M: 26.88 (3.18)	NA NA

SUPPLEMENTARY TABLE 2

Association results from the discovery phase.

Supplementary Table 2A: Significant ($P < 5 \times 10^{-7}$) and suggestive ($P < 5 \times 10^{-6}$) single variant association results in previously published regions associated with FI levels or FG levels. The published association statistics are shaded in gray. The association results for each region in our analyses are presented in the non-shaded rows.

Insulin

GWAS Loci	Location	rsID	Gene	Consequence	Protein Change	ETH	Allele	Allele Freq	Beta estimate	Standard Error	P value	N
LYPLAL1		rs4846565					G	0.67	0.013		1.8E-09	99014
	1:219644224	rs2605100	NA	NA	NA	1	A	0.31	-0.019	0.0039	4.5E-07	30825
	1:219652033	rs2791552	NA	NA	NA	1	A	0.32	-0.018	0.0039	8.7E-07	30824
GCKR		rs780094					C	0.62	0.015		3.6E-20	96126
	2:27730940	rs1260326	*GCKR*	missense,splice_region	p.L446P	5	T	0.39	-0.021	0.0036	2.2E-10	35380
	2:27741237	rs780094	*GCKR*	intron	NA	1	T	0.37	-0.023	0.0038	6.3E-11	30825
	2:27742603	rs780093	*GCKR*	intron	NA	1	T	0.37	-0.023	0.0038	5.4E-11	30815
	2:27801493	rs1919127	*C2orf16*	missense	p.V685A	5	T	0.73	0.022	0.0047	4.7E-07	26227
	2:27801759	rs1919128	*C2orf16*	missense	p.I774V	5	A	0.73	0.021	0.0040	1.9E-08	35381
	2:27851918	rs3749147	*GPN1*	missense	p.R12K	5	A	0.25	-0.020	0.0044	5.4E-07	30846
GRB14		rs10195252					T	0.60	0.017		1.3E-16	
	2:165540800	rs12328675	*COBLL1*	downstream_gene	NA	1	T	0.89	0.029	0.0058	1.6E-07	30739
	2:165551201	rs7607980	*COBLL1*	missense	p.N939D	4	T	0.88	0.031	0.0056	3.1E-09	34278
	2:165528876	rs7578326	NA	NA	NA	1	T	0.38	-0.019	0.0038	4.2E-07	30824
IRS1		rs2943645					T	0.63	0.019		2.3E-19	99023
	2:227020653	rs7578326	NA	NA	NA	1	A	0.65	0.023	0.0038	5.8E-11	30823
	2:227068080	rs2943634	NA	NA	NA	1	A	0.34	-0.025	0.0038	7.7E-13	30816
	2:227093745	rs2943641	NA	NA	NA	1	T	0.37	-0.028	0.0038	1.4E-15	30825
	2:227100698	rs2972146	NA	NA	NA	1	T	0.63	0.028	0.0038	1.1E-15	30818
	2:227105921	rs2943650	NA	NA	NA	1	T	0.62	0.049	0.0083	3.8E-09	6792
ANKRD55:MAP3K1	rs459193						G	0.73	0.015		1.1E-12	
	5:55806751	rs459193	*AC022431.2.1*	downstream_gene	NA	1	A	0.29	-0.019	0.0040	1.5E-06	30825
GCKR		rs780094					C	0.62	0.03		5.8E-38	118032
	2:27424636	rs1395	*SLC5A6*	missense	p.S481F	5	A	0.69	-0.02	0.0036	4.0E-08	38338
	2:27550967	rs1049817	*GTF3C2*	synonymous	p.P782P	5	A	0.58	-0.02	0.0033	1.4E-07	38339
	2:27711893	rs1260327	*IFT172*	intron	NA	1	A	0.52	-0.02	0.0035	2.9E-09	33231
	2:27730940	rs1260326	*GCKR*	missense,splice_region	p.L446P	5	T	0.37	-0.03	0.0034	3.1E-18	38338
	2:27741237	rs780094	*GCKR*	intron	NA	1	T	0.37	-0.03	0.0037	1.4E-18	33231
	2:27742603	rs780093	*GCKR*	intron	NA	1	T	0.37	-0.03	0.0037	8.0E-18	33221
	2:27801493	rs1919127	*C2orf16*	missense	p.V685A	5	T	0.72	0.02	0.0043	2.6E-07	29085
	2:27801759	rs1919128	*C2orf16*	missense	p.I774V	5	A	0.72	0.02	0.0037	6.0E-10	38339
	2:27851918	rs3749147	*GPN1*	missense	p.R12K	5	A	0.25	-0.02	0.004	7.7E-09	33763
	2:28344285	rs12104449	*BRE*	intron	NA	1	A	0.11	-0.03	0.0056	2.2E-06	33231
	2:27972833	rs4401177	NA	NA	NA	1	A	0.88	0.02	0.0054	3.7E-06	33200
G6PC2		rs560887					C	0.70	0.08		8.7E-218	119169
	2:169763148	rs560887	*G6PC2*	intron	NA	5	T	0.30	-0.07	0.0036	7.9E-87	38339
	2:169763262	rs138726309	*G6PC2*	missense	p.H177Y	1	T	0.01	-0.10	0.0193	7.4E-08	34574
	2:169764141	rs2232323	*G6PC2*	missense	p.Y207S	3	A	0.99	0.13	0.0227	1.7E-09	35227
	2:169764176	rs492594	*G6PC2*	missense	p.V219L	5	C	0.48	0.02	0.0032	1.4E-08	38339
	2:169791438	rs552976	*ABCB11*	intron	NA	1	A	0.35	-0.06	0.0037	5.1E-66	33231
	2:169774071	rs563694	NA	NA	NA	1	A	0.65	0.06	0.0037	4.3E-68	33231
PCSK1		rs4869272					T	0.69			1.0E-15	13,872
	5:95728898	rs6235	*PCSK1*	missense	p.S690T	5	C	0.72	0.02	0.0036	2.1E-09	38339
	5:95728974	rs6234	*PCSK1*	missense	p.Q665E	5	C	0.28	-0.02	0.0036	2.0E-09	38339
	5:95539448	rs4869272	NA	NA	NA	1	T	0.68	0.02	0.0038	8.3E-07	33231
CDKAL1		rs9368222					A	0.28	0.01		1.0E-09	128453
	6:20679709	rs7756992	*CDKAL1*	intron	NA	1	A	0.70	-0.02	0.0038	3.9E-06	33219
GLP1R												
	6:39046794	rs10305492	*GLP1R*	missense	p.A316T	2	A	0.02	-0.07	0.0139	4.5E-07	36218
DGKB:TMEM195	rs2191349						T	0.52	0.03		3.0E-44	
	7:15063833	rs10244051	NA	NA	NA	1	T	0.51	-0.03	0.0035	1.5E-14	33230
	7:15064309	rs2191349	NA	NA	NA	1	T	0.49	0.03	0.0035	1.3E-14	33231

Glucose

GWAS Loci	Location	rsID	Gene	Consequence	Protein Change	ETH		Allele Freq	Beta estimate	Standard Error	P value	N
GCK		rs4607517					A	0.16	0.06		6.5E-92	118500
	7:44183187	rs2971681	*MYL7*	upstream_gene	NA	1	A	0.79	-0.02	0.0044	2.8E-07	33231
	7:44223721	rs730497	*GCK*	intron	NA	1	A	0.14	0.06	0.0052	4.7E-31	33231
	7:44229068	rs1799884	*GCK*	upstream_gene	NA	1	T	0.13	0.06	0.0064	4.9E-21	24042
	7:44231886	rs6975024	*GCK*	upstream_gene	NA	1	T	0.86	-0.06	0.0052	2.2E-31	33228
	7:44235668	rs4607517	*YKT6*	upstream_gene	NA	1	A	0.14	0.06	0.0052	2.2E-31	33231
GRB10		rs6943153					T	0.34	0.02		1.6E-12	131795
	7:50730452	rs2715094	*GRB10*	intron	NA	1	A	0.69	-0.02	0.0039	6.5E-07	33231
	7:50751090	rs10248619	*GRB10*	intron	NA	1	T	0.30	0.02	0.004	8.6E-09	33225
	7:50786663	rs2108349	*GRB10*	intron	NA	1	A	0.61	-0.02	0.0037	5.8E-08	33226
	7:50791579	rs6943153	*GRB10*	intron	NA	1	T	0.39	0.02	0.0037	6.7E-08	33230
	7:50758245	rs933360	NA	NA	NA	1	T	0.68	-0.02	0.0046	1.3E-06	23984
PPP1R3B		rs983309					T	0.12	0.03		6.3E-15	127470
	8:9183358	rs9987289	NA	NA	NA	1	A	0.13	0.03	0.0058	3.8E-07	26841
	8:9183596	rs4841132	NA	NA	NA	1	A	0.13	0.03	0.0054	2.0E-07	33231
	8:9184691	rs6601299	NA	NA	NA	1	T	0.14	0.03	0.0055	3.9E-07	28698
	8:9185146	rs2126259	NA	NA	NA	1	T	0.14	0.02	0.0052	3.4E-06	33230
SLC30A8		rs11558471					A	0.68	0.03		2.6E-11	
	8:118184783	rs13266634	*SLC30A8*	missense	p.R276W	5	T	0.36	-0.02	0.0034	1.6E-11	38338
	8:118185025	rs3802177	*SLC30A8*	3_prime_UTR	NA	1	A	0.36	-0.02	0.0036	2.5E-10	33230
	8:118185733	rs11558471	*SLC30A8*	3_prime_UTR	NA	1	A	0.64	0.02	0.0036	2.1E-10	33231
CDKN2B		rs10811661					T	0.82	0.02		5.6E-18	
	9:22133284	rs10965250	NA	NA	NA	1	A	0.15	-0.03	0.0059	7.9E-07	22658
IKBKAP		rs16913693					T	0.97	0.04		3.5E-11	
	9:111679940	rs17853166	*IKBKAP*	missense	p.S251G	2	T	0.97	0.04	0.0097	3.7E-06	36218
ADRA2A		rs10885122					G	0.87	0.04		2.9E-16	
	10:113022555	rs10885117	NA	NA	NA	1	T	0.91	0.03	0.006	9.5E-07	33211
TCF7L2		rs7903146					C	0.72	-0.02		2.7E-20	127477
	10:114758349	rs7903146	*TCF7L2*	intron	NA	1	T	0.23	0.02	0.0042	4.3E-07	33231
CRY2		rs11605924					A	0.49	0.02		1.0E-14	
	11:45878992	rs7945565	*CRY2*	intron	NA	1	A	0.51	0.02	0.0035	1.8E-10	33230
MADD		rs7944584					A	0.75	0.03		2.0E-18	118741
	11:47270255	rs2167079	*ACP2*	missense	p.R29Q	5	T	0.38	0.02	0.0034	1.9E-07	38338
	11:47286290	rs7120118	*NR1H3*	intron	NA	1	T	0.63	-0.02	0.0037	2.8E-06	33231
	11:47290984	rs1449627	*MADD*	5_prime_UTR	NA	1	T	0.62	-0.02	0.0036	4.6E-06	33231
	11:47298360	rs326214	*MADD*	synonymous	p.E347E	5	A	0.61	-0.02	0.0033	3.8E-07	38339
	11:47336320	rs7944584	*MADD*	intron	NA	1	A	0.77	0.03	0.0043	2.6E-11	33231
	11:47354787	rs1052373	*MYBPC3*	synonymous	p.E1096E	5	T	0.39	0.02	0.0033	1.1E-06	38337
FADS1		rs174550					T	0.64	0.02		1.7E-15	118908
	11:61557803	rs102275	*C11orf10*	intron	NA	1	T	0.62	0.02	0.0036	1.5E-07	33231
	11:61569830	rs174546	*FADS1*	3_prime_UTR	NA	1	T	0.38	-0.02	0.0037	4.1E-07	33231
	11:61570783	rs174547	*FADS1*	intron	NA	5	T	0.62	0.02	0.0034	2.1E-09	38339
	11:61571478	rs174550	*FADS1*	intron	NA	1	T	0.62	0.02	0.0037	3.4E-07	33230
	11:61597972	rs1535	*FADS2*	intron	NA	1	A	0.62	0.02	0.0036	6.1E-07	33230
	11:61609750	rs174583	*FADS2*	intron	NA	1	T	0.38	-0.02	0.0036	3.0E-07	33231
ARAP1		rs11603334					G	0.83	0.02		1.1E-11	
	11:72432985	rs11603334	*ARAP1*	5_prime_UTR	NA	1	A	0.21	-0.02	0.0044	1.5E-08	33231
	11:72433098	rs1552224	*ARAP1*	5_prime_UTR	NA	1	A	0.79	0.02	0.0044	1.2E-08	33230
MTNR1B		rs10830963					G	0.30	0.08		5.8E-175	
	11:92708710	rs10830963	*MTNR1B*	intron	NA	1	C	0.69	-0.09	0.0038	2.8E-118	33230
	11:92651002	rs7950811	NA	NA	NA	1	A	0.05	0.06	0.0087	6.8E-11	33231
	11:92668826	rs3847554	NA	NA	NA	1	T	0.43	0.06	0.0035	1.6E-62	33231
	11:92673828	rs1387153	NA	NA	NA	1	T	0.30	0.07	0.0038	5.6E-76	33231
	11:92691532	rs2166706	NA	NA	NA	1	T	0.60	-0.06	0.0036	5.5E-57	33231
	11:92722761	rs1447352	NA	NA	NA	1	A	0.53	0.04	0.0035	5.3E-31	33214
C2CD4B		rs11071657					A	0.63	0.02		3.6E-08	
	15:62383155	rs4502156	NA	NA	NA	1	T	0.50	0.02	0.0035	1.4E-10	33231
	15:62396389	rs7172432	NA	NA	NA	1	A	0.51	0.02	0.0035	3.8E-11	33231
	15:62404382	rs1436955	NA	NA	NA	1	T	0.28	-0.02	0.0039	1.0E-06	33231
FOXA2		rs6113722					G	0.96	0.35		2.5E-11	123665
	20:39832628	rs17265513	ZHX3	missense	p.N310S	4	T	0.76	-0.02	0.0039	1.4E-07	37233

Supplementary Table 2B: Significant ($P< 5\times10^{-7}$) and suggestive ($P<5\times10^{-6}$) single variant association results that are not in previously published regions. Results are shown for variants with association $P< 5\times10^{-6}$ that fall outside the regions of previously published genetic associations.

	Location	rsID	Gene	Consequence	Protein Change	ETH	Allele	Allele Freq	Effect	Standard Error	P value	N
Glucose	7:2854547	rs116515234	*GNA12*	intron	NA	1	A	0.98	-0.30	0.07	6.7E-07	508
	15:43714320	rs140119148	*TP53BP1*	missense	p.T1278I	1	A	0.002	0.34	0.07	9.0E-07	13286
	1:2535397	rs150660153	*MMEL1*	missense	p.E323Q	2	C	1.00	-0.24	0.05	1.1E-06	17659
	6:43806609	rs881858	*VEGFA*	NA	NA	1	A	0.69	-0.02	0.004	4.1E-06	33231
	19:3754020	rs61731066	*APBA3*	synonymous	p.S282S	4	C	0.02	-0.16	0.03	4.1E-06	4004
	Location	**rsID**	**Gene**	**Consequence**	**Protein Change**	**ETH**	**Allele**	**Allele Freq**	**Effect**	**Standard Error**	**P value**	**N**
Insulin	19:40762860	rs184042322	*AKT2*	missense	p.P50T	1	T	0.01	0.12	0.02	1.2E-07	28118
	6:43758873	rs6905288	*VEGFA*	downstream gene	NA	1	A	0.56	0.02	0.00	4.2E-07	17898
	9:116764392	rs143246917	*ZNF618*	intron	NA	1	A	0.99	-0.77	0.14	9.2E-07	507
	8:23004629	rs3924519	*TNFRSF10D*	intron	NA	5	T	0.56	-0.06	0.01	9.8E-07	4556
	6:30414848	rs1362115	*HLA-E*	NA	NA	1	T	0.15	-0.02	0.01	1.9E-06	30825
	10:116331030	rs3824819	*ABLIM1*	intron	NA	1	T	0.07	-0.28	0.05	2.0E-06	1103
	6:30428351	rs2077573	*HLA-E*	NA	NA	1	A	0.85	0.02	0.01	2.3E-06	30825

Supplementary Table 2C: Single variant association results at previously published genome-wide association loci. Each row contains a previously reported GWAS association with FG level or FI level. Not all previously published SNPs were available for analysis in the exome array or exome sequencing data (denoted with - for our analyses).

Fasting Glucose				Published							WES + ExomeArray				
rsID	Gene	BMI	PHENO	Eff /Neff	Effective Freq	Effect	P	N	Ancestry	CITATION	Freq	Effect	Std Error	P	N
rs340874	*PROX1*	No	FGlu	C/T	0.52	0.021	6.6E-12	116882	European	Dupuis et al. (Nat Genet 2010)	0.49	0.01	0.00	2.23E-02	33231
rs560887	*G6PC2*	No	FGlu	C/T	0.7	0.075	8.7E-218	119169	European	Dupuis et al. (Nat Genet 2010); Prokopenko et al. (Nat Genet 2008); Sabatti et al. (Nat Genet 2008); Kristiansson et al. (Circ Cardiovasc Genet 2012); Bouatia et al. (Science 2008)	0.70	0.07	0.00	7.88E-87	38339
rs1371614	*DPYSL5*	Yes	FGluBMIadj	T/C	0.25	0.020/ 0.022	2.3E-12	96496	European	Manning et al. (Nat Genet 2012)					
rs780094	*GCKR*	No	FGlu	C/T	0.62	0.026	5.6E-38	118032	European	Dupuis et al. (Nat Genet 2010); Prokopenko et al. (Nat Genet, 2008)	0.63	0.03	0.00	1.37E-18	33231
rs3736594	*MRPL33*	Yes	FGluBMIadj	A/C	0.27		1.1E-15	96487	European	Manning et al. (Nat Genet 2012)	-	-	-	-	-
rs895636	*SIX3 - SIX2*	No	FGlu	C/T	0.38	0.039	1.0E-12	17617	East Asian	Kim et al. (Nat Genet 2011)	-	-	-	-	-
rs11715915	*AMT*	No	FGlu	C/T	0.68	0.012	4.9E-08	131523	European	Scott et al. (Nat Genet 2012)	0.63	0.01	0.00	5.34E-02	38337
rs11708067	*ADCY5*	No	FGlu	A/G	0.78	0.027	7.1E-22	118475	European	Dupuis et al. (Nat Genet 2010)	0.79	0.02	0.00	1.33E-04	33228
rs11920090	*SLC2A2*	No	FGlu	T/A	0.87	0.025	8.1E-13	119024	European	Dupuis et al. (Nat Genet 2010)	0.87	0.02	0.01	4.87E-05	33231
rs7651090	*IGF2BP2*	No	FGlu	G/A	0.3	0.013	1.8E-08	104019	European	Scott et al. (Nat Genet 2012)	0.31	0.00	0.00	7.51E-01	33231
rs7708285	*ZBED3*	Yes	FGluBMIadj	G/A	0.27	0.015	1.2E-08	117931	European	Scott et al. (Nat Genet 2012)	-	-	-	-	-
rs4869272	*PCSK1*	No	FGlu	T/C	0.69	0.018	1.0E-15	131872	European	Scott et al. (Nat Genet 2012)	0.68	0.02	0.00	8.32E-07	33231
rs13179048	*PCSK1*		FGluBMIadj	C/A	0.69	0.022/ 0.018	1.6E-10	96496	European	Manning et al. (Nat Genet 2012)	0.70	0.01	0.01	2.41E-01	4532
rs9368222	*CDKAL1*	No	FGlu	A/C	0.28	0.014	1.0E-09	128453	European	Scott et al. (Nat Genet 2012)	-	-	-	-	-
rs17762454	*RREB1*	Yes	FGluBMIadj	T/C	0.26	0.014	9.6E-09	123247	European	Scott et al. (Nat Genet 2012)	-	-	-	-	-
rs1127065	*CAMK2B*	No	FGlu	G/A	0.49	0.08	8.9E-11	11616	European	Kristiansson et al. (Circ Cardiovasc Genet, 2012)	0.59	0.03	0.01	5.20E-04	5108
rs2191349	*DGKB/TMEM195*	No	FGlu	T/G	0.52	0.03	3.0E-44	122743	European	Dupuis et al. (Nat Genet 2010)	0.49	0.03	0.00	1.25E-14	33231
rs6947830	*DGKB/TMEM195*	No	FGlu	A/G	0.46	0.1	1.4E-13	11616	European	Kristiansson et al. (Circ Cardiovasc Genet, 2012)	-	-	-	-	-
rs1799884	*GCK*	No	FGlu	A/G	0.85	0.063	4.5E-18	14211	East Asian	Go et al. (J Hum Genet 2013)	0.13	0.06	0.01	4.94E-21	24042
rs3757840	*GCK*	No	FGlu	A/C	0.46	0.1	4.9E-13	11616	European	Kristiansson et al. (Circ Cardiovasc Genet, 2012)	-	-	-	-	-
rs6975024	*GCK*	No	FGlu	C/T	0.15	0.061	2.9E-99	103517	European	Scott et al. (Nat Genet 2012)	0.14	0.06	0.01	2.25E-31	33228
rs4607517	*GCK*	No	FGlu	A/G	0.16	0.062	6.5E-92	118500	European	Dupuis et al. (Nat Genet 2010)	0.14	0.06	0.01	2.25E-31	33231
rs6943153	*GRB10*	No	FGlu	T/C	0.34	0.015	1.6E-12	131795	European	Scott et al. (Nat Genet 2012)	0.39	0.02	0.00	6.66E-08	33230
rs11558471	*SLC30A8*	No	FGlu	A/G	0.68	0.027	2.6E-11	45996	European	Dupuis et al. (Nat Genet 2010)	0.64	0.02	0.00	2.09E-10	33231
rs983309	*PPP1R3B*	No	FGlu	T/G	0.12	0.026	6.3E-15	127470	European	Scott et al. (Nat Genet 2012)	-	-	-	-	-
rs4841132	*PPP1R3B*		FGluBMIadj	A/G	0.1	0.027/0.030	7.6E-09	96496	European	Manning et al. (Nat Genet 2012)	0.13	0.03	0.01	1.96E-07	33231
rs2126259	*PPP1R3B*	No	FGlu	T/C	0.11	0.51	6.3E-15	124740	European	Scott et al. (Nat Genet 2012)	0.14	0.02	0.01	3.38E-06	33230
rs16913693	*IKBKAP*	No	FGlu	T/G	0.97	0.043	3.5E-11	125115	European	Scott et al. (Nat Genet 2012)	0.96	0.04	0.01	7.46E-05	28667
rs3829109	*DNLZ*	No	FGlu	G/A	0.71	0.017	1.1E-10	115310	European	Scott et al. (Nat Genet 2012)	0.68	0.01	0.00	1.24E-02	33229
rs10811661	*CDKN2B*	No	FGlu	T/C	0.82	0.024	5.6E-18	128488	European	Scott et al. (Nat Genet 2012)	-	-	-	-	-
rs7034200	*GLIS3*	No	FGlu	A/C	0.49	0.014	1.0E-12	106250	European	Dupuis et al. (Nat Genet 2010)	0.48	0.01	0.00	4.89E-04	33173
rs10885122	*ADRA2A*	No	FGlu	G/T	0.87	0.038	2.9E-16	118410	European	Dupuis et al. (Nat Genet 2010)	0.87	0.02	0.01	7.89E-05	33230
rs4506565	*TCF7L2*	No	FGlu	T/A	0.31	0.023	1.2E-08	46181	European	Dupuis et al. (Nat Genet 2010)	0.26	0.02	0.00	9.85E-06	33230
rs7903146	*TCF7L2*	No	FGlu	C/T	0.72	-0.022	2.7E-20	127477	European	Scott et al. (Nat Genet 2012)	0.77	-0.02	0.00	4.31E-07	33231
rs11605924	*CRY2*	No	FGlu	A/C	0.49	0.022	1.0E-14	116479	European	Dupuis et al. (Nat Genet 2010)	0.52	0.02	0.01	3.46E-04	8772
rs7944584	*MADD*	No	FGlu	A/T	0.75	0.025	2.0E-18	118741	European	Dupuis et al. (Nat Genet 2010)	0.77	0.03	0.00	2.62E-11	33231
rs1483121	*OR4S1*	Yes	FGluBMIadj	G/A	0.86	0.021/0.015	1.6E-08	96496	European	Manning et al. (Nat Genet 2012)	0.87	0.01	0.01	1.75E-02	28692
rs174550	*FADS1*	No	FGlu	T/C	0.64	0.022	1.7E-15	118908	European	Dupuis et al. (Nat Genet 2010)	0.62	0.02	0.00	3.37E-07	33230
rs11603334	*ARAP1*	No	FGluBMIadj	G/A	0.83	0.022/0.030	2.4E-14	96496	European	Manning et al. (Nat Genet 2012)	0.79	0.02	0.00	1.55E-08	33231
rs11603334	*ARAP1*	No	FGlu	G/A	0.83	0.019	1.1E-11	128139	European	Scott et al. (Nat Genet 2012)	0.79	0.02	0.00	1.55E-08	33231
rs2166706	*FAT3 - MTNR1B*	No	FGlu	G/A	0.450-0.462	0.05	2.1E-09	6776	South Asian	Chambers et al. (Diabetes 2009)	0.40	0.06	0.00	5.48E-57	33231
rs10830962	*MTNR1B*	No	FGlu	G/C	0.4	0.12	5.0E-16	11616	European	Kristiansson et al. (Circ Cardiovasc Genet, 2012)	-	-	-	-	-

Fasting Glucose				Published							WES + ExomeArray				
rsID	Gene	BMI	PHENO	Eff /Neff	Effective Freq	Effect	P	N	Ancestry	CITATION	Freq	Effect	Std Error	P	N
rs10830962	*MTNR1B*	No	FGlu	C/G	0.531	0.041	4.8E-13	14081	EastAsian	Go et al. (J Hum Genet 2013)	-	-	-	-	-
rs10830963	*MTNR1B*	No	FGlu	G/C	0.205	0.048	3.7E-08	815	Hispanic	Comuzzie et al. (PLoS One 2012)	0.31	0.09	0.00	2.79E-118	33230
rs10830963	*MTNR1B*	No	FGlu	G/C	0.3	0.079	5.8E-175	112844	European	Dupuis et al. (Nat Genet 2010); Prokopenko et al. (Nat Genet, 2008)	0.31	0.09	0.00	2.79E-118	33230
rs2657879	*GLS2*	Yes	FGluBMIadj	G/A	0.18	0.016	3.9E-08	123247	European	Scott et al. (Nat Genet 2012)	0.18	0.00	0.00	1.87E-01	38339
rs2074356	*C12orf51*	No	FGlu	T/C	0.199	-0.061	6.0E-14	14193	East Asian	Go et al. (J Hum Genet 2013)	-	-	-	-	-
rs10747083	*P2RX2*	No	FGlu	A/G	0.66	0.013	7.6E-09	127111	European	Scott et al. (Nat Genet 2012)	0.64	0.01	0.01	1.89E-02	16158
rs11619319	*PDX1*	No	FGlu	G/A	0.23	0.02	1.3E-15	132996	European	Scott et al. (Nat Genet 2012)	0.24	0.02	0.00	7.73E-06	33226
rs2293941	*PDX1*	No	FGluBMIadj	A/G	0.22	0.019/0.016	5.3E-10	96496	European	Manning et al. (Nat Genet 2012)	-	-	-	-	-
rs576674	*KL*	No	FGlu	G/A	0.15	0.017	2.3E-08	131856	European	Scott et al. (Nat Genet 2012)	0.14	0.02	0.01	1.48E-03	28601
rs3783347	*WARS*	No	FGlu	G/T	0.79	0.017	1.3E-10	132544	European	Scott et al. (Nat Genet 2012)	0.78	0.01	0.00	1.01E-03	33231
rs11071657	*C2CD4B*	No	FGlu	A/G	0.63	0.021	3.6E-08	114454	European	Dupuis et al. (Nat Genet 2010)	0.64	0.01	0.00	1.01E-03	33230
rs2302593	*GIPR*	No	FGlu	C/G	0.5	0.014	9.3E-10	116141	European	Scott et al. (Nat Genet 2012)	0.53	-0.01	0.01	4.74E-01	5108
rs6113722	*FOXA2*	No	FGlu	G/A	0.96	0.353	2.5E-11	123665	European	Scott et al. (Nat Genet 2012)	0.96	0.02	0.01	1.12E-02	33231
rs6048205	*FOXA2*	No	FGluBMIadj	A/G	0.95	0.040/0.029	1.6E-12	96496	European	Manning et al. (Nat Genet 2012)	-	-	-	-	-
rs1209523	*FOXA2*	No	FGlu	T/C	0.037-0.391	-	2.2E-11	14853	African American + European	Xing et al. (Am J Hum Genet 2013)	-	-	-	-	-
rs6072275	*TOP1*	No	FGlu	A/G	0.16	0.016	1.7E-08	128616	European	Scott et al. (Nat Genet 2012)	0.20	0.02	0.00	6.06E-05	33231

Fasting Insulin				Published							WES + ExomeArray				
rsID	Gene	BMI	PHENO	Eff /Neff	Effective Freq	Effect	P	N	Ancestry	CITATION	Freq	Effect	Std Error	P	N
rs2820436	*LYPLAL1*		FIns	C/A	0.67	0.015	4.4E-09	104044	European	Scott et al. (Nat Genet 2012)	-	-	-	-	-
rs2785980	*LYPLAL1*	Yes	FInsBMIadj	T/C	0.67	0.016/0.017	2.0E-08	83116	European	Manning et al. (Nat Genet 2012)	0.66	0.01	0.00	3.4E-02	17731
rs4846565	*LYPLAL1*		FInsBMIadj	G/A	0.67	0.013	1.8E-09	99014	European	Scott et al. (Nat Genet 2012)	-	-	-	-	-
rs7607980	*GRB14*	Yes	FInsBMIadj	T/C	0.88	0.023/0.039	4.3E-20	83116	European	Manning et al. (Nat Genet 2012)	0.88	0.03	0.01	3.1E-09	34278
rs1530559	*YSK4*	No	FIns	T/C	0.52	0.015	3.4E-08	107281	European	Scott et al. (Nat Genet 2012)	-	-	-	-	-
rs10195252	*GRB14*		FIns	T/C	0.59	0.016	4.9E-10	99126	European	Scott et al. (Nat Genet 2012)	0.60	0.02	0.00	3.0E-05	21680
rs10195252	*GRB14*	Yes	FInsBMIadj	T/C	0.6	0.017	1.3E-16	98997	European	Scott et al. (Nat Genet 2012)	0.60	0.02	0.00	3.0E-05	21680
rs2943634	*IRS1*	Yes	FInsBMIadj	C/A	0.66	0.018/0.025	2.5E-14	83116	European	Manning et al. (Nat Genet 2012)	0.66	0.03	0.00	7.7E-13	30816
rs2943645	*IRS1*		FInsBMIadj	T/C	0.63	0.019	2.3E-19	99023	European	Scott et al. (Nat Genet 2012)	-	-	-	-	-
rs2972143	*IRS1*		FIns	G/A	0.62	0.014	3.2E-08	99566	European	Scott et al. (Nat Genet 2012)	-	-	-	-	-
rs780094	*GCKR*	No	FIns	C/T	0.62	0.015	3.6E-20	96126	European	Dupuis et al. (Nat Genet 2010)	0.63	0.02	0.00	6.3E-11	30825
rs17036328	*PPARG*	Yes	FInsBMIadj	T/C	0.86	0.021	3.6E-12	98497	European	Scott et al. (Nat Genet 2012)	-	-	-	-	-
rs974801	*TET2*		FInsBMIadj	G/A	0.38	0.014	3.3E-11	103489	European	Scott et al. (Nat Genet 2012)	-	-	-	-	-
rs9884482	*TET2*	No	FIns	C/T	0.39	0.017	1.4E-11	108420	European	Scott et al. (Nat Genet 2012)	0.39	0.01	0.00	8.7E-03	26330
rs4691380	*PDGFC*		FInsBMIadj	C/T	0.67	0.016/0.021	5.3E-09	83116	European	Manning et al. (Nat Genet 2012)	0.71	0.01	0.00	2.3E-03	30825
rs6822892	*PDGFC*	Yes	FInsBMIadj	A/G	0.69	0.014	2.6E-10	103432	European	Scott et al. (Nat Genet 2012)	0.70	0.01	0.01	3.3E-02	17280
rs17046216	*SC4MOL*	No	FIns	A/T	0.48	0.18	1.7E-08	1497	African American	Chen et. al (Hum Mol Genet 2012)	-	-	-	-	-
rs3822072	*FAM13A*	Yes	FInsBMIadj	A/G	0.48	0.012	1.8E-08	99977	European	Scott et al. (Nat Genet 2012)	-	-	-	-	-
rs4865796	*ARL15*	No	FIns	A/G	0.67	0.015	2.1E-08	100001	European	Scott et al. (Nat Genet 2012)	-	-	-	-	-
rs4865796	*ARL15*		FInsBMIadj	A/G	0.67	0.015	2.2E-12	98314	European	Scott et al. (Nat Genet 2012)	-	-	-	-	-
rs459193	*ANKRD55/MAP3K1*	Yes	FInsBMIadj	G/A	0.73	0.015	1.1E-12	103378	European	Scott et al. (Nat Genet 2012)	0.71	0.02	0.00	1.5E-06	30825
rs2745353	*RSPO3*	No	FIns	T/C	0.51	0.014	5.5E-09	104075	European	Scott et al. (Nat Genet 2012)	0.52	0.01	0.00	3.7E-03	30825
rs6912327	*UHRF1BP1*		FInsBMIadj	T/C	0.8	0.017	2.3E-08	80010	European	Scott et al. (Nat Genet 2012)	-	-	-	-	-
rs4646949	*UHRF1BP1*	Yes	FInsBMIadj	T/G	0.75	0.014/0.020	3.7E-08	83116	European	Manning et al. (Nat Genet 2012)	0.77	0.01	0.00	7.5E-02	30824
rs1167800	*HIP1*	No	FIns	A/G	0.54	0.016	2.6E-09	90927	European	Scott et al. (Nat Genet 2012)	0.55	0.01	0.00	7.1E-02	30825
rs983309	*PPP1R3B*		FInsBMIadj	T/G	0.12	0.022	1.2E-12	99024	European	Scott et al. (Nat Genet 2012)	-	-	-	-	-
rs983309	*PPP1R3B*		FIns	T/G	0.12	0.029	3.8E-14	103030	European	Scott et al. (Nat Genet 2012)	-	-	-	-	-
rs4841132	*PPP1R3B*		FInsBMIadj	A/G	0.1	0.021/0.031	1.7E-10	83116	European	Manning et al. (Nat Genet 2012)	0.13	0.02	0.01	1.4E-04	30825
rs2126259	*PPP1R3B*	Yes	FInsBMIadj	T/C	0.11	0.024	3.3E-13	99021	European	Scott et al. (Nat Genet 2012)	0.14	0.02	0.01	2.2E-04	30824
rs7903146	*TCF7L2*	No	FIns	C/T	0.72	0.018	6.1E-11	103037	European	Scott et al. (Nat Genet 2012)	0.77	0.01	0.00	2.8E-03	30825
rs7077836	*TCERG1L*	No	FIns	T/C	0.12	0.28	7.5E-09	1497	African American	Chen et. al (Hum Mol Genet 2012)	-	-	-	-	-
rs35767	*IGF1*	No	FIns	G/A	0.85	0.028	3.3E-08	94590	European	Dupuis et al. (Nat Genet 2010)	0.81	0.01	0.00	6.1E-04	30825
rs1421085	*FTO*	No	FIns	C/T	0.42	0.02	1.9E-15	104062	European	Scott et al. (Nat Genet 2012)	0.41	0.00	0.00	5.7E-01	30825
rs731839	*PEPD*	Yes	FInsBMIadj	G/A	0.34	0.015	5.1E-12	103252	European	Scott et al. (Nat Genet 2012)	0.34	0.02	0.00	6.7E-05	30825

Supplementary Table 2D: Significant and suggestive gene based association signals. Results for all data and mask combinations are shown for any gene that attains exome-wide significant (** $P<2.5\times10^{-6}$) or exome-wide suggestive levels (* $P<2.5\times10^{-5}$).

Fasting Insulin		PTV+missense			PTV+NS$_{broad}$			PTV+NS$_{strict}$			PTV-only		
Gene	Ancestry	MAC (No. vars)	P value SKAT	P value Burden	MAC (No. vars)	P value SKAT	P value Burden	MAC (No. vars)	P value SKAT	P value Burden	MAC (No. vars)	P value SKAT	P value Burden
AKT2	AfrAm	1	0.67	0.67	1	0.67	0.67	-	-	-	-	-	-
19q13.1-q13.2	E.Asian	5	0.33	0.15	5	0.33	0.15	0	0.65	0.65	-	-	-
	Europ	31	0.53	0.31	31	0.53	0.31	-	-	-	-	-	-
	Hisp	7	0.42	0.13	7	0.42	0.13	-	-	-	-	-	-
	S.Asian	2	0.86	0.83	1	0.6	0.6	-	-	-	-	-	-
	WES (all)	46(36)	0.6	0.051	45(33)	0.57	0.052	0(5)	0.65	0.65	-	-	-
	ExArray	398(4)	**6.10E-07**	**3.60E-06**	398(4)	**6.10E-07**	**3.60E-06**	-	-	-	-	-	-
	WES (all) + ExArray	444	0.00056	**7.30E-06**	443	0.00048	**7.50E-06**	0	0.65	0.65	-	-	-
NDUFAF1	AfrAm	15	0.25	0.92	7	0.29	0.24	2	0.35	0.29	-	-	-
15q11.2-q21.3	E.Asian	12	0.4	0.96	6	0.62	0.51	4	0.54	0.4	-	-	-
	Europ	36	9.60E-05	4.10E-05	35	9.90E-05	0.0001	31	9.30E-05	9.30E-05	-	-	-
	Hisp	18	0.056	0.011	14	0.045	0.0058	5	0.033	0.011	-	-	-
	S.Asian	10	0.44	0.32	1	0.2	0.2	-	-	-	-	-	-
	WES (all)	91(58)	6.10E-05	0.0001	63(38)	6.20E-05	**9.20E-07**	42(14)	7.60E-05	**2.20E-06**	0(2)	0	0
	ExArray	555(9)	0.02	0.094	535(6)	0.021	0.044	418(2)	0.017	0.018	-	-	-
	WES (all) + ExArray	646	**1.50E-05**	0.00019	598	**1.60E-05**	**2.30E-06**	460	**1.50E-05**	**1.10E-06**	-	-	-
ALPK1	AfrAm	42	0.83	0.7	30	0.89	0.26	5	0.3	0.059	3	0.26	0.11
4q25	E.Asian	82	0.85	0.26	55	0.63	0.4	32	0.85	0.84	23	0.68	0.69
	Europ	59	0.25	0.77	93	0.46	0.97	51	0.73	0.7	5	0.36	0.2
	Hisp	43	0.73	0.44	41	0.49	0.65	14	0.16	0.13	3	0.39	0.39
	S.Asian	26	0.036	**6.50E-06**	22	0.033	**1.70E-05**	14	0.24	0.011	4	0.16	0.017
	WES (all)	252(158)	0.65	0.062	241(105)	0.55	0.014	116(36)	0.7	0.071	38(16)	0.6	0.17
	ExArray	5514(26)	0.87	0.75	3237(17)	0.74	0.76	291(4)	0.91	0.83	-	-	-
	WES (all) + ExArray	5766	0.86	0.27	3478	0.71	0.15	407	0.91	0.36	38	0.6	0.17
ZBTB10	AfrAm	2	0.56	0.37	2	0.56	0.37	-	-	-	-	-	-
8q13-q21.1	E.Asian	5	0.18	0.26	5	0.18	0.26	-	-	-	-	-	-
	Europ	7	0.97	0.95	7	0.97	0.95	-	-	-	-	-	-
	Hisp	20	0.82	0.64	18	0.74	0.53	2	0.92	0.92	-	-	-
	S.Asian	5	0.45	0.41	3	0.21	0.46	0	0.73	0.73	-	-	-
	WES (all)	39(44)	0.86	0.41	35(34)	0.76	0.39	2(4)	0.92	0.91	-	-	-
	ExArray	646(5)	**7.40E-06**	**1.90E-05**	646(5)	**7.40E-06**	**1.90E-05**	-	-	-	-	-	-
	WES (all) + ExArray	685	0.011	0.0011	681	0.0051	0.00094	2	0.92	0.91	-	-	-
PLCB3	AfrAm	15	0.0061	0.00012	11	0.0078	**2.10E-05**	5	0.0072	0.00056	1	0.0056	0.0056
11q13	E.Asian	24	0.13	0.2	19	0.12	0.6	7	0.16	0.99	-	-	-
	Europ	77	0.59	0.86	65	0.57	0.84	3	0.074	0.093	1	0.9	0.9
	Hisp	36	0.23	0.62	32	0.19	0.54	2	0.86	0.59	-	-	-
	S.Asian	19	0.65	0.48	15	0.45	0.29	4	0.97	0.62	-	-	-
	WES (all)	173(121)	0.27	0.35	144(87)	0.23	0.16	21(27)	0.02	0.1	2(2)	0.024	0.043
	ExArray	730(12)	0.64	0.84	668(6)	0.67	0.82	8(1)	0.58	0.58	-	-	-
	WES (all) + ExArray	903	0.42	0.67	812	0.42	0.48	29	0.093	0.23	2	0.024	0.043

Fasting glucose		PTV+missense			PTV+NS$_{broad}$			PTV+NS$_{strict}$			PTV-only		
Gene	Ancestry	MAC (No. vars)	P value SKAT	P value Burden	MAC (No. vars)	P value SKAT	P value Burden	MAC (No. vars)	P value SKAT	P value Burden	MAC (No. vars)	P value SKAT	P value Burden
G6PC2	AfrAm	17	0.11	0.42	13	0.043	0.29	5	0.17	0.78	2	0.21	0.083
2q24.3	E.Asian	26	0.26	0.1	21	0.2	0.022	4	0.11	0.034	3	0.18	0.18
	Europ	93	0.23	0.11	90	0.22	0.14	69	0.2	0.16	7	0.63	0.63
	Hisp	23	0.2	0.24	22	0.19	0.17	21	0.19	0.22	5	0.049	0.53
	S.Asian	11	0.059	0.053	9	0.046	0.02	8	0.049	0.047	-	-	-
	WES (all)	170(69)	0.12	0.0028	155(53)	0.1	0.00078	107(19)	0.11	0.01	17(8)	0.22	0.07
	ExArray	1174(15)	**1.80E-13**	**4.10E-16**	1129(12)	**2.00E-13**	**1.20E-17**	913(4)	**3.60E-12**	**5.10E-13**	71(1)	0.67	0.67
	WES (all) + ExArray	1344	**1.30E-09**	**9.90E-15**	1284	**8.30E-10**	**9.60E-17**	1020	**5.40E-09**	**1.30E-11**	88	0.41	0.23
GIMAP8	AfrAm	24	0.49	0.28	19	0.35	0.38	3	0.04	0.055	1	0.0019	0.0019
7q36.1	E.Asian	75	0.58	0.92	38	0.71	0.15	3	0.37	0.15	3	0.37	0.15
	Europ	18	0.95	0.54	12	0.75	0.53	4	0.54	0.56	1	0.13	0.13
	Hisp	24	0.35	0.88	22	0.3	0.85	6	0.077	0.068	4	0.048	0.048
	S.Asian	10	0.031	0.61	6	0.0096	0.28	3	0.011	0.0022	3	0.011	0.0022
	WES (all)	151(87)	0.6	0.43	97(52)	0.47	0.088	19(15)	0.012	0.00013	12(11)	0.0029	**2.30E-06**
	ExArray	240(14)	0.25	0.84	219(7)	0.25	0.77	17(2)	0.29	0.19	-	-	-
	WES (all) + ExArray	391	0.38	0.72	316	0.3	0.34	36	0.023	0.00065	12	0.0029	**2.30E-06**
OR4S1	AfrAm	43	0.69	0.095	18	0.8	0.2	-	-	-	-	-	-

Fasting glucose			PTV+missense			PTV+NS$_{broad}$			PTV+NS$_{strict}$			PTV-only		
Gene	Ancestry	MAC (No. vars)	P value SKAT	P value Burden	MAC (No. vars)	P value SKAT	P value Burden	MAC (No. vars)	P value SKAT	P value Burden	MAC (No. vars)	P value SKAT	P value Burden
11p11.2	E.Asian	11	0.032	0.16	4	0.033	0.027	-	-	-	-	-	-
	Europ	19	0.15	0.34	13	0.36	0.87	-	-	-	-	-	-
	Hisp	22	0.21	0.87	15	0.1	0.53	1	0.56	0.56	1	0.56	0.56
	S.Asian	20	0.27	0.057	6	0.16	0.029	-	-	-	-	-	-
	WES (all)	115(75)	0.15	0.0074	56(52)	0.16	0.023	1(3)	0.56	0.56	1(3)	0.56	0.56
	ExArray	201(8)	0.00051	3.70E-05	33(5)	0.075	0.036	-	-	-	-	-	-
	WES (all) + ExArray	316	0.0011	**3.10E-06**	89	0.041	0.0036	1	0.56	0.56	1	0.56	0.56
G6PC	AfrAm	1	0.62	0.62	1	0.62	0.62	-	-	-	-	-	-
17q21	E.Asian	10	0.73	0.41	9	0.7	0.53	6	0.49	0.27	1	0.74	0.74
	Europ	47	0.48	0.62	46	0.47	0.52	6	0.048	0.98	1	0.33	0.33
	Hisp	16	0.075	0.052	16	0.075	0.052	14	0.088	0.2	12	0.084	0.057
	S.Asian	5	0.76	0.71	4	0.63	0.84	-	-	-	-	-	-
	WES (all)	79(54)	0.38	0.51	76(48)	0.36	0.6	26(21)	0.063	0.14	14(5)	0.092	0.034
	ExArray	643(7)	**1.90E-05**	**9.30E-06**	643(7)	**1.90E-05**	**9.30E-06**	17(3)	0.072	0.077	3(1)	0.0056	0.0056
	WES (all) + ExArray	722	0.00086	0.0013	719	0.00076	0.0022	43	0.017	0.039	17	0.0031	0.001
PIK3AP1	AfrAm	15	0.39	0.87	9	0.47	0.74	0	0.34	0.34	-	-	-
10q24.1	E.Asian	22	0.42	0.19	10	0.074	0.12	3	0.18	0.36	-	-	-
	Europ	28	0.78	0.84	7	0.25	0.23	0	0.37	0.37	-	-	-
	Hisp	18	0.00018	0.0011	13	0.00049	**1.70E-05**	11	0.00045	**1.40E-05**	-	-	-
	S.Asian	11	0.92	0.8	4	0.85	0.3	3	0.8	0.41	-	-	-
	WES (all)	94(68)	0.019	0.054	43(42)	0.00059	0.017	17(15)	0.00048	0.005	-	-	-
	ExArray	204(9)	0.85	0.35	96(6)	0.57	0.27	35(2)	0.9	0.68	-	-	-
	WES (all) + ExArray	298	0.23	0.078	139	0.015	0.027	52	0.075	0.068	-	-	-
ZNF44	AfrAm	9	0.0093	0.5	7	0.071	0.084	-	-	-	-	-	-
19p13.2	E.Asian	11	0.72	0.79	7	0.63	0.41	2	0.16	0.054	2	0.16	0.054
	Europ	68	0.002	0.0058	50	0.0024	0.02	3	0.41	0.41	3	0.41	0.41
	Hisp	14	7.50E-05	0.32	14	7.50E-05	0.32	4	**1.40E-05**	**1.40E-05**	4	**1.40E-05**	**1.40E-05**
	S.Asian	21	0.51	0.004	16	0.54	0.015	1	0.26	0.26	1	0.26	0.26
	WES (all)	123(80)	0.00044	0.6	94(56)	0.0002	0.94	10(9)	**2.10E-05**	0.0086	10(9)	**2.10E-05**	0.0086
	ExArray	570(7)	0.84	0.88	307(5)	0.77	0.52	-	-	-	-	-	-
	WES (all) + ExArray	693	0.05	0.84	401	0.023	0.88	10	**2.10E-05**	0.0086	10	**2.10E-05**	0.0086
OR13A1	AfrAm	71	0.073	0.046	70	0.069	0.072	67	0.06	0.068	62	0.25	0.3
10q11.21	E.Asian	39	0.74	0.75	30	0.64	0.57	-	-	-	-	-	-
	Europ	184	0.16	0.024	180	0.15	0.029	152	0.82	0.82	151	0.77	0.77
	Hisp	93	0.31	0.89	87	0.18	0.99	81	0.1	0.52	80	0.14	0.14
	S.Asian	24	0.17	0.13	22	0.14	0.13	16	0.15	0.53	15	0.18	0.18
	WES (all)	412(58)	0.16	0.89	390(40)	0.12	0.9	317(6)	0.3	0.57	309(2)	0.45	0.96
	ExArray	290(9)	4.30E-05	4.20E-05	257(5)	3.70E-05	**1.50E-05**	-	-	-	-	-	-
	WES (all) + ExArray	702	0.00024	0.029	647	0.00013	0.021	317	0.3	0.57	309	0.45	0.96
ANKH	AfrAm	10	0.65	0.37	10	0.65	0.37	-	-	-	-	-	-
5p15.1	E.Asian	4	0.82	0.37	4	0.82	0.37	1	0.95	0.95	-	-	-
	Europ	22	0.16	0.95	16	0.24	0.4	-	-	-	-	-	-
	Hisp	9	0.55	0.37	9	0.55	0.37	-	-	-	-	-	-
	S.Asian	6	0.74	0.69	6	0.74	0.69	1	0.53	0.53	-	-	-
	WES (all)	51(46)	0.41	0.27	45(45)	0.61	0.082	2(11)	0.83	0.7	0(4)	0	0
	ExArray	371(5)	2.60E-05	0.016	202(4)	**1.70E-05**	**5.70E-06**	-	-	-	-	-	-
	WES (all) + ExArray	422	0.0013	0.025	247	0.0031	2.20E-05	2	0.83	0.7	-	-	-
MAP3K7CL	AfrAm	13	0.065	0.052	3	0.2	0.8	-	-	-	-	-	-
21q22.3	E.Asian	0	0.91	0.91	0	0.91	0.91	-	-	-	-	-	-
	Europ	3	0.38	0.23	2	0.55	0.62	-	-	-	-	-	-
	Hisp	4	0.34	0.59	4	0.34	0.59	1	0.07	0.07	1	0.07	0.07
	S.Asian	3	0.92	0.94	1	0.83	0.83	1	0.83	0.83	1	0.83	0.83
	WES (all)	23(24)	0.11	0.1	10(19)	0.42	0.97	2(4)	0.18	0.15	2(3)	0.18	0.15
	ExArray	9(2)	**1.90E-05**	**1.90E-05**	8(1)	**2.00E-05**	**2.00E-05**	-	-	-	-	-	-
	WES (all) + ExArray	32	7.60E-05	7.10E-05	18	0.0012	0.053	2	0.18	0.15	2	0.18	0.15
CDC42BPA	AfrAm	16	0.16	0.32	9	0.26	0.36	3	0.09	0.057	1	0.27	0.27
1q42.11	E.Asian	48	0.041	0.17	38	0.043	0.08	6	0.0007	**2.30E-05**	-	-	-
	Europ	22	0.79	0.88	18	0.64	0.97	9	0.44	0.93	-	-	-
	Hisp	20	0.21	0.49	12	0.24	0.36	6	0.23	0.22	-	-	-
	S.Asian	23	0.61	0.75	22	0.61	0.92	3	0.12	0.36	-	-	-
	WES (all)	130(154)	0.078	0.57	100(124)	0.086	0.29	27(38)	0.0025	0.11	1(2)	0.27	0.27
	ExArray	111(13)	0.76	0.086	93(9)	0.77	0.24	17(4)	0.11	0.3	-	-	-
	WES (all) + ExArray	241	0.31	0.2	193	0.33	0.19	44	0.0022	0.11	1	0.27	0.27

AfrAm: African American ancestry
E.Asian: East asian ancestry
Europ: European ancestry
Hisp: Hispanic ancestry
S.Asian: South Asican ancestry
WES (all): Whole exome sequencing meta-analysis
ExArray: Exome array meta-analysis
WES (all) + ExArray: Whole exome sequencing and exome array meta-analysis

Variant masks:
PTV: containing only variants predicted to introduce a premature stop codon
PTV+NS: containing variants in the PTV group and protein-altering variants with MAF<1%
PTV+NSstrict: composed of variants in "PTV" and protein-altering variants predicted damaging by SIFT, LRT, MutationTaster, polyphen2 HDIV, and polyphen2 HVAR
PTV+NSbroad: composed of "**PTV+NSstrict**" and NS variants with MAF<1% and predicted damaging by at least one prediction algorithm.

Supplementary Table 2E: Replication of *AKT2* p.Pro50Thr in independent Finnish cohorts and association results in the discovery and replication studies combined.

					Replication Analysis		Combined Discovery and Replication Analysis	
Trait	Location	Gene	Protein change	MAC	P	N	P	N
Fasting Insulin	19:40762860	*AKT2*	p.P50T	114	0.00054	5747	9.98E-10	25,316

MAC: Minor Allele Count
P: P-value
N: Sample size

SUPPLEMENTARY TABLE 3

Protein altering variation in AKT2. Displayed are all variants predicted to cause a nonsynonymous substitution or alter a splice site in 12,940 samples with whole exome sequencing data. Annotations were obtained using dbNSFP.

rsID	pos on chr19	Protein change	1000 Genomes Observations	MAF ExAC	MAC	MAC cases/ MAC controls	SIFT	LRT	Mutation Taster	Polyphen 2 HDIV	Polyphen2 HVAR	Cancer Tissue	Monogenic	Functional domain
-	40771156	p.I7V	1 Eur	5.69E-05	6	3/3	tolerated	D	D	B,B,B	B,B,B	NA		PH domain
rs387906659	40762959	E17K	-	0	0	0/0	deleterious	D	D	D,D,D	D,D,D	Thyroid; Breast	hypoketotic hypoglycemia with hemihypertrophy (Arya 2014, Hussain 2011)	PH domain
-	40762875	p.P45S	-	8.23E-06	1	0/1	tolerated	N	N	B,B,B	B,B,B	NA		PH domain
rs184042322	40762860	p.P50T	4 Eur	1.01E-03	61	39/22	tolerated	D	D	B,B,B	B,B,B	NA		PH domain
-	40761140	p.N71S	1 Amr	1.98E-04	4	1/3	tolerated	D	D	P,D,P,B	P,P,B,B	NA		PH domain
-	40761132	p.V74F	-	8.24E-06	1	0/1	tolerated	D	D	B,B,B,B	P,B,B,B	NA		PH domain
-	40761069	p.E95K	-	4.94E-05	1	1/0	deleterious	D	D	D,P,D,D	D,B,P,P	NA		PH domain
-	40761059	splice	-	8.24E-06	1	1/0	NA	NA	NA	NA	NA	NA		PH domain
-	40748581	p.R101W	-	4.16E-05	1	0/1	deleterious	N	D	B,B,B,B	B,B,B,B	NA		PH domain
-	40748568	p.M105T	-	8.29E-06	1	1/0	tolerated	D	D	B,B,B,B	B,B,B,B	NA		PH domain
rs141209878	40748535	p.G116A	1 Eur	2.64E-04	3	1/2	tolerated	D	N	B,B,B,B	B,B,B,B	NA		
-	40748529	p.D118G	-	8.26E-06	1	0/1	tolerated	D	D	B,B,B,B	B,B,B,B	NA		
-	40748526	p.P119L	-	8.26E-06	1	0/1	tolerated	N	D	B,B,B,B	B,B,B,B	NA		
-	40748518	p.Y122H	-	4.95E-05	4	2/2	tolerated	N	N	B,B,B,B	B,B,B,B	NA		
-	40748517	p.Y122C	1 Eur	1.49E-04	4	2/2	tolerated	N	D	B,B,B,B	B,B,B,B	NA		
-	40748480	p.E134D	-	0	1	0/1	tolerated	D	D	B,B,B,B	B,B,B,B	NA		
-	40748470	p.V138L	-	8.25E-06	1	1/0	tolerated	D	D	B,B,B,B	B,B,B,B	NA		
-	40747984	splice	-	4.87E-04	5	3/2	NA	NA	NA	NA	NA	NA		
-	40747892	p.R176C	-	2.48E-05	1	0/1	deleterious	D	D	D,P,D,D	D,P,P,P	NA		Protein kinase
-	40747891	p.R176L	-	1.65E-05	2	1/1	tolerated	D	D	B,B,B,B	B,B,B,B	NA		Protein kinase
-	40747846	p.K191R	-	3.33E-05	1	1/0	tolerated	NA	NA	NA	NA	NA		Protein kinase
-	40747837	splice	-	2.52E-05	3	1/2	NA	NA	NA	NA	NA	NA		Protein kinase
-	40746015	p.D192E	-	8.24E-06	1	1/0	tolerated	D	D	D,B,P,B	D,B,P,B	NA		Protein kinase
rs35817154	40745968	p.R208K	-	2.88E-04	4	2/2	tolerated	D	D	B,B,B,B	B,B,B,B	NA	Severe IR and acanthosis nigricans* (Tan 2007)	Protein kinase
-	40744879	p.A214V	-	2.49E-05	1	1/0	tolerated	D	D	B,B,B	B,B,B	Prostate		Protein kinase
-	40744805	splice	-	1.65E-05	1	1/0	NA	NA	NA	NA	NA	NA		Protein kinase
-	40744001	splice	-	2.50E-04	2	1/1	NA	NA	NA	NA	NA	NA		Protein kinase
-	40743973	p.R245H	-	2.85E-05	2	1/1	deleterious	D	D	P,D,D	B,P,D	NA		Protein kinase
-	40743956	p.R251W	-	0	2	2/0	deleterious	D	D	D,D,D	D,D,D	CCLE		Protein kinase
-	40743953	p.A252T	-	1.22E-05	2	1/1	tolerated	D	D	B,B,B	B,B,B	NA		Protein kinase
-	40743887	p.R274C	-	1.75E-05	2	1/1	deleterious	D	D	D,D,D	D,D,D	NA		Protein kinase

rsID	pos on chr19	Protein change	1000 Genomes Observations	MAF ExAC	MAC	MAC cases/ MAC controls	SIFT	LRT	Mutation Taster	Polyphen 2 HDIV	Polyphen2 HVAR	Cancer Tissue	Monogenic	Functional domain
rs121434593	40743886	p.R274H	-	0	0	0/0	deleterious	D	A	D,D,D	D,P,D	NA	severe insulin resistance and diabetes (George 2004)	Protein kinase
-	40743872	splice	-	1.11E-04	6	4/2	NA	NA	NA	NA	NA	NA		Protein kinase
-	40742207	p.T306S	-	1.40E-04	5	1/4	tolerated	D	D	B,B	B,B	NA		Protein kinase glycosylation site
-	40741992	p.Y327C	-	0	1	1/0	deleterious	D	D	D,D,D	D,D,D	NA		Protein kinase
-	40741915	p.Q353E	-	8.26E-06	1	0/1	tolerated	D	D	B,B,B	B,B,B	NA		Protein kinase
-	40741876	p.E366K	-	2.49E-05	3	1/2	tolerated	D	D	B,B,B	B,B,B	NA		Protein kinase
-	40741270	splice	-	6.70E-05	2	1/1	NA	NA	NA	NA	NA	NA		Protein kinase
-	40741222	p.M404T	-	8.26E-06	1	0/1	tolerated	D	D	P,P,B	P,B,B	NA		Protein kinase
-	40741212	p.R407S	-	0	1	0/1	tolerated	D	D	B,B,B	B,B,B	NA		Protein kinase
-	40741181	p.V418F	-	8.25E-06	1	1/0	tolerated	N	D	B,B,B	B,B,B	NA		AGC-kinase C-terminal
-	40741176	p.Q419H	-	8.25E-06	1	0/1	tolerated	N	D	B,B,B	B,B,B	NA		AGC-kinase C-terminal
-	40741058	splice	-	9.90E-05	2	0/2	NA	NA	NA	NA	NA	NA		AGC-kinase C-terminal
-	40741026	p.T431M	-	2.48E-05	1	1/0	deleterious	D	D	B,P,B	B,B,B	NA		AGC-kinase C-terminal
rs191069336	40739865	splice	-	9.55E-05	2	1/1	NA	NA	NA	NA	NA	NA		AGC-kinase C-terminal
-	40739862	spice	-	8.65E-06	1	1/0	NA	NA	NA	NA	NA	NA		AGC-kinase C-terminal
-	40739853	p.S458C	-	1.71E-05	2	0/2	tolerated	N	D	B,B	B,B	NA		AGC-kinase C-terminal
rs142926499	40739826	p.R467W	-	1.01E-04	1	0/1	deleterious	D	D	D,D	P,P	NA	T2D and partial lipodystrophy* (Tan 2007)	AGC-kinase C-terminal

SUPPLEMENTARY TABLE 4

Association of *AKT2* p.Pro50Thr with diabetes-related metabolic traits in Finnish Cohorts.

Supplementary Table 4A: Association with quantitative metabolic traits.

Trait Group	Trait	N	MAF	Effect (Std. Err) on inverse-normalized trait residuals	P	Padjusted
Anthropometric Traits	Waist-hip ratio	31966	0.012	0.045 (0.0383)	0.24	1
	Waist-hip ratio - females	12445	0.011	0.0822 (0.065)	0.21	1
	Waist-hip ratio - males	19521	0.013	0.0299 (0.0473)	0.53	1
	Waist circumference	31970	0.012	0.0354 (0.0384)	0.36	1
	Waist circumference - females	12448	0.011	0.0741 (0.065)	0.25	1
	Waist circumference - males	19522	0.013	0.0227 (0.0475)	0.63	1
	Hip circumference	31972	0.012	-0.00851 (0.0384)	0.83	1
	Hip circumference - females	12448	0.011	-0.0254 (0.0648)	0.70	1
	Hip circumference - males	19524	0.013	-0.00317 (0.0476)	0.95	1
	Body mass index	34597	0.012	-0.0978 (0.0371)	0.01	0.19
	Height	34601	0.012	-0.105 (0.0373)	4.7E-03	0.11
Lipid Traits	HDL-C	36923	0.012	0.027 (0.0348)	0.44	1
	LDL-C	31045	0.012	0.0604 (0.0372)	0.11	1
	Total cholesterol	36939	0.012	0.0926 (0.0348)	0.01	0.18
	Triglycerides	31303	0.012	-0.0418 (0.0371)	0.26	1
	Adiponectin	10036	0.013	-0.0320 (0.0290)	0.27	1
Glycemic Traits	Fasting Glucose	22015	0.011	0.0163 (0.0468)	0.73	1
	Fasting Insulin	21792	0.011	0.286 (0.0473)	1.5E-09	3.5E-08
	2 hour Glucose	16715	0.0119	0.0717 (0.0952)	0.40	1
	2 Hour Insulin	14150	0.0121	0.2337 (0.0435)	7.86E-08	1.8E-06
	Matsuda index *	8566	0.012	-0.3448 (0.0709)	1.2E-06	2.8E-05
Blood Pressure Traits	Systolic blood pressure	31840	0.012	0.0115 (0.0384)	0.77	1
	Diastolic blood pressure	31840	0.012	0.0705 (0.0384)	0.07	1

N: sample size contributing to association
MAF: minor allele frequency
Effect (Std. Err): regression estimate of the additive genetic effect and standard error of the estimate
P: P-value testing the significance of the association
Padjusted: A Bonferroni P value correction for 23 tests was applied

Supplementary Table 4B: T2D and hypertension association analysis with AKT2 p.Pro50Thr. These analyses was performed in a staged meta-analysis modeling the approach taken in the discovery and replication of the FI association with AKT2 p.Pro50Thr, with the European exome sequence data, the Finnish exome chip cohorts and the Finnish replication cohorts.

Outcome	Adjustment	Genotypes in Cases / Controls	MAF	N	Odds Ratio (95% CI)	P	Padjusted
Type 2 Diabetes	BMI	9554/224/5 22223/437/2	0.01	32421	1.05 (1.01, 1.09)	8.10E-05	0.0019
	Unadjusted	14180/306/5 17691/357/2	0.01	32578	1.05 (1.01, 1.09)	9.80E-04	0.022
Hypertension	BMI	34963/846/12 17765/371/3	0.011	53960	1.03 (0.98, 1.08)	0.31	1

Outcome: dichotomous outcome tested
Adjustment: indicates if BMI was used as a covariate in addition to sex and age.
MAF: minor allele frequency
Odds Ratio (95% CI): odds ratio estimate for increased risk of outcome and 95% confidence interval of the estimate
Padjusted: A Bonferroni P value correction for 23 tests was applied.

Supplementary Table 4C: Statistics for differences in HbA1c, fasting glucose, and fasting insulin distributions in the sample sub-cohorts with the *AKT2* P50T allele from the T2D-GENES whole exome sequencing data. Here, we provide genotype counts, median values of the scaled trait value, and tests difference in distributions using the non-parametric Kruskal-Wallis rank sum test and Monte Carlo permutation test.

		Control Group				Type 2 Diabetes Group				
Trait	Cohort	AKT2 P50T Genotype counts: 0/0; 0/1; 1/1	Median scaled trait value: 0/0; 0/1; 1/1	Kruskal-Wallis Test P	Monte Carlo Permutation Test P	AKT2 P50T Genotype counts: 0/0; 0/1; 1/1	Median scaled trait value: 0/0; 0/1; 1/1	Kruskal-Wallis Test P	Monte Carlo Permutation Test P	Percentile value for homozygous carrier (1/1)
HbA1c	METSIM	363; 10; 0	-0.15; -0.15; NA	0.78	0.88	465; 18; 1	-0.055; -0.06; 0.18	0.28	0.098	95%
Fasting Glucose	Botnia	220; 1; 0	-0.41; -0.33; NA	0.38	0.52	0; 0; 0				
	FUSION	467; 9; 0	-0.32; -0.43; NA	0.12	0.12	0; 0; 0				
	METSIM	486; 12; 0	-0.28; -0.22; NA	0.016	0.071	465; 18; 1	0.41; 0.60; 4.6	0.06	0.002	99.8%
Fasting Insulin	Botnia	205; 1; 0	-0.35; -0.30; NA	0.82	0.91	0; 0; 0				
	FUSION	464; 9; 0	1.1; 0.96; NA	0.86	0.46	0; 0; 0				
	METSIM	485; 12; 0	-0.49; -0.44; NA	0.32	0.56	465; 18; 1	-0.17; -0.29; 5.3	0.17	0.017	98.8%

Genotype categories: 0/0 indicates the group of individuals who are homozygote for the reference allele at rs184042322 (C/C); 0/1 indicates the group of individuals who are heterozygote at rs184042322 (C/T); 1/1 indicates the group of individuals who are homozygote for the AKT2 p.Pro50Thr allele at rs184042322 (T/T).

SUPPLEMENTARY TABLE 5

Phenotype exploration of *AKT2* p.Pro50Thr carriers electronic medical records.
Phenotype exploration of *AKT2* p.Pro50Thr carriers electronic medical records were queried in two cohorts for diseases plausibly related to AKT2. The genotype counts for the *AKT2* p.Pro50Thr variant are displayed for individuals not coded for an outcome (Controls) and individuals coded for an outcome (Cases). * Other related phenotype outcome included Lipodystrophy (E88.1), Acanthosis nigricans (L83), and Malignant neoplasm of male breast (C50.*2). No cases were reported for these outcomes in both METSIM and FINRISK. ** ICD 10 codes are used to obtain diagnoses of the phenotype outcome from hospital discharge records or electronic health records.

			Genotype counts (GG/TG/TT)	
			Controls	Cases
Malignant neoplasm of digestive organs and peritoneum	C15 – C26	METSIM	8708/215/3	42/1/0
		FINRISK	8200/182/1	146/1/0
Malignant neoplasm of genitourinary organs	C55 – C68	METSIM	8620/213/3	130/3/0
		FINRISK	8154/180/1	192/3/0
Malignant neoplasm of female breast	C50.*1	FINRISK	4167/87/0	70/1/0
Ovaries, polycystic	E28.2	FINRISK	4236/88/0	1/0/0
Cyst of ovary, follicular	N83.0	FINRISK	4233/88/0	4/0/0

ICD = International Classification of Diseases
OR = Odds ratio
95% CI = 95% Confidence interval
METSIM = Metabolic Syndrome in Men Study
FINRISK = The National FINRISK Study

SUPPLEMENTARY TABLE 6

Aggregate test of variants in monogenic gene sets and in the Insulin Receptor Signaling Pathway.

Supplementary Table 6A: List of the genes in the monogenic gene sets and the Insulin Receptor Signaling Pathway.

Chr	Location	Gene	Monogenic diabetes classification	Monogenic All	Monogenic Glucose	Monogenic insulin	Insulin Receptor Signaling Pathway
1	1p12	*SLC16A1/MCT1*	hyperinsulinsim	1	1	1	0
1	1p21	*S1PR1*		0	0	0	1
1	1p22	*BCL10*		0	0	0	1
1	1p31	*LEPR*		1	0	0	0
1	1p32	*TAL1*		0	0	0	1
1	1p32-p31	*JUN*		0	0	0	1
1	1p34	*PTPRF*		0	0	0	1
1	1p34	*YBX1*		0	0	0	1
1	1p34	*ZMPSTE24*		1	1	1	0
1	1p34.1	*PIK3R3*		0	0	0	1
1	1p36.11	*SFN*		0	0	0	1
1	1p36.2	*MTOR*		0	0	0	1
1	1p36.2	*PIK3CD*		0	0	0	1
1	1p36.21	*CASP9*		0	0	0	1
1	1p36.33	*SKI*		0	0	0	1
1	1q21	*CLK2*		0	0	0	1
1	1q21	*MCL1*		0	0	0	1
1	1q21	*SHC1*		0	0	0	1
1	1q21	*THEM4*		0	0	0	1
1	1q22	*DAP3*		0	0	0	1
1	1q22	*LMNA*		1	1	1	0
1	1q23.3	*SLC19A2*	NDM	1	0	0	0
1	1q25	*NCF2*		0	0	0	1
1	1q25.2-q25.3	*PTGS2*		0	0	0	1
1	1q32	*PIK3C2B*		0	0	0	1
2	2p12	*EIF2AK3*	NDM	1	1	0	0
2	2p13	*ALMS1*	syndromic	1	1	1	0
2	2p13	*HK2*		0	0	0	1
2	2p16.1	*CCDC88A*		0	0	0	1
2	2p21	*RHOQ*		0	0	0	1
2	2p23.3	*POMC*		1	0	0	0
2	2p25	*KLF11*	MODY7	1	0	0	0
2	2q12.3	*LIMS1*		0	0	0	1
2	2q31.1	*BBS5*		1	0	0	0
2	2q31.1	*C2ORF37/DCAF17*		1	0	0	0
2	2q32	*NEUROD1*	MODY6 NDM	1	1	1	0
2	2q32.2	*STAT1*		0	0	0	1
2	2q34	*PIKFYVE*		0	0	0	1
2	2q36	*IRS1*		0	0	0	1
3	3p21	*CTNNB1*		0	0	0	1
3	3p21.3	*USP4*		0	0	0	1
3	3p25	*PPARG*		1	1	1	0
3	3p25	*RAF1*		0	0	0	1
3	3p25.3	*CIDEC*		1	0	0	0
3	3q11.2	*ARL6*		1	0	0	0
3	3q13.3	*GSK3B*		0	0	0	1
3	3q21	*NCK1*		0	0	0	1
3	3q22.1	*TOPBP1*		0	0	0	1
3	3q22.3	*PIK3CB*		0	0	0	1
3	3q26.1-q26.2	*SLC2A2/GLUT2*	NDM	1	1	1	0
3	3q26.3	*PIK3CA/PI3K*		1	1	1	1
4	4p15.1	*PPARGC1A*		0	0	0	1
4	4p16.1	*WFS1*	NDM	1	1	1	0
4	4p16.3	*HTT*		0	0	0	1
4	4q22-q26	*HADH*	hyperinsulinsim	1	1	1	0
4	4q23	*EIF4E*		0	0	0	1
4	4q24	*CISD2 (WFS2)*		1	1	1	0
4	4q25	*SEC24B*		0	0	0	1
4	4q27	*BBS12*		1	0	0	0
4	4q27	*BBS7*		1	0	0	0
4	4q31.21	*GAB1*		0	0	0	1
4	4q34	*CASP3*		0	0	0	1
4	4q35.1	*SORBS2*		0	0	0	1
5	5p12	*PRKAA1*		0	0	0	1
5	5p15.33	*TERT*		0	0	0	1
5	5q11.1	*ISL1*		1	0	0	0
5	5q13.1	*PIK3R1*		1	1	1	1
5	5q13.3	*RASA1*		0	0	0	1
5	5q15-q21	*PCSK1*		1	0	0	0
5	5q31	*SMAD5*		0	0	0	1
5	5q32	*SPINK1/PST1*		1	0	0	0
5	5q33	*HAND1*		0	0	0	1
5	5q35.1	*NPM1*		0	0	0	1
6	6p21	*RUNX2*		0	0	0	1
6	6p21.1	*SRF*		0	0	0	1
6	6p21.2	*CDKN1A*		0	0	0	1
6	6p21.31	*POU5F1*		0	0	0	1
6	6p22.1	*ZFP57*	NDM	1	0	0	0
6	6p25	*FOXC1*		0	0	0	1
6	6q21	*FOXO3*		0	0	0	1
6	6q21	*FYN*		0	0	0	1
6	6q22.1	*RFX6*	NDM	1	1	1	0
6	6q22.31	*GJA1*		0	0	0	1
6	6q22.33	*MAP3K5*		0	0	0	1
6	6q23	*SGK1*		0	0	0	1
6	6q24-q25	*PLAGL1*	NDM	1	0	0	0
6	6q24.2	*HYMAI*	NDM	1	0	0	0
6	6q25.1	*ESR1*		0	0	0	1
6	6q26	*IGF2R*		0	0	0	1
6	6q27	*MLLT4*		0	0	0	1
7	7p12	*EGFR*		0	0	0	1
7	7p12.2	*GRB10*		0	0	0	1
7	7p14	*BBS9*		1	0	0	0
7	7p15.3-p15.1	*GCK*	MODY2 NDM	1	1	1	0
7	7p21.2	*TWIST1*		0	0	0	1
7	7p22	*RAC1*		0	0	0	1
7	7q11.23	*NCF1*		0	0	0	1
7	7q22	*DLX5*		0	0	0	1
7	7q22	*SH2B2*		0	0	0	1
7	7q22-q31.1	*SRPK2*		0	0	0	1
7	7q22.1	*COPS6*		0	0	0	1
7	7q22.3	*PIK3CG*		0	0	0	1
7	7q31.1	*CAV1*		1	1	1	0
7	7q31.1	*PPP1R3*		1	1	1	0
7	7q31.2	*CFTR*		1	0	0	0
7	7q31.3	*LEP*		1	0	0	0
7	7q32	*PAX4*	MODY9	1	0	0	0
7	7q34	*BRAF*		0	0	0	1
7	7q34	*PRSS1*		1	0	0	0
7	7q36	*MNX1*	NDM	1	1	1	0
7	7q36	*NOS3*		0	0	0	1
7	7q36	*RHEB*		0	0	0	1
8	8p11	*KAT6A*		0	0	0	1
8	8p12	*EIF4EBP1*		0	0	0	1
8	8p12	*WRN/RECQL2*		1	1	0	0
8	8p21.1	*PTK2B*		0	0	0	1
8	8p22-p21	*DPYSL2*		0	0	0	1
8	8p23-p22	*BLK*	MODY11	1	0	0	0
8	8p23.1-p22	*GATA4*	NDM	1	1	1	0

Chr	Location	Gene	Monogenic diabetes classification	Monogenic All	Monogenic Glucose	Monogenic insulin	Insulin Receptor Signaling Pathway
8	8q22.2	*STK3*		0	0	0	1
8	8q23.1	*YWHAZ*		0	0	0	1
8	8q24.3	*NDRG1*		0	0	0	1
8	8q24.3	*PTK2*		0	0	0	1
9	9p21	*RPS6*		0	0	0	1
9	9p24.2	*GLIS3*	NDM	1	1	1	0
9	9q33.1	*TRIM32/BBS11*		1	0	0	0
9	9q33.3	*MAPKAP1*		0	0	0	1
9	9q34	*TSC1*		0	0	0	1
9	9q34.3	*AGPAT2*		1	0	0	0
9	9q34.3	*CEL*	MODY8	1	1	0	0
9	9q34.3	*RAPGEF1*		0	0	0	1
10	10p11.23	*BMI1*		0	0	0	1
10	10p11.23	*MAP3K8*		0	0	0	1
10	10p12.2	*PTF1A*	NDM	1	1	1	0
10	10q11.22	*MAPK8*		0	0	0	1
10	10q21.3	*NEUROG3*	NDM	1	1	1	0
10	10q21.3	*SIRT1*		1	0	0	0
10	10q23.3	*GLUD1*	hyperinsulinsim	1	1	1	0
10	10q23.3	*PTEN*		1	1	1	1
10	10q24-q25	*CHUK*		0	0	0	1
11	11p11.2	*MAPK8IP1*		0	0	0	1
11	11p13	*PAX6*	NDM	1	0	0	0
11	11p15	*ARFIP2*		0	0	0	1
11	11p15.1	*ABCC8*	MODY NDM	1	1	1	0
11	11p15.1	*KCNJ11*	MODY NDM	1	1	1	0
11	11p15.1	*PDE3B*		0	0	0	1
11	11p15.4	*ILK*		0	0	0	1
11	11p15.5	*CDKN1C*		0	0	0	1
11	11p15.5	*INS*	MODY10 NDM	1	1	1	0
11	11p15.5-p14	*PIK3C2A*		0	0	0	1
11	11q13	*BBS1*		1	0	0	0
11	11q13	*BSCL2*		1	1	1	0
11	11q13	*CCND1*		0	0	0	1
11	11q13	*RELA*		0	0	0	1
11	11q13	*UCP2*	hyperinsulinsim	1	1	1	0
11	11q13	*YAP1*		0	0	0	1
11	11q13.1	*BAD*		0	0	0	1
11	11q13.1-q13.3	*MAP3K11*		0	0	0	1
11	11q23.3	*CBL*		0	0	0	1
11	11q24.2	*CHEK1*		0	0	0	1
12	12p12	*PIK3C2G*		0	0	0	1
12	12p13.1-p12	*CDKN1B*		0	0	0	1
12	12p13.31	*NANOG*		0	0	0	1
12	12q12-q14	*PRKAG1*		0	0	0	1
12	12q13	*NR4A1*		0	0	0	1
12	12q13.1	*SP1*		0	0	0	1
12	12q14.3-q15	*MDM2*		0	0	0	1
12	12q21.2	*BBS10*		1	0	0	0
12	12q21.32	*CEP290*		1	0	0	0
12	12q23.2	*IGF1*		0	0	0	1
12	12q24	*PTPN11*		0	0	0	1
12	12q24.1-q24.3	*PRKAB1*		0	0	0	1
12	12q24.2	*HNF1A*	MODY3	1	1	0	0
12	12q24.31	*PXN*		0	0	0	1
12	12q24.33	*CHFR*		0	0	0	1
13	13q12.1	*PDX1/IPF1*	MODY4 NDM	1	1	1	0
13	13q13.1	*STARD13*		0	0	0	1
13	13q14.1	*FOXO1*		0	0	0	1
13	13q14.2	*RB1*		0	0	0	1
13	13q22.2	*TBC1D4*		0	0	0	1
13	13q34	*IRS2*		0	0	0	1
14	14q11.2	*NDRG2*		0	0	0	1
14	14q12	*LTB4R2*		0	0	0	1
14	14q13	*NFKBIA*		0	0	0	1
14	14q23.2	*HIF1A*		0	0	0	1
14	14q24	*SRSF5*		0	0	0	1
14	14q24.3	*FOS*		0	0	0	1
14	14q31.3	*TTC8/BBS8*		1	0	0	0
14	14q32.32	*AKT1*		0	0	0	1
15	15q	*NEDD4*		0	0	0	1
15	15q21	*MYO5A*		0	0	0	1
15	15q21.2	*USP8*		0	0	0	1
15	15q22.3-q23	*BBS4*		1	0	0	0
15	15q22.33	*SMAD3*		0	0	0	1
15	15q24.1	*EDC3*		0	0	0	1
15	15q26	*PLIN*		1	1	1	0
15	15q26.3	*IGF1R*		0	0	0	1
16	16p11.2	*SH2B1*		1	0	0	0
16	16p11.2	*STX4*		0	0	0	1
16	16p13.3	*TSC2*		0	0	0	1
16	16q21	*BBS2*		1	0	0	0
17	17p11.2	*SREBF1*		0	0	0	1
17	17p12	*MAP2K4*		0	0	0	1
17	17p13	*SLC2A4*		0	0	0	1
17	17p13.1	*PIK3R5*		0	0	0	1
17	17p13.1	*PIK3R6*		0	0	0	1
17	17p13.1	*TP53*		0	0	0	1
17	17p13.1	*VAMP2*		0	0	0	1
17	17p13.3	*YWHAE*		0	0	0	1
17	17q12	*HNF1B*	MODY5 NDM	1	1	1	0
17	17q21	*BRCA1*		0	0	0	1
17	17q21.1	*MAPT*		0	0	0	1
17	17q21.2	*ACLY*		0	0	0	1
17	17q21.2	*PTRF*		1	1	1	0
17	17q21.31	*STAT3*		0	0	0	1
17	17q22	*MKS1*		1	0	0	0
17	17q22	*SRSF1*		0	0	0	1
17	17q22	*STXBP4*		0	0	0	1
17	17q23.1	*RPS6KB1*		0	0	0	1
17	17q24-q25	*GRB2*		0	0	0	1
17	17q25.3	*RPTOR*		0	0	0	1
17	17q25.3	*SOCS3*		0	0	0	1
18	18q11.1-q11.2	*GATA6*	NDM	1	1	1	0
18	18q12	*IER3IP1*	NDM	1	0	0	0
18	18q21.3	*BCL2*		0	0	0	1
18	18q22	*MC4R*		1	0	0	0
19	19p13.11	*GDF15*		0	0	0	1
19	19p13.2	*CDC37*		0	0	0	1
19	19p13.3	*STK11*		0	0	0	1
19	19p13.3	*TRIP10*		0	0	0	1
19	19p13.3-p13.2	*INSR*		1	1	1	1
19	19q13.1-q13.2	*AKT2*		1	1	1	1
19	19q13.12	*NFKBID*		0	0	0	1
19	19q13.2	*GSK3A*		0	0	0	1
19	19q13.2	*LIPE*		0	0	0	1
19	19q13.2-q13.4	*PIK3R2*		0	0	0	1
19	19q13.3	*DMPK*		1	0	0	0
19	19q13.3	*POLD1*		1	1	1	0
19	19q13.3-q13.4	*BAX*		0	0	0	1
19	19q13.3-q13.4	*IRF3*		0	0	0	1
19	19q13.33	*AKT1S1*		0	0	0	1
20	20p12	*MKKS*		1	0	0	0
20	20q11.2-q13.2	*STK4*		0	0	0	1
20	20q11.21	*BCL2L1*		0	0	0	1
20	20q12-q13	*SRC*		0	0	0	1
20	20q13.1-q13.2	*PTPN1*		0	0	0	1
20	20q13.12	*HNF4A*	MODY1	1	1	1	0
20	20q13.2	*SGK2*		0	0	0	1
20	20q13.31	*RBM38*		0	0	0	1
20	20q13.33	*DNAJC5*		0	0	0	1
21	21q22.3	*AIRE*		1	0	0	0
21	21q22.3	*PCNT*		1	0	0	0
23	Xp11.23	*FOXP3*	NDM	1	0	0	0
NA	NA	*C8orf44-SGK3/SGK3*		0	0	0	1
X	Xp11.2	*ELK1*		0	0	0	1
X	Xq13.1	*FOXO4*		0	0	0	1
X	Xq22.3	*IRS4*		0	0	0	1

Supplementary Table 6B: Global test of monogenic genes from exome chip analysis. Aggregate tests of rare variants based on functional annotation were performed using exome array variants in all the genes in each gene set. We performed conditional analyses to understand the variants contributing to the significant association signals.

Trait	Gene set	Test	PTV	PTV+NS$_{strict}$	PTV+NS$_{broad}$	PTV+Missense
Fasting Insulin	All Monogenic	SKAT	0.275	0.494	0.014*	0.028
		BURDEN	0.972	0.012	**0.00024*** **	0.019
	Monogenic Insulin	SKAT	0.173	0.618	0.002*	0.011
		BURDEN	0.136	0.147	0.001*	0.01
	Insulin Receptor Signaling Pathway	SKAT	0.361901	0.826451	0.011	**0.00066**** **
		BURDEN	0.595991	0.800962	0.278479	0.072434
Fasting Glucose	All Monogenic	SKAT	0.073	0.078	0.635	0.712
		BURDEN	0.00697**	0.131	0.041	0.375
	Monogenic Glucose	SKAT	0.073	0.026	0.224	0.189
		BURDEN	0.0098**	0.431	0.051	0.346

* After conditioning on *ATK2* p.Pro50Thr, the global test P values for the Monogenic gene set was P=0.38 (SKAT). For the Monogenic Insulin gene set, the conditional P values were P = 0.02 (SKAT) and P = 0.017 (BURDEN).
** After conditioning on *BSCL2* p.Q271*, the global test was P = 0.019 (BURDEN) for the Monogenic gene set and P = 0.039 (BURDEN) for the Monogenic Glucose gene set.
*** Conditional analysis of this test is presented in Supplementary Table 6C.
**** After conditioning on *AKT2* p.Pro50Thr, the global test P values for the Insulin Receptor Signaling Pathway was P=0.01.

Supplementary Table 6C: Global test of monogenic genes from exome sequencing analysis.

Trait	Gene set	Test	PTV	PTV+NS$_{strict}$	PTV+NS$_{broad}$	PTV+NS
Fasting Insulin	Monogenic	SKAT	0.25	0.15	0.15	0.48
		BURDEN	0.91	0.2	0.87	0.55
	Monogenic Insulin	SKAT	0.44	0.39	0.49	0.71
		BURDEN	0.95	0.31	0.05	0.62
	Insulin Receptor Signaling Pathway	SKAT	0.52	0.04	0.26	0.69
		BURDEN	0.61	0.04	0.79	0.12
Fasting Glucose	Monogenic	SKAT	0.49	0.93	0.82	0.6
		BURDEN	0.86	0.1	0.92	0.83
	Monogenic Glucose	SKAT	0.22	0.74	0.52	0.49
		BURDEN	0.97	0.5	0.96	0.33

Variant masks:
PTV: containing only variants predicted to introduce a premature stop codon
PTV+NS: containing variants in the PTV group and protein-altering variants with MAF<1%
PTV+NSstrict: composed of variants in "PTV" and protein-altering variants predicted damaging by SIFT, LRT, MutationTaster, polyphen2 HDIV, and polyphen2 HVAR
PTV+NSbroad: composed of "**PTV+NSstrict**" and NS variants with MAF<1% and predicted damaging by at least one prediction algorithm.

Supplementary Table 6D: Sequential conditional analysis of the exome chip global BURDEN test with the monogenic all gene set for FI with PTV + NSstrict + Nsbroad variants. Variants that contributed the most to the association, as reported by RAREMETALS v.4.7, were added to the model sequentially. Single variant association results of these variants are provided in **Supplementary Table 7B**.

Location	rsID	REF	ALT	Gene	Protein change	Global Test P value after conditioning
No conditioning						0.00024
19:40762860	rs184042322	G	T	*AKT2*	p.P50T	0.0017
7:117282582	rs11971167	G	A	*CFTR*	p.D1270N	0.0029
19:7125518	rs1799816	C	T	*INSR*	p.V1012M	0.0087
1:40756572	rs41268053	G	A	*ZMPSTE24*	p.R369Q	0.0089
6:29641139	rs199589695	G	A	*ZFP57*	p.R178H	0.0098
7:117171169	rs78756941	G	T	*CFTR*	Splice donor	0.0089
21:47831307	rs201709021	G	A	*PCNT*	p.E1785K	0.0104

Supplementary Table 6E: Association results of the variants contributing to the exome chip global burden test association of the "Monogenic" genes for FI level.

Location	rsID	REF	ALT	Gene	Protein change	Effect Allele; Effect allele frequency	Effect (Standard error)	BF	P	N
19:40762860	rs184042322	G	T	*AKT2*	p.P50T	T; 0.011	0.112 (0.023)	5.4	2.1×10^{-7}	28118
7:117282582	rs11971167	G	A	*CFTR*	p.D1270N	A; 0.008	0.143 (0.048)	1.7	1.5×10^{-3}	9898
19:7125518	rs1799816	C	T	*INSR*	p.V1012M	T; 0.01	0.065 (0.02)	1.1	5.4×10^{-3}	32685
1:40756572 *	rs41268053	G	A	*ZMPSTE24*	p.R369Q	-	-	-	7.1×10^{-3} **	-
6:29641139 *	rs199589695	G	A	*ZFP57*	p.R178H	-	-	-	7.2×10^{-3} **	-
7:117171169 *	rs78756941	G	T	*CFTR*	Splice donor	T; 0.001	-0.426 (0.161)	1	9.7×10^{-3}	4136
21:47831307 *	rs201709021	G	A	*PCNT*	p.E1785K	-	-	-	7.9×10^{-3} **	-

* Single variant association tests were not performed because variant did not meet the inclusion criteria (MAC > 5 within each cohort).
** P values from the RAREMETALS v.4.7 software.
BF: log10(Bayes factor) for association
P: P value for association test
N: Total Sample size contributing to analysis

SUPPLEMENTARY TABLE 7

Gene-based and single-variant association results from genes highlighted in the enrichment analyses.

Supplementary Table 7A: Gene based results of the monogenic genes or insulin receptor signaling genes exhibiting enrichment of association signals.

Fasting insulin		PTV+missense			PTV+NS_{broad}			PTV+NS_{strict}			PTV-only		
Gene	Ancestry	MAC (No. vars)	P value SKAT	P value Burden	MAC (No. vars)	P value SKAT	P value Burden	MAC (No. vars)	P value SKAT	P value Burden	MAC (No. vars)	P value SKAT	P value Burden
AKT2	AfrAm	1	0.67	0.67	1	0.67	0.67	-	-	-	3	0.043	0.52
19q13.1-q13.2	E.Asian	5	0.33	0.15	5	0.33	0.15	<1	0.65	0.65	1	0.95	0.95
	Europ	31	0.53	0.31	31	0.53	0.31	-	-	-	3	0.12	0.12
	Hisp	7	0.42	0.13	7	0.42	0.13	-	-	-	2	0.55	0.88
	S.Asian	2	0.86	0.83	1	0.6	0.6	-	-	-	-	-	-
	WES (all)	46(36)	0.6	0.051	45(33)	0.57	0.052	<1(5)	0.65	0.65	9(14)	0.083	0.52
	ExArray	398(4)	6.10E-07	3.60E-06	398(4)	6.10E-07	3.60E-06	-	-	-	5(2)	0.63	0.99
	WES (all) + ExArray	444	0.00056	7.30E-06	443	0.00048	7.50E-06	<1	0.65	0.65	14	0.23	0.96
INSR	AfrAm	29	0.43	0.98	20	0.29	0.79	1	0.75	0.75	1	0.75	0.75
19p13.3-p13.2	E.Asian	29	0.015	0.29	24	0.02	0.095	-	-	-	-	-	-
	Europ	42	0.46	0.76	35	0.42	0.89	1	0.73	0.73	-	-	-
	Hisp	7	0.48	0.68	6	0.66	0.26	-	-	-	-	-	-
	S.Asian	16	0.39	0.029	5	0.14	0.021	-	-	-	-	-	-
	WES (all)	123(127)	0.17	0.62	90(96)	0.12	0.99	2(9)	0.9	0.64	1(4)	0.75	0.75
	ExArray	767(10)	0.0066	0.035	667(6)	0.0074	0.033	-	-	-	-	-	-
	WES (all) + ExArray	890	0.0074	0.14	757	0.0055	0.61	2	0.9	0.64	1	0.75	0.75
ZMPSTE24	AfrAm	1	0.28	0.28	1	0.28	0.28	1	0.28	0.28	1	0.28	0.28
1p34	E.Asian	6	0.62	0.86	6	0.62	0.86	4	0.83	0.79	-	-	-
	Europ	10	0.35	0.65	9	0.54	0.84	6	0.54	0.53	5	0.42	0.52
	Hisp	8	0.75	0.49	8	0.75	0.49	5	0.53	0.87	4	0.49	0.74
	S.Asian	8	0.072	0.94	8	0.072	0.94	1	0.18	0.18	-	-	-
	WES (all)	33(51)	0.23	0.82	32(46)	0.3	0.56	17(22)	0.73	0.62	10(9)	0.54	0.74
	ExArray	8(2)	0.011	0.078	8(2)	0.011	0.078	-	-	-	-	-	-
	WES (all) + ExArray	41	0.016	0.36	40	0.024	0.18	17	0.73	0.62	10	0.54	0.74
CFTR	AfrAm	37	0.39	0.5	43	0.34	0.4	30	0.2	0.19	2	0.16	0.16
7q31.2	E.Asian	99	0.45	0.76	67	0.25	0.43	20	0.32	0.045	1	0.55	0.55
	Europ	179	0.27	0.26	109	0.17	0.7	52	0.35	0.41	7	0.98	0.57
	Hisp	107	0.015	0.66	74	0.0096	0.043	42	0.0073	0.074	-	-	-
	S.Asian	50	0.0021	0.92	41	0.0016	0.36	23	0.0039	0.13	2	0.23	0.8
	WES (all)	474(248)	0.031	0.36	335(216)	0.012	0.11	168(100)	0.011	0.027	12(27)	0.76	0.48
	ExArray	3410(54)	0.58	0.82	3851(50)	0.53	0.31	2140(25)	0.27	0.049	28(7)	0.076	0.34
	WES (all) + ExArray	3884	0.12	0.65	4186	0.063	0.11	2308	0.021	0.0057	40	0.3	0.38
ZFP57	AfrAm	30	0.45	0.74	5	1	0.93	-	-	-	-	-	-
6p22.1	E.Asian	74	0.58	0.42	1	0.21	0.21	-	-	-	-	-	-
	Europ	11	0.49	0.4	-	-	-	-	-	-	-	-	-
	Hisp	20	1	1	-	-	-	-	-	-	-	-	-
	S.Asian	6	0.15	0.24	4	0.093	0.093	-	-	-	-	-	-
	WES (all)	141(55)	0.77	0.76	10(17)	0.27	0.59	-	-	-	-	-	-
	ExArray	243(10)	0.65	0.41	4(1)	0.0077	0.0077	-	-	-	-	-	-
	WES (all) + ExArray	384	0.78	0.63	14	0.016	0.061	-	-	-	-	-	-

Fasting insulin		PTV+missense			PTV+NS$_{broad}$			PTV+NS$_{strict}$			PTV-only		
Gene	Ancestry	MAC (No. vars)	P value SKAT	P value Burden	MAC (No. vars)	P value SKAT	P value Burden	MAC (No. vars)	P value SKAT	P value Burden	MAC (No. vars)	P value SKAT	P value Burden
PCNT	AfrAm	129	0.31	0.34	36	0.28	0.057	3	0.043	0.52	-	-	-
21q22.3	E.Asian	252	0.51	0.85	92	0.64	0.61	1	0.95	0.95	-	-	-
	Europ	174	0.043	0.61	75	0.16	0.11	3	0.12	0.12	-	-	-
	Hisp	110	0.32	0.87	32	0.53	0.36	2	0.55	0.88	-	-	-
	S.Asian	40	0.99	0.54	18	0.88	0.52	-	-	-	-	-	-
	WES (all)	706(531)	0.14	0.94	254(230)	0.4	0.16	9(14)	0.083	0.52	-	-	-
	ExArray	3805(86)	0.58	0.65	2205(39)	0.88	0.98	5(2)	0.63	0.99	-	-	-
	WES (all) + ExArray	4511	0.26	0.91	2459	0.75	0.75	14	0.23	0.96	-	-	-
PTGS2	Afr. Amer.	2	0.74	0.74	-	-	-	-	-	-	-	-	-
1q25.2-q25.3	E.Asian	23	0.042	0.0062	4	0.27	0.29	-	-	-	-	-	-
	European	13	0.0024	0.0043	7	0.72	0.49	1	0.41	0.41	-	-	-
	Hispanic	6	0.29	0.39	-	-	-	-	-	-	-	-	-
	S.Asian	2	0.43	0.43	2	0.43	0.43	2	0.43	0.43	-	-	-
	all sequencing	46(31)	0.0041	0.00011	13(21)	0.64	0.16	3(5)	0.51	0.26	-	-	-
	ExArray	200(5)	0.71	0.28	110(2)	0.61	0.57	-	-	-	-	-	-
	WES (all) + ExArray	246	0.069	0.0013	123	0.68	0.28	3	0.51	0.26	-	-	-

Fasting glucose		PTV+missense			PTV+NS$_{broad}$			PTV+NS$_{strict}$			PTV-only		
Gene	Ancestry	MAC (No. vars)	P value SKAT	P value Burden	MAC (No. vars)	P value SKAT	P value Burden	MAC (No. vars)	P value SKAT	P value Burden	MAC (No. vars)	P value SKAT	P value Burden
BSCL2	AfrAm	10	0.77	0.46	4	0.66	0.48	-	-	-	-	-	-
11q13	E.Asian	26	0.15	0.16	23	0.17	0.18	2	0.026	0.026	2	0.026	0.026
	Europ	38	0.00072	0.00034	4	0.88	0.63	-	-	-	-	-	-
	Hisp	29	0.49	0.58	14	0.77	0.88	<1	0.41	0.41	<1	0.41	0.41
	S.Asian	12	0.6	0.16	8	0.36	0.058	1	0.9	0.9	1	0.9	0.9
	WES (all)	116(60)	0.0013	0.048	53(36)	0.4	0.74	3(5)	0.049	0.24	3(5)	0.049	0.24
	ExArray	574(13)	0.08	0.022	288(9)	0.021	0.0043	102(2)	0.033	0.0067	102(2)	0.033	0.0067
	WES (all) + ExArray	690	0.00088	0.0046	341	0.051	0.081	105	0.0068	0.012	105	0.0068	0.012
CAV1	AfrAm	2	0.14	0.05	2	0.14	0.05	1	0.095	0.095	-	-	-
7q31.1	E.Asian	5	0.027	0.23	5	0.027	0.23	4	0.022	0.1	-	-	-
	Europ	9	0.17	0.0065	6	0.13	0.018	3	0.17	0.064	2	0.36	0.36
	Hisp	11	0.098	0.1	11	0.098	0.1	-	-	-	-	-	-
	S.Asian	5	0.69	0.36	2	0.92	0.68	1	0.79	0.79	-	-	-
	WES (all)	32(18)	0.032	0.00017	26(16)	0.025	0.00065	9(8)	0.019	0.0051	2(1)	0.36	0.36
	ExArray	77(4)	0.31	0.35	77(4)	0.31	0.35	-	-	-	-	-	-
	WES (all) + ExArray	109	0.049	0.0025	103	0.041	0.0055	9	0.019	0.0051	2	0.36	0.36

MAC (No. vars): Minor allele count (number of variants in the test)
Variant masks:
PTV: containing only variants predicted to introduce a premature stop codon
PTV+NS: containing variants in the PTV group and protein-altering variants with MAF<1%
PTV+NSstrict: composed of variants in "PTV" and protein-altering variants predicted damaging by SIFT, LRT, MutationTaster, polyphen2 HDIV, and polyphen2 HVAR
PTV+NSbroad: composed of "**PTV+NSstrict**" and NS variants with MAF<1% and predicted damaging by at least one prediction algorithm.

Supplementary Table 7B: Single variant association results with FG levels from the monogenic genes exhibiting enrichment of association signals.

Gene set and Variant Group	Location	SNP	REF	ALT	Gene	Protein change	Inverse Normalized Effect (Standard error, Effect Allele, Effect allele frequency)	Untransformed Effect (Standard Error)	BF	P	N
Monogenic - PTV	11:62458267	rs149907021	G	A	*BSCL2*	p.Q271*	1.621 (0.39; A; 0.001)	0.844 (0.185)	3.3	3.3E-05	4513
Monogenic - PTV + Nsstrict	11:62458267	rs149907021	G	A	*BSCL2*	p.Q271*	1.621 (0.39; A; 0.001)	0.844 (0.185)	3.3	3.3E-05	4513
	7:33545217	rs61764068	A	T	*BBS9*	p.E753V	-0.576 (0.19; A; 0.998)	-0.27 (0.086)	1.6	2.4E-03	8754
Monogenic - PTV + Nsstrict + Nsbroad	11:62458267	rs149907021	G	A	*BSCL2*	p.Q271*	1.621 (0.39; A; 0.001)	0.844 (0.185)	3.3	3.3E-05	4513
	2:73786157	rs34398445	G	C	*ALMS1*	p.K3423N	0.673 (0.188; C; 0.018)	0.221 (0.065)	2.3	3.4E-04	5935
	3:170715865	rs140138702	G	C	*SLC2A2*	p.L468V	0.641 (0.197; C; 0.012)	0.267 (0.081)	1.7	1.2E-03	1104
	7:33545217	rs61764068	A	T	*BBS9*	p.E753V	-0.576 (0.19; A; 0.998)	-0.27 (0.086)	1.6	2.4E-03	8754
Monogenic glucose - PTV+Nsstrict+Nsbroad	11:66287196	rs35520756	G	A	*BBS1*	p.E234K	0.275 (0.095; A; 0.102)	0.086 (0.032)	1.4	3.8E-03	1352
	11:62458267	rs149907021	G	A	*BSCL2*	p.Q271*	1.621 (0.39; A; 0.001)	0.844 (0.185)	3.3	3.3E-05	4513
	2:73786157	rs34398445	G	C	*ALMS1*	p.K3423N	0.673 (0.188; C; 0.018)	0.221 (0.065)	2.3	3.4E-04	5935
	3:170715865	rs140138702	G	C	*SLC2A2*	p.L468V	0.641 (0.197; C; 0.012)	0.267 (0.081)	1.7	1.2E-03	1104
Monogenic glucose - PTV+Missense	11:62458267	rs149907021	G	A	*BSCL2*	p.Q271*	1.621 (0.39; A; 0.001)	0.844 (0.185)	3.3	3.3E-05	4513
	2:73786157	rs34398445	G	C	*ALMS1*	p.K3423N	0.673 (0.188; C; 0.018)	0.221 (0.065)	2.3	3.4E-04	5935
	3:170715865	rs140138702	G	C	*SLC2A2*	p.L468V	0.641 (0.197; C; 0.012)	0.267 (0.081)	1.7	1.2E-03	1104
	9:4286344	rs113754532	T	C	*GLIS3*	p.I28V	0.418 (0.144; T; 0.998)	0.213 (0.071)	1.2	3.6E-03	19883
	2:73677876	var_2_73677876	G	A	*ALMS1*	p.V1407I	-0.949 (0.357; A; 0.001)	-0.44 (0.169)	1.2	7.8E-03	4513

BF: log10(Bayes factor) for association
P: P value for association test
N: Total Sample size contributing to analysis

SUPPLEMENTARY TABLE 8

GTEx tissue differential expression of AKT2 compared to AKT1 and AKT3. Listed are the tissues from the GTEx project pilot phase release where AKT2 expression was assessed.

Tissue abbreviation *	Tissue description **	N	P (AKT2 > AKT1)	P (AKT2 > AKT3)
ADPSBQ	Adipose - Subcutaneous	94	1	5.08×10^{-15}
ADPVSC	Adipose - Visceral (Omentum)	19	1	2.74×10^{-3}
ADRNLG	Adrenal Gland	12	1	5.37E-10
ARTAORT	Artery - Aorta	24	1	0.03
ARTCRN	Artery - Coronary	9	1	0.8
ARTTBL	Artery - Tibial	112	1	1
BREAST	Breast - Mammary Tissue	27	1	2.12×10^{-6}
BRNACC	Brain - Anterior cingulate cortex (BA24)	17	1	1
BRNAMY	Brain - Amygdala	23	1	1
BRNCDT	Brain - Caudate (basal ganglia)	36	1	0.12
BRNCHA #	Brain - Cerebellum	30	3.04×10^{-7}	8.94×10^{-17}
BRNCHB #	Brain - Cerebellar Hemisphere	24	6.60×10^{-4}	2.41×10^{-9}
BRNCTXA	Brain - Cortex	23	1	1
BRNCTXB	Brain - Frontal Cortex (BA9)	24	1	1
BRNHPP	Brain - Hippocampus	24	1	0.99
BRNHPT	Brain - Hypothalamus	23	1	0.99
BRNNCC	Brain - Nucleus accumbens (basal ganglia)	28	1	2.15×10^{-3}
BRNPTM	Brain - Putamen (basal ganglia)	20	1	0.02
BRNSNG	Brain - Substantia nigra	25	1	0.67
BRNSPC	Brain - Spinal cord (cervical c-1)	16	1	0.16
CLNTRN	Colon - Transverse	12	1	2.24×10^{-5}
ESPMCS	Esophagus - Mucosa	18	1	3.13×10^{-12}
ESPMSL	Esophagus - Muscularis	20	1	3.39×10^{-3}
FIBRBLS	Cells - Transformed fibroblasts	14	1	1.78×10^{-4}
HRTAA	Heart - Atrial Appendage	25	1	1.45×10^{-9}
HRTLV	Heart - Left Ventricle	83	1	9.20×10^{-53}
KDNCTX	Kidney - Cortex	3	0.71	0.1
LCL	Cells - EBV-transformed lymphocytes	39	1	$1.74\times10^{-}$
LIVER	Liver	5	0.97	6.56×10
LUNG	Lung	119	1	$5.24\times10^{-}$
MSCLSK #	Muscle - Skeletal	138	1.47×10^{-19}	$7.76\times10^{-}$
NERVET	Nerve - Tibial	88	1	$3.19\times10^{-}$
OVARY	Ovary	6	0.53	4.03×10
PNCREAS	Pancreas	19	1	$1.19\times10^{-}$
PRSTTE	Prostate	9	1	2.38×10
PTTARY #	Pituitary	13	0.03	$8.55\times10^{-}$
SKINNS	Skin - Not Sun Exposed (Suprapubic)	23	1	$1.05\times10^{-}$
SKINS	Skin - Sun Exposed (Lower leg)	96	1	$1.99\times10^{-}$
STMACH	Stomach	12	1	3.64×10
TESTIS	Testis	14	0.84	2.87×10
THYROID	Thyroid	105	0.13	$7.22\times10^{-}$
UTERUS	Uterus	7	0.99	0.0
VAGINA	Vagina	6	0.99	1.09×10
WHLBLD	Whole Blood	156	1	1.43×10^{-1}

N = sample size per tissue; P(*AKT2* > *AKT1*) = P value for the test of expression in *AKT2* compared to *AKT1*; P(*AKT2* > *AKT3*) = P value for the test of expression in *AKT2* compared to *AKT3*. * The tissue abbreviation used in Fig. S13 and Fig. S14. ** The corresponding tissue description. *** The one-sided paired t-test P-values for the comparison of *AKT2* expression with *AKT1* and *AKT3*. # The tissues where *AKT2* expression is significantly (P < 0.05) higher than both *AKT1* and *AKT3* expression. BRNCHA/BRNCHB and BRNCTXA/BRNCTXB are sampled from the same regions, cerebellum and cortex, respectively, but in separate collections.

SUPPLEMENTARY TABLE 9

Expression analyses in adipose tissue in the METSIM, EuroBATS and GTEx studies.

Supplementary Table 9A: The associations of the two eSNPs discovered in METSIM (rs8104727) and EuroBATS (rs11880261) with *AKT2* transcript levels. Results are presented for all the three cohorts queried (METSIM, EuroBATS and GTEx). The eSNPs are in linkage disequilibrium: R2 = 0.847 and D' = 0.92 in 1000 Genomes European population samples and R2 = 1 and D' = 1 in 1000 Genomes Finnish population samples.

GeneID	Cohort	Tissue	N	SNP	SNP origin	Effect allele	Other allele	EAF	Beta effect	SE	P-value (SNP-AKT2)
AKT2	GTEx	Adipose Subcutaneous	94	rs11880261	EuroBATS eSNP	T	C	0.25	0.186	0.103	7.56E-02
AKT2	EuroBATS	Adipose	720	rs11880261	EuroBATS eSNP	T	C	NA	0.206	0.037	2.27E-08
AKT2	METSIM	Adipose	770	rs8104727	METSIM eSNP	T	C	0.35312	0.4026	0.05214	3.595E-14
AKT2	METSIM	Adipose	770	rs11880261	EuroBATS eSNP	T	C	0.35239	0.3983	0.05219	6.882E-14

Supplementary Table 9B: Associations of the *AKT2* eSNPs with FI are displayed for the METSIM and EuroBATS studies.

GeneID	Cohort	N	SNP	SNP origin	Effect allele	Other allele	Adjustment	Effect	SE	P-value (eSNP-FI)
AKT2	METSIM	10081	rs8104727	METSIM eSNP	T	C	Age, BMI	-0.016	0.01523	0.2857
AKT2	METSIM	10081	rs11880261	EuroBATS eSNP	T	C	Age, BMI	-0.017	0.01527	0.2661
AKT2	EuroBATS	710	rs11880261	EuroBATS eSNP	T	C	Age, BMI	-0.015	0.0555131	0.7842
AKT2	METSIM	10081	rs8104727	METSIM eSNP	T	C	Age	-0.00088	0.01523	0.9541
AKT2	METSIM	10081	rs11880261	EuroBATS eSNP	T	C	Age	-0.0011	0.01527	0.9436
AKT2	EuroBATS	710	rs11880261	EuroBATS eSNP	T	C	Age	-0.0094	0.05497855	0.8649

Supplementary Table 9C: Associations of *AKT2* expression with FI are shown for the METSIM and EuroBATS studies.

GeneID	Cohort	N	Adjustment	Effect	SE	P-value (AKT2-FI)
AKT2	METSIM	770	Age, BMI	-0.33	0.07	0.00000949
AKT2	METSIM	770	Age	-0.42	0.06	3.293E-11
AKT2	EuroBATS	710	Age, BMI	-0.05	0.11	6.28E-04
AKT2	EuroBATS	710	Age	-0.04	0.01	1.14E-03

Supplementary Table 9D: The association between *AKT2* expression and age was queried in adipose tissue in the METSIM, EuroBATS and GTEx cohorts.

GeneID	Study	Tissue	N	ChiSq (age)	P-value (age)	Effect (age)
AKT2	METSIM	Adipose	770	8.46	0.00362	0.02
AKT2	EuroBATS	Adipose	720	0.143	0.71	0.001
AKT2	GTEx	Adipose Subcutaneous	89	3.49	0.06	-0.02

Supplementary Table 9E: The association between *AKT2* expression and BMI was queried in adipose tissue in the METSIM, EuroBATS and GTEx cohorts.

GeneID	Study	Tissue	N	ChiSq (BMI)	P-value (BMI)	Effect (BMI)
AKT2	METSIM	Adipose	770	28.772	8.143E-08	-0.06
AKT2	EuroBATS	Adipose	720	120.07	6.10E-28	-0.07
AKT2	GTEx	Adipose Subcutaneous	89	0.30	0.58	-0.01

NA: The data was not available
GeneID: The name of the gene investigated
Cohort: The cohort the association was studied in
Tissue: The tissue the expression data is from
N: The sample size in analysis
SNP: The rsID of the SNP for which the association is shown
SNP origin: The cohort where the SNP was most associated with AKT2 expression
Effect allele and Other allele: The effect and non-effect alleles of the SNP
EAF: The frequency of the effect allele

Beta effect: The effect estimate for the effect allele
SE: Standard error for the effect estimate
P-value (SNP-AKT2): The P-value for the SNP-expression association
Study: Study in which the association was studied
Adjustment: The covariate adjustment for fasting insulin
P-value (eSNP-FI): The P-value for the SNP-fasting insulin association
P-value (AKT2-FI): The P-value for the gene-fasting insulin association
ChiSq (age): Chi squared test statistic for the expression-age association
P-value (age): P-value for the SNP-expression association

Effect (age): Effect estimate for the age in the model
ChiSq (BMI): Chi squared test statistic for the expression-BMI association
P-value (BMI): P-value for the SNP-expression association
Effect (BMI): Effect estimate for the BMI in the model

SUPPLEMENTARY TABLE 10

Mendelian randomization analysis to assess the causality of *AKT2* expression for fasting insulin (FI) levels.

The results from the meta-analysis of the EuroBATS and METSIM data and for the instrumental variable (IV) estimator are shown for the EuroBATs eSNPs (rs11880261) additionally separated by whether BMI adjustment was used for SNP-FI and *AKT2*-FI analyses.

		No BMI adjustment				BMI adjusted			
Association	N	Effect	SE	P-value	P-value for difference	Effect	SE	P-value	P-value for difference
SNP-AKT2	1490	0.270	0.030	1.89E-19		0.270	0.030	1.89E-19	
SNP-FI	10791	-0.002	0.014	9.13E-01		-0.017	0.014	2.44E-01	
AKT2-FI	1480	-0.050	0.011	4.39E-06		-0.064	0.013	5.95E-07	
IV		-0.006	0.054	9.13E-01	0.41	-0.063	0.054	2.48E-01	0.99

Association: The pair of traits tested or the instrumental variable (IV)
N: The sample size in meta-analysis
Effect: The effect estimate in the association
SE: Standard error
P-value: The P-value for the association
P-value for difference: The P-value for the difference between the IV estimator and the *AKT2*-FI estimate

Ethics Statements

All human research was approved by the relevant institutional review boards, and conducted according to the Declaration of Helsinki and all patients provided written informed consent. FIN-D2D 2007, DPS, DR's EXTRA, FINRISK 2007, FUSION, and METSIM were approved by the University of Michigan Health Sciences and Behavioral Sciences Institutional Review Board (ID: H03-00001613-R2). The Danish studies (Health 2006, Inter99, and Vejle Biobank) were approved by the local Ethical Committees of Capital Region (approval # H-3-2012-155, KA 98155 and KA-20060011) and Region of Southern Denmark (approval # S-20080097). The GoDARTS study was approved by EoS REC 09/S1402/44. The Twins UK study was approved by EC04/015. The OBB study was approved by South Central, Oxford C, 08/H0606/107+5, IRAS project 136602. The PIVUS study is approved by 00–419 and ULSAM study by 251/90 and 2007/338. The PPP study was approved by the Committee On the Use of Humans as Experimental Subjects at MIT (IRB 0912003615). T2D-GENES and GoT2D exome sequencing was approved by local institutional review boards. The study protocol of the Health 2000 survey was approved by the Epidemiology Ethics Committee of the Hospital District of Helsinki and Uusimaa. All participants gave signed informed consent. The YFS study was approved by local ethics committees. The HBCS study was approved by the Ethics Committee of Hospital District of Helsinki and Uusimaa and conducted according to the guidelines in the Declaration of Helsinki. The EuroBATS study was approved by St Thomas' Hospital Research Ethics Committee (ref. EC04/015).

Additional Acknowledgements

Individuals (in author order):

Alisa K. Manning was supported by American Diabetes Association grant #7-12-MN-02. Taru Tukiainen was supported by Orion-Farmos Research Foundation and the Finnish Cultural Foundation. Manuel A Rivas received the NDM Prize Studentship, Clarendon Award. Tune H Pers is supported by the Benzon Foundation and the Lundbeck Foundation.

Ana Viñuela has been funded by the EU FP7 grant EuroBATS (Grant No. 259749).

Andrew Anand Brown has been funded by the EU FP7 grant EuroBATS (Grant No. 259749) and by the South East Norway Health Authority (Grant No. 2011060). Eric. R. Gazamon was supported by NIH Grants for GTEx: R01 MH101820 and MH090937. Hae Kyung Im was in part funded by R01MH107666 and K12CA139160, and travel was funded by P30DK020595. John R B Perry was supported by the Sir Henry Wellcome Postdoctoral Fellowship. Martijn van de Bunt is supported by the NDM Prize Studentship. Martin Hrabe de Angelis was supported by the German Center for Diabetes Research (DZD). Reedik Magi was supported by the Estonian Research Council (grant IUT20-60), the Development Fund of the University of Tartu (grant SP1GVARENG), EU structural support through Archimedes Foundation, grant no: 3.2.1001.11-0033, EU 7FP grant 278913, and H2020 grants 633589, 676550, 654248. Panos Deloukas's work forms part of the research themes contributing to the translational research portfolio of Barts Cardiovascular Biomedical Research Unit, which is supported and funded by the National Institute for Health Research. Katharine R Owen is a NIHR Clinician Scientist. Andrew Farmer is a NIHR Senior Investigator. Gilean McVean is a Wellcome Trust Senior Investigator. Eleftheria Zeggini was supported by The Wellcome Trust (098051). Heikki A. Koistinen has received funding from Academy of Finland (support for clinical research careers, grant no 258753). Veikko Salomaa is funded by the Finnish Foundation for Cardiovascular Research and the Academy of Finland (grant # 139635). Andrew P Morris is a Wellcome Trust Senior Fellow in Basic Biomedical Science (under award WT098017). Fredrik Karpe is supported by NIHR Oxford Biomedical Research Centre and NIHR National Bioresource. Graeme I. Bell was supported by NIH P30DK020595 (for genotyping, and analysis). James B. Meigs was supported by National Institute for Diabetes and Digestive and Kidney Diseases (NIDDK) R01 DK078616, NIDDK K24 DK080140. Mark I McCarthy is a Wellcome Trust Senior Investigator and a NIHR Senior Investigator. Anna L Gloyn is a Wellcome Trust Senior Fellow in Basic Biomedical Science and is supported by Wellcome Trust (200837/Z/16/Z). Cecilia Lindgren is supported in part by Wellcome Trust (WT086596/Z/08/A and 086596/Z/08/Z) and the Li Ka Shing Foundation.

Study and Cohort Acknowledgements

Funding for the GoT2D and T2D-GENES studies was provided by grants: NIH U01s DK085526, DK085501, DK085524, DK085545, and DK085584 (Multiethnic Study of Type 2 Diabetes Genes) and DK088389 (Low-Pass Sequencing and High-Density SNP Genotyping for Type 2 Diabetes). The work at the University of Oxford, UK was supported by the European Commission (ENGAGE: HEALTH-F4-2007-201413; Marie-Curie Fellowship PIEF-GA-2012-329156), MRC (G0601261, G0900747-91070), National Institutes of Health (RC2-DK088389, DK085545, DK098032), and Wellcome Trust (064890, 083948, 085475, 086596, 090367, 090532, 092447, 095101, 095552, 098017, 098381). The work at the Wellcome Trust Sanger Institute, UK was supported by the National Institute for Health Research and the Wellcome Trust (098051).

The Jackson Heart Study is supported by contracts HHSN268201300046C, HHSN268201300047C, HHSN268201300048C, HHSN268201300049C, HHSN268201300050C from the National Heart, Lung, and Blood Institute and the National Institute on Minority Health and Health Disparities.

The work at Wake Forest School of Medicine (WFSM) was supported by NIH grant R01 DK066358 (DWB).

The Korea Association Research Project was supported at Center for Genome Science, National Institute of Health, Republic of Korea by the Korea National Institute of Health (2012-N73002-00) and the Korea National Institute of Health and Korea Centers for Disease Control and Prevention (4845–301); and at Hallym University Chuncheon, Republic of Korea by theNational Research Foundation of Korea (NRF-2012R1A2A1A03006155). This study was provided with

biospecimens and data from the Korean Genome Analysis Project (4845-301), the Korean Genome and Epidemiology Study (4851-302), and the Korea Biobank Project (4851-307, KBP-2013-11 and KBP-2014-68) that were supported by the Korea Centers for Disease Control and Prevention, Republic of Korea.

The work at the University of Texas Health Science Center at Houston, USA was supported by the National Institutes of Health (U01DK085501, R01HL102830, R01DK073541)

The work at Imperial College London, UK was supported by Action on Hearing Loss (G51), the British Heart Foundation (SP/04/002), European Union FP7 (EpiMigrant, 279143), Medical Research Council (G0601966, G0700931), MRC-PHE Centre for Environment and Health, The National Institute for Health Research (NIHR) (RP-PG-0407-10371), NIHR Biomedical Research Centre at Imperial College Health Care NHS Trust, NIHR Health Protection Research Unit on Health Impact of Environmental Hazards, and the Wellcome Trust (084723). Personal support includes Paul Elliot: NIHR Senior Investigator.

The LOLIPOP study is supported by the National Institute for Health Research (NIHR) Comprehensive Biomedical Research Centre Imperial College Healthcare NHS Trust. The work was carried out in part at the NIHR/Wellcome Trust Imperial Clinical Research Facility. We thank the participants and research staff who made the study possible.

The work at the National University of Singapore was supported by Biomedical Research Council (BMRC) Individual Research Gran, National Medical Research Council (NMRC) Individual Research Grant, NMRC Centre Grant. Personal support includes: Ching-Yu Cheng: NMRC Clinician Scientist award; E Shyong Tai: NMRC Clinician Scientist award; YY Teo: National Research Foundation Fellowship; TY Wong: NMRC Singapore Translational Research Investigator award.

The work at Helmholtz Zentrum München – German Research Center for Environmental Health, Germany was supported by The German Center for Diabetes Research (DZD), Helmholtz Zentrum München (German Research Center for Environmental Health), which is supported by the German Federal Ministry of Education and Research (BMBF) and by the State of Bavaria, and the Munich Center of Health Sciences (MC-Health), Ludwig-Maximilians-Universität, as part of LMUinnovativ.

The KORA research platform (KORA, Cooperative Research in the Region of Augsburg) was initiated and financed by the Helmholtz Zentrum München – German Research Center for Environmental Health, which is funded by the German Federal Ministry of Education and Research and by the State of Bavaria. Furthermore, KORA research was supported within the Munich Center of Health Sciences (MC Health), Ludwig-Maximilians-Universität, as part of LMUinnovativ.

The work at Lund University, Sweden was supported by the Academy of Finland, a European Research Council Advanced Research Grant, the Folkhälsan Research Foundation, Novo Nordisk, the Påhlssons Foundation, the Sigrid Juselius Foundation, the Skåne Regional Health Authority, the Swedish Heart-Lung Foundation, and the Swedish Research Council (Linné and Strategic Research Grant).

TwinsUK was funded by the Wellcome Trust; European Community's Seventh Framework Programme (FP7/2007-2013). The study also receives support from the National Institute for Health Research (NIHR) funded BioResource, Clinical Research Facility and Biomedical Research Centre based at Guy's and St Thomas' NHS Foundation Trust in partnership with King's College London. SNP Genotyping was performed by The Wellcome Trust Sanger Institute and National Eye Institute via NIH/CIDR.

Some computations were performed at the Vital-IT (http://www.vital-it.ch) Center for high-performance computing of the SIB Swiss Institute of Bioinformatics; and at the ACEnet, the regional high performance computing consortium for universities in Atlantic Canada. ACEnet is funded by the Canada Foundation for Innovation (CFI), the Atlantic Canada Opportunities Agency (ACOA), and the provinces of Newfoundland and Labrador, Nova Scotia, and New Brunswick.

FIN-D2D was supported by financing from the hospital districts of Pirkanmaa, Southern Ostrobothnia, North Ostrobothnia, Central Finland, and Northern Savo; the Finnish National Public Health Institute; the Finnish Diabetes Association; the Ministry of Social Affairs and Health in Finland; Finland's Slottery Machine Association; the Academy of Finland (grant No. 129293) and Commission of the European Communities, Directorate C-Public Health (grant agreement No. 2004310) in cooperation with the FIN-D2D Study Group.

The Finnish DPS study was supported by the Academy of Finland (grants 128315, 129330, 131593).

The METSIM study was supported by the Academy of Finland (contract 124243), the Finnish Heart Foundation, the Finnish Diabetes Foundation, Tekes (contract 1510/31/06), and the Commission of the EuropeanCommunity(HEALTH-F2-2007-201681), and the US National Institutes of Health grants DK093757, DK072193, DK062370, and 1Z01 HG000024. Genotyping of the METSIM and DPS studies was conducted at the Genetic Resources Core Facility (GRCF) at the Johns Hopkins Institute of Genetic Medicine.

The DR's EXTRA Study was supported by grants to RR by the Ministry of Education and Culture of Finland (627;2004-2011), Academy of Finland (102318; 123885), Kuopio University Hospital, Finnish Diabetes Association, Finnish Heart Association, Päivikki and Sakari Sohlberg Foundation and by grants from European Commission FP6 Integrated Project (EXGENESIS); LSHM-CT-2004-005272, City of Kuopio and Social Insurance Institution of Finland (4/26/ 2010).

The National FINRISK 2007 study was supported by Finnish Foundation for Cardiovascular Research, the Academy of Finland (grant # 139635).

The FUSION study was supported by DK093757, DK072193, DK062370, and 1Z01 HG000024.

The Inter99 study Data collection in the Inter99 study was supported economically by The Danish Medical Research Council, The Danish Centre for Evaluation and Health Technology Assessment, Novo Nordisk, Copenhagen County, The Danish Heart Foundation, The Danish Pharmaceutical Association, Augustinus foundation, Ib Henriksen foundation and Becket foundation. The Danish studies (Inter99, Health2006, and Vejle Biobank) were supported by the Lundbeck Foundation (Lundbeck Foundation Centre for Applied Medical Genomics in Personalised Disease Prediction, Prevention and Care (LuCamp); http://www. lucamp.org/) and the Danish Council for Independent Research. The Novo Nordisk Foundation Center for Basic Metabolic Research is an independent Research Center at the University of Copenhagen, partially funded by an unrestricted donation from the Novo Nordisk Foundation (http://www.metabol.ku.dk/).

GoDARTS study was funded by The Wellcome Trust Study Cohort Wellcome Trust Functional Genomics Grant (2004-2008) (Grant No: 072960/2/ 03/2) and The Wellcome Trust Scottish Health Informatics Programme (SHIP) (2009-2012) (Grant No: 086113/Z/08/Z). Analysis and genotyping of the British UK cohorts was supported by Wellcome Trust funding 090367, 098381, 090532, 083948, 085475, MRC (G0601261), EU (Framework 7) HEALTH-F4-2007-201413, and NIDDK DK098032.

TwinsUK study was funded by the Wellcome Trust; European Community's Seventh Framework Programme (FP7/2007–2013). The study also receives support from the National Institute for Health Research (NIHR) BioResource Clinical Research Facility and Biomedical Research Centre based at Guy's and St Thomas' NHS Foundation Trust and King's College London.

The Oxford Biobank is supported by the Oxford Biomedical Research Centre and part of the National NIHR Bioresource.

The PIVUS/ULSAM cohort was supported by Wellcome Trust Grants WT098017, WT064890, WT090532, Uppsala University, Uppsala University Hospital, the Swedish Research Council and the Swedish Heart-Lung Foundation.

The Botnia study has been financially supported by grants from the Sigrid Juselius Foundation, Folkhälsan Research Foundation, Nordic Center of Excellence in Disease Genetics, an EU grant (EXGENESIS), Signe and Ane Gyllenberg Foundation, Swedish Cultural Foundation in Finland, Finnish Diabetes Research Foundation, Foundation for Life and Health in Finland, Finnish Medical Society, Paavo Nurmi Foundation, Helsinki University Central Hospital Research Foundation, Perklén Foundation, Ollqvist Foundation, Närpes Health Care Foundation and Ahokas Foundation. The study has also been supported by the Ministry of Education in Finland, Municipal Heath Care Center and Hospital in Jakobstad and Health Care Centers in Vasa, Närpes and Korsholm.

The Cardiovascular Risk in Young Finns Study was financially supported by the Academy of Finland (grants 121584, 126925, 124282, and 129378), the Social Insurance Institution of Finland, the Turku University Foundation, special federal grants for University Hospitals, the Juho Vainio Foundation, Paavo Nurmi Foundation, the Finnish Foundation of Cardiovascular Research, Orion-Farmos Research Foundation, and the Finnish Cultural Foundation.

The Hesinki Birth Cohort Study was supported by Emil Aaltonen Foundation, Finnish Foundation for Diabetes Research, Novo Nordisk Foundation, Signe and Ane Gyllenberg Foundation, Samfundet Folkhälsan, Finska Läkaresällskapet, Liv och Hälsa, Finnish Foundation for Cardiovascular Research.

Additional Acknowledgements

We thank the High-Throughput Genomics Group at the Wellcome Trust Centre for Human Genetics for the generation of array and sequencing data. The High-Throughput Genomics Group at the Wellcome Trust Centre for Human Genetics is funded by a Wellcome Trust grant (reference 090532/Z/09/Z)

The Genotype-Tissue Expression (GTEx) Project was supported by the Common Fund of the Office of the Director of the National Institutes of Health. Additional funds were provided by the NCI, NHGRI, NHLBI, NIDA, NIMH, and NINDS. Donors were enrolled at Biospecimen Source Sites funded by NCI\SAIC-Frederick, Inc. (SAIC-F) subcontracts to the National Disease Research Interchange (10XS170), Roswell Park Cancer Institute (10XS171), and Science Care, Inc. (X10S172). The Laboratory, Data Analysis, and Coordinating Center (LDACC) was funded through a contract (HHSN268201000029C) to The Broad Institute, Inc. Biorepository operations were funded through an SAIC-F subcontract to Van Andel Institute (10ST1035). Additional data repository and project management were provided by SAIC-F (HHSN261200800001E). The Brain Bank was supported by a supplements to University of Miami grants DA006227 & DA033684 and to contract N01MH000028. Statistical Methods development grants were made to the University of Geneva (MH090941 & MH101814), the University of Chicago (MH090951, MH090937, MH101820, MH101825), the University of North Carolina – Chapel Hill (MH090936 & MH101819), Harvard University (MH090948), Stanford University (MH101782), Washington University St Louis (MH101810), and the University of Pennsylvania (MH101822). The data used for the analyses described in this manuscript were obtained from dbGaP (accession number phs000424.v3.p1).

Funding support for "Building on GWAS for NHLBI-diseases: the U.S. CHARGE consortium" was provided by the NIH through the American Recovery and Reinvestment Act of 2009 (ARRA) (5RC2HL102419). Sequence data for "Building on GWAS for NHLBI-diseases: the U.S. CHARGE consortium" was provided by Eric Boerwinkle on behalf of the Atherosclerosis Risk in Communities (ARIC) Study, L. Adrienne Cupples, principal investigator for the Framingham Heart Study, and Bruce Psaty, principal investigator for the Cardiovascular Health Study. Sequencing was carried out at the Baylor Genome Center (U54 HG003273). Further support came from HL120393, "Rare variants and NHLBI traits in deeply phenotyped cohorts" (Bruce Psaty, principal investigator). Supporting funding was also provided by NHLBI with the CHARGE infrastructure grant HL105756.