The Trans-Ancestral Genomic Architecture of Glycemic Traits

Ji Chen[1,2#], Cassandra N. Spracklen[3,4#], Gaëlle Marenne[2,5#], Arushi Varshney[6#], Laura J Corbin[7,8#], Jian'an Luan[9], Sara M Willems[9], Ying Wu[3], Xiaoshuai Zhang[9,10], Momoko Horikoshi[11,12,13], Thibaud S Boutin[14], Reedik Mägi[15], Johannes Waage[16,17], Achilleas Pitsilides[18], Ruifang Li-Gao[19], Kei Hang Katie Chan[20,21,22], Jie Yao[23], Mila D Anasanti[24], Audrey Y Chu[25], Annique Claringbould[26], Jani Heikkinen[24], Jaeyoung Hong[18], Jouke-Jan Hottenga[27,28], Shaofeng Huo[29], Marika A. Kaakinen[30,24], Tin Louie[31], Winfried März[32,33,34], Hortensia Moreno-Macias[35], Anne Ndungu[12], Sarah C. Nelson[31], Ilja M. Nolte[36], Kari E North[37], Chelsea K. Raulerson[3], Debashree Ray[38], Rebecca Rohde[37], Denis Rybin[18], Claudia Schurmann[39,40], Xueling Sim[41,42,43], Loz Southam[2], Isobel D Stewart[9], Carol A. Wang[44], Yujie Wang[37], Peitao Wu[18], Weihua Zhang[45,46], Tarunveer S. Ahluwalia[16,17,47], Emil VR Appel[48], Lawrence F. Bielak[49], Jennifer A. Brody[50], Noël P Burtt[51], Claudia P Cabrera[52,53], Brian E Cade[54,55], Jin Fang Chai[41], Xiaoran Chai[56,57], Li-Ching Chang[58], Chien-Hsiun Chen[58], Brian H Chen[59], Kumaraswamy Naidu Chitrala[60], Yen-Feng Chiu[61], Hugoline G. de Haan[19], Graciela E Delgado[34], Ayse Demirkan[62,30], Qing Duan[3,63], Jorgen Engmann[64], Segun A Fatumo[65,66,67], Javier Gayán[68], Franco Giulianini[69], Jung Ho Gong[20], Stefan Gustafsson[70], Yang Hai[71], Fernando P Hartwig[72,7], Jing He[73], Yoriko Heianza[74], Tao Huang[75], Alicia Huerta-Chagoya[76,77], Mi Yeong Hwang[78], Richard A. Jensen[50], Takahisa Kawaguchi[79], Katherine A Kentistou[80,81], Young Jin Kim[78], Marcus E Kleber[34], Ishminder K Kooner[46], Shuiqing Lai[20], Leslie A Lange[82], Carl D Langefeld[83], Marie Lauzon[23], Man Li[84], Symen Ligthart[62], Jun Liu[62,85], Marie Loh[86,45], Jirong Long[87], Valeriya Lyssenko[88,89], Massimo Mangino[90,91], Carola Marzi[92,93], May E Montasser[94], Abhishek Nag[12], Masahiro Nakatochi[95], Damia Noce[96], Raymond Noordam[97], Giorgio Pistis[98], Michael Preuss[39,99], Laura Raffield[3], Laura J. Rasmussen-Torvik[100], Stephen S Rich[101,102], Neil R Robertson[11,12], Rico Rueedi[103,104], Kathleen Ryan[94], Serena Sanna[98,26], Richa Saxena[105,106,107], Katharina E Schraut[80,81], Bengt Sennblad[108], Kazuya Setoh[79], Albert V Smith[109,110], Lorraine Southam[111,112], Thomas Sparsø[48], Rona J Strawbridge[113,114], Fumihiko Takeuchi[115], Jingyi Tan[23], Stella Trompet[97,116], Erik van den Akker[117,118,119], Peter J van der Most[36], Niek Verweij[120,121], Mandy Vogel[122], Heming Wang[54,55], Chaolong Wang[123,124], Nan Wang[125,126], Helen R Warren[52,53], Wanqing Wen[87], Tom Wilsgaard[127], Andrew Wong[128], Andrew R Wood[1], Tian Xie[36], Mohammad Hadi Zafarmand[129,130], Jing-Hua Zhao[131], Wei Zhao[49], Najaf Amin[62,85], Zorayr Arzumanyan[23], Arne Astrup[132], Stephan JL Bakker[133], Damiano Baldassarre[134,135], Marian Beekman[117], Richard N Bergman[136], Alain Bertoni[137], Matthias Blüher[138], Lori L. Bonnycastle[139], Stefan R Bornstein[140], Donald W Bowden[141], Qiuyin Cai[73], Archie Campbell[142,143], Harry Campbell[80], Yi Cheng Chang[144,145,146], Eco J.C. de Geus[27,28], Abbas Dehghan[62], Shufa Du[147], Gudny Eiriksdottir[110], Aliki Eleni Farmaki[148,149], Mattias Frånberg[150], Christian Fuchsberger[96], Yutang Gao[151], Anette P Gjesing[48], Anuj Goel[152,12], Sohee Han[78], Catharina A Hartman[153], Christian Herder[154,155,156], Andrew A. Hicks[96], Chang-Hsun Hsieh[157,158], Willa A. Hsueh[159], Sahoko Ichihara[160], Michiya Igase[161], M. Arfan Ikram[62], W. Craig Johnson[31], Marit E Jørgensen[17,162], Peter K Joshi[80], Rita R Kalyani[163], Fouad R. Kandeel[164], Tomohiro Katsuya[165,166], Chiea Chuen Khor[124], Wieland Kiess[122], Ivana Kolcic[167], Teemu Kuulasmaa[168], Johanna Kuusisto[169], Kristi Läll[15], Kelvin Lam[23], Deborah A Lawlor[170,8], Nanette R. Lee[171,172], Rozenn N. Lemaitre[50], Honglan Li[173], Lifelines Cohort Study[174], Shih-Yi Lin[175,176,177], Jaana Lindström[178], Allan Linneberg[179,180], Jianjun Liu[124,181], Carlos Lorenzo[182], Tatsuaki Matsubara[183], Fumihiko Matsuda[79], Geltrude Mingrone[184], Simon Mooijaart[97], Sanghoon Moon[78], Toru Nabika[185], Girish N. Nadkarni[39], Jerry L. Nadler[186], Mari Nelis[15], Matt J Neville[11,187], Jill M Norris[188], Yasumasa Ohyagi[189], Annette Peters[190,93,191], Patricia A. Peyser[49], Ozren Polasek[167,192], Qibin Qi[193], Dennis Raven[153], Dermot F Reilly[194], Alex Reiner[195], Fernando Rivideneira[196], Kathryn Roll[23], Igor Rudan[197], Charumathi Sabanayagam[56,198], Kevin Sandow[23], Naveed Sattar[199], Annette Schürmann[200,201], Jinxiu Shi[202], Heather M Stringham[43,42], Kent D. Taylor[23], Tanya M. Teslovich[203], Betina Thuesen[179], Paul RHJ Timmers[80,204], Elena Tremoli[135], Michael Y Tsai[205], Andre Uitterlinden[196], Rob M van Dam[41,181,206], Diana van Heemst[97], Astrid van Hylckama Vlieg[19], Jana V Van Vliet-Ostaptchouk[36], Jagadish Vangipurapu[207], Henrik Vestergaard[48,208], Tao Wang[193], Ko Willems van Dijk[209,210,211], Tatijana Zemunik[212], Goncalo R Abecasis[43], Linda S. Adair[147,213], Carlos Alberto Aguilar-

Salinas[214,215,216], Marta E Alarcón-Riquelme[217,218], Ping An[219], Larissa Aviles-Santa[220], Diane M Becker[221], Lawrence J Beilin[222], Sven Bergmann[103,104,223], Hans Bisgaard[16], Corri Black[224], Michael Boehnke[43,42], Eric Boerwinkle[225,226], Bernhard O Böhm[227,228], Klaus Bønnelykke[16], D I. Boomsma[27,28], Erwin P. Bottinger[39,229,230], Thomas A Buchanan[231,232,126], Mickaël Canouil[233,234], Mark J Caulfield[52,53], John C. Chambers[86,45,46,235,236], Daniel I. Chasman[69,237], Yii-Der Ida Chen[23], Ching-Yu Cheng[56,198], Francis S. Collins[139], Adolfo Correa[238], Francesco Cucca[98], H. Janaka de Silva[239], George Dedoussis[240], Sölve Elmståhl[241], Michele K. Evans[242], Ele Ferrannini[243], Luigi Ferrucci[244], Jose C Florez[245,246,107], Paul W Franks[89,247], Timothy M Frayling[1], Philippe Froguel[233,234,248], Bruna Gigante[249], Mark O. Goodarzi[250], Penny Gordon-Larsen[147,213], Harald Grallert[92,93], Niels Grarup[48], Sameline Grimsgaard[127], Leif Groop[251,252], Vilmundur Gudnason[110,253], Xiuqing Guo[23], Anders Hamsten[114], Torben Hansen[48], Caroline Hayward[204], Susan R. Heckbert[254], Bernardo L Horta[72], Wei Huang[202], Erik Ingelsson[255], Pankow S James[256], Marjo-Ritta Jarvelin[257,258,259,260], Jost B Jonas[261,262,263], J. Wouter Jukema[116,264], Pontiano Kaleebu[265], Robert Kaplan[193,195], Sharon L.R. Kardia[49], Norihiro Kato[115], Sirkka M. Keinanen-Kiukaanniemi[266,267], Bong-Jo Kim[78], Mika Kivimaki[268], Heikki A. Koistinen[269,270,271], Jaspal S. Kooner[46,235,236,272], Antje Körner[122], Peter Kovacs[138,273], Diana Kuh[128], Meena Kumari[274], Zoltan Kutalik[275,104], Markku Laakso[169], Timo A. Lakka[276,277,278], Lenore J Launer[60], Karin Leander[279], Huaixing Li[29], Xu Lin[29], Lars Lind[280], Cecilia Lindgren[12,281,282], Simin Liu[20], Ruth J.F. Loos[39,99], Patrik KE Magnusson[283], Anubha Mahajan[12], Andres Metspalu[15], Dennis O Mook-Kanamori[19,284], Trevor A Mori[222], Patricia B Munroe[52,53], Inger Njølstad[127], Jeffrey R O'Connell[94], Albertine J Oldehinkel[153], Ken K Ong[9], Sandosh Padmanabhan[285], Colin N.A. Palmer[286], Nicholette D Palmer[141], Oluf Pedersen[48], Craig E Pennell[44], David J Porteous[142,287], Peter P. Pramstaller[96], Michael A. Province[219], Bruce M. Psaty[50,254,288], Lu Qi[289], Leslie J. Raffel[290], Rainer Rauramaa[278], Susan Redline[54,55], Paul M Ridker[69,291], Frits R. Rosendaal[19], Timo E. Saaristo[292,293], Manjinder Sandhu[294], Jouko Saramies[295], Neil Schneiderman[296], Peter Schwarz[140,297,201], Laura J. Scott[43,42], Elizabeth Selvin[38], Peter Sever[272], Xiao-ou Shu[87], P Eline Slagboom[117], Kerrin S Small[90], Blair H Smith[298], Harold Snieder[36], Tamar Sofer[299,246], Thorkild I.A. Sørensen[48,300,7,8], Tim D Spector[90], Alice Stanton[301], Claire J Steves[90,302], Michael Stumvoll[138], Liang Sun[29], Yasuharu Tabara[79], E Shyong Tai[181,41,303], Nicholas J Timpson[7,8], Anke Tönjes[138], Jaakko Tuomilehto[304,305,306], Teresa Tusie[77,307], Matti Uusitupa[308], Pim van der Harst[120,26], Cornelia van Duijn[85,62], Veronique Vitart[204], Peter Vollenweider[309], Tanja GM Vrijkotte[129], Lynne E Wagenknecht[310], Mark Walker[311], Ya X Wang[262], Nick J Wareham[9], Richard M Watanabe[125,232,126], Hugh Watkins[152,12], Wen B Wei[312], Ananda R Wickremasinghe[313], Gonneke Willemsen[27,28], James F Wilson[80,204], Tien-Yin Wong[56,198], Jer-Yuarn Wu[58], Anny H Xiang[314], Lisa R Yanek[221], Loïc Yengo[315], Mitsuhiro Yokota[316], Eleftheria Zeggini[111,317,318], Wei Zheng[87], Alan B Zonderman[60], Jerome I Rotter[23], Anna L Gloyn[11,12,187,319], Mark I. McCarthy[11,320,187,12@], Josée Dupuis[18], James B Meigs[321,246,107], Robert A Scott[9], Inga Prokopenko[30,24], Aaron Leong[322,323,237], Ching-Ti Liu[18], Stephen CJ Parker[6,324#], Karen L. Mohlke[3#], Claudia Langenberg[9#], Eleanor Wheeler[2,9#], Andrew P. Morris[325,326,327,12#], Inês Barroso[1,2,9#] on behalf of the Meta-Analysis of Glucose and Insulin-related Traits Consortium (MAGIC)

[1]Exeter Centre of Excellence for Diabetes Research (EXCEED), Genetics of Complex Traits, University of Exeter Medical School, University of Exeter, Exeter, UK, [2]Department of Human Genetics, Wellcome Sanger Institute, Hinxton, Cambridge, UK, [3]Department of Genetics, University of North Carolina, Chapel Hill, NC, USA, [4]Department of Biostatistics and Epidemiology, University of Massachusetts, Amherst, MA, USA, [5]Inserm, Univ Brest, EFS, UMR 1078, GGB, Brest, France, [6]Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI, USA, [7]MRC Integrative Epidemiology Unit, University of Bristol, Bristol, UK, [8]Department of Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, Bristol, UK, [9]MRC Epidemiology Unit, Institute of Metabolic Science, University of Cambridge, Cambridge, UK, [10]Department of Biostatistics, School of Public Health, Shandong University, Jinan, Shandong, China, [11]Oxford Centre for Diabetes, Endocrinology and Metabolism, Radcliffe Department of Medicine, University of Oxford, Oxford, UK, [12]Wellcome Centre for Human Genetics, University of Oxford, Oxford, UK, [13]Laboratory for Genomics of Diabetes and Metabolism, RIKEN Centre for Integrative

103 Medical Sciences, Yokohama, Japan, [14]Medical Research Council Human Genetics Unit, Institute for
104 Genetics and Molecular Medicine, Edinburgh, UK, [15]Estonian Genome Center, Institute of Genomics,
105 University of Tartu, Tartu, Estonia, [16]COPSAC, Copenhagen Prospective Studies on Asthma in
106 Childhood, Herlev and Gentofte Hospital, University of Copenhagen, Copenhagen, Denmark, [17]Steno
107 Diabetes Center Copenhagen, Gentofte, Denmark, [18]Department of Biostatistics, Boston University
108 School of Public Health, Boston, MA, USA, [19]Department of Clinical Epidemiology, Leiden University
109 Medical Center, Leiden, The Netherlands, [20]Department of Epidemiology, Brown University School of
110 Public Health, Brown University, Providence, RI, USA, [21]Department of Biomedical Sciences, City
111 University of Hong Kong, Hong Kong SAR, China, [22]Department of Electrical Engineering, City
112 University of Hong Kong, Hong Kong SAR, China, [23]The Institute for Translational Genomics and
113 Population Sciences, Department of Pediatrics, The Lundquist Institute for Biomedical Innovation at
114 Harbor-UCLA Medical Center, Torrance, CA, USA, [24]Department of Metabolism, Digestion, and
115 Reproduction, Imperial College London, London, UK, [25]Division of Preventive Medicine, Brigham and
116 Women's Hospital, Boston, MA, USA, [26]Department of Genetics, University of Groningen, University
117 Medical Center Groningen, Groningen, The Netherlands, [27]Department of Biological Psychology,
118 Faculty of Behaviour and Movement Sciences, Vrije Universiteit Amsterdam, Amsterdam, The
119 Netherlands, [28]Amsterdam Public Health Research Institute, Amsterdam Universities Medical Center,
120 Amsterdam, The Netherlands, [29]CAS Key Laboratory of Nutrition, Metabolism and Food Safety,
121 Shanghai Institute of Nutrition and Health, University of Chinese Academy of Sciences, Chinese
122 Academy of Sciences, Shanghai, China, [30]Section of Statistical Multi-omics, Department of Clinical
123 and Experimental Research, University of Surrey, Guildford, Surrey, UK, [31]Department of
124 Biostatistics, University of Washington, Seattle, WA, USA, [32]SYNLAB Academy, SYNLAB Holding
125 Deutschland GmbH, Mannheim, Germany, [33]Clinical Institute of Medical and Chemical Laboratory
126 Diagnostics, Medical University Graz, Graz, Austria, [34]Vth Department of Medicine (Nephrology,
127 Hypertensiology, Rheumatology, Endocrinology, Diabetology), Medical Faculty Mannheim,
128 Heidelberg University, Mannheim, Baden-Württemberg, Germany, [35]Department of Economics,
129 Metropolitan Autonomous University, Mexico City, Mexico, [36]Department of Epidemiology,
130 University of Groningen, University Medical Center Groningen, Groningen, The Netherlands, [37]CVD
131 Genetic Epidemiology Computational Laboratory, Gillings School of Global Public Health, University
132 of North Carolina, Chapel Hill, NC, USA, [38]Department of Epidemiology, Johns Hopkins Bloomberg
133 School of Public Health, Baltimore, MD, USA, [39]The Charles Bronfman Institute for Personalized
134 Medicine, Icahn School of Medicine at Mount Sinai, New York, NY, USA, [40]HPI Digital Health Center,
135 Digital Health and Personalized Medicine, Hasso Plattner Institute, Potsdam, Germany, [41]Saw Swee
136 Hock School of Public Health, National Univeristy of Singapore and National University Health
137 System, Singapore, Singapore, [42]Center for Statistical Genetics, University of Michigan, Ann Arbor,
138 MI, USA, [43]Department of Biostatistics, School of Public Health, University of Michigan, Ann Arbor,
139 MI, USA, [44]School of Medicine and Public Health, College of Health, Medicine and Wellbeing, The
140 University of Newcastle, Newcastle, NSW, Australia, [45]Department of Epidemiology and Biostatistics,
141 Imperial College London, London, UK, [46]Department of Cardiology, Ealing Hospital, London North
142 West Healthcare NHS Trust, Middlesex, UK, [47]The Bioinformatics Centre, Department of Biology,
143 University of Copenhagen, Copenhagen, Denmark, [48]Novo Nordisk Foundation Center for Basic
144 Metabolic Research, Faculty of Health and Medical Sciences, University of Copenhagen,
145 Copenhagen, Denmark, [49]Department of Epidemiology, School of Public Health, University of
146 Michigan, Ann Arbor, MI, USA, [50]Department of Medicine, Cardiovascular Health Research Unit,
147 University of Washington, Seattle, WA, USA, [51]Metabolism Program, Program in Medical and
148 Population Genetics, Broad Institute, Cambridge, MA, USA, [52]Department of Clinical Pharmacology,
149 William Harvey Research Institute, Barts and The London School of Medicine and Dentistry, Queen
150 Mary University of London, London, UK, [53]NIHR Barts Cardiovascular Biomedical Research Centre,
151 Queen Mary University of London, London, UK, [54]Department of Medicine, Sleep and Circadian
152 Disorders, Brigham and Women's Hospital, Boston, MA, USA, [55]Department of Medicine, Sleep
153 Medicine, Harvard Medical School, Boston, MA, USA, [56]Ocular Epidemiology, Singapore Eye Research

154 Institute, Singapore National Eye Centre, Singapore, Singapore, [57]Department of Ophthalmology,
155 National University of Singapore and National University Health System, Singapore, Singapore,
156 [58]Institute of Biomedical Sciences, Academia Sinica, Taipei, Taiwan, Taiwan, [59]Department of
157 Epidemiology, The Herbert Wertheim School of Public Health and Human Longevity Science, UC San
158 Diego, La Jolla, CA, USA, [60]Laboratory of Epidemiology and Population Sciences, National Institute on
159 Aging, National Institutes of Health, Baltimore, MD, USA, [61]Institute of Population Health Sciences,
160 National Health Research Institutes, Miaoli, Taiwan, [62]Department of Epidemiology, Erasmus Medical
161 Center, Rotterdam, The Netherlands, [63]Department of Statistics, University of North Carolina at
162 Chapel Hill, Chapel Hill, NC, USA, [64]Institute of Cardiovascular Science, UCL, London, UK, [65]Uganda
163 Medical Informatics Centre (UMIC), MRC/UVRI and London School of Hygiene & Tropical Medicine
164 (Uganda Research Unit), Entebbe, Uganda, [66]London School of Hygiene & Tropical Medicine, London,
165 UK, [67]H3Africa Bioinformatics Network (H3ABioNet) Node, Centre for Genomics Research and
166 Innovation, NABDA/FMST, Abuja, Nigeria, [68]Bioinfosol, Sevilla, Spain, [69]Division of Preventive
167 Medicine, Brigham and Women's Hospital, Boston, MA, USA, [70]Department of Medical Sciences,
168 Molecular Epidemiology and Science for Life Laboratory, Uppsala University, Uppsala, Sweden,
169 [71]Department of Statistics, The University of Auckland, Science Center, Auckland, New Zealand,
170 [72]Postgraduate Program in Epidemiology, Federal University of Pelotas, Pelotas, RS, Brazil,
171 [73]Department of Medicine, Epidemiology, Vanderbilt University Medical Center, Nashville, TN, USA,
172 [74]Department of Epidemiology, Tulane University Obesity Research Center,, Tulane University, New
173 Orleans, USA, [75]Department of Epidemiology and Biostatistics, School of Public Health, Peking
174 University, Beijing, China, [76]Molecular Biology and Genomic Medicine Unit, National Council for
175 Science and Technology, Mexico City, Mexico, [77]Molecular Biology and Genomic Medicine Unit,
176 National Institute of Medical Sciences and Nutrition, Mexico City, Mexico, [78]Division of Genome
177 Science, Department of Precision Medicine, National Institute of Health, Cheongju-si,
178 Chungcheongbuk-do, South Korea, [79]Center for Genomic Medicine, Kyoto University Graduate
179 School of Medicine, Kyoto, Japan, [80]Centre for Global Health Research, Usher Institute, University of
180 Edinburgh, Edinburgh, Scotland, [81]Centre for Cardiovascular Sciences, Queen's Medical Research
181 Institute, University of Edinburgh, Edinburgh, Scotland, [82]Department of Medicine, Divison of
182 Biomedical Informatics and Personalized Medicine, University of Colorado Anschutz Medical
183 Campus, Denver, CO, USA, [83]Department of Biostatistics and Data Science, Wake Forest School of
184 Medicine, Winston-Salem, NC, USA, [84]Department of Medicine, Division of Nephrology and
185 Hypertension, University of Utah, Salt Lake City, UT, USA, [85]Nuffield Department of Population
186 Health, University of Oxford, Oxford, UK, [86]Lee Kong Chian School of Medicine, Nanyang
187 Technological University, Singapore, Singapore, [87]Division of Epidemiology, Department of Medicine,
188 Vanderbilt Epidemiology Center, Vanderbuilt University Medical Center, Nashville, TN, USA,
189 [88]Department of Clinical Science, Center for Diabetes Research, University of Bergen, Bergen,
190 Norway, [89]Department of Clinical Sciences, Lund University Diabetes Centre, Lund University,
191 Malmo, Sweden, [90]Department of Twin Research and Genetic Epidemiology, School of Life Course
192 Sciences, King's College London, London, UK, [91]NIHR Biomedical Research Centre, Guy’s and St
193 Thomas’ Foundation Trust, London, UK, [92]Institute of Epidemiology, Research Unit of Molecular
194 Epidemiology, Helmholtz Zentrum München Research Center for Environmental Health, Neuherberg,
195 Bavaria, Germany, [93]German Center for Diabetes Research (DZD), Neuherberg, Bavaria, Germany,
196 [94]Department of Medicine, Division of Endocrinology, Diabetes, and Nutrition, University of
197 Maryland School of Medicine, Baltimore, MD, USA, [95]Public Health Informatics Unit, Department of
198 Integrated Sciences, Nagoya University Graduate School of Medicine, Nagoya, Japan, [96]Institute for
199 Biomedicine, Eurac Research, Bolzano, BZ, Italy, [97]Department of Internal Medicine, Section of
200 Gerontology and Geriatrics, Leiden University Medical Center, Leiden, The Netherlands, [98]Istituto di
201 Ricerca Genetica e Biomedica (IRGB), Consiglio Nazionale delle Ricerche (CNR), Monserrato, Italy,
202 [99]The Mindich Child Health and Development Institute for Personalized Medicine, Icahn School of
203 Medicine at Mount Sinai, New York, NY, USA, [100]Department of Preventive Medicine, Northwestern
204 University Feinberg School of Medicine, Chicago, IL, USA, [101]Center for Public Health Genomics,

205 University of Virginia, Charlottesville, VA, USA, [102]Department of Public Health Sciences, University of
206 Virginia, Charlottesville, VA, USA, [103]Department of Computational Biology, University of Lausanne,
207 Lausanne, Switzerland, [104]Swiss Institute of Bioinformatics, Lausanne, Switzerland, [105]Center for
208 Genomic Medicine, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA,
209 [106]Department of Anesthesia, Critical Care and Pain Medicine, Massachusetts General Hospital,
210 Boston, MA, USA, [107]Program in Medical and Population Genetics,, Broad Institute, Cambridge, MA,
211 USA, [108]Department of Cell and Molecular Biology., National Bioinformatics Infrastructure Sweden,,
212 Science for Life Laboratory, Uppsala University, Uppsala, Sweden, [109]Department of Biostatistics,
213 University of Michigan, Ann Arbor, MI, USA, [110]Icelandic Heart Association, Kopavogur, Iceland,
214 [111]Institute of Translational Genomics, Helmholtz Zentrum München – German Research Center for
215 Environmental Health, Neuherberg, Germany, [112]Wellcome Sanger Institute, Hinxton, Cambridge,
216 UK, [113]Institute of Health and Wellbeing, University of Glasgow, Glasgow, Glasgow, UK,
217 [114]Department of Medicine Solna, Cardiovascular medicine, Karolinska Institutet, Stockholm,
218 Sweden, [115]National Center for Global Health and Medicine, Tokyo, Japan, [116]Department of
219 Cardiology, Leiden University Medical Center, Leiden, The Netherlands, [117]Department of Biomedical
220 Data Sciences, Molecular Epidemiology, Leiden University Medical Center, Leiden, The Netherlands,
221 [118]Department of Pattern Recognition & Bioinformatics, Delft University of Technology, Delft, The
222 Netherlands, [119]Department of Biomedical Data Sciences, Leiden Computational Biology Center,
223 Leiden University Medical Center, Leiden, The Netherlands, [120]Department of Cardiology, University
224 of Groningen, University Medical Center Groningen, Groningen, The Netherlands, [121]Genomics plc,
225 Oxford, UK, [122]Center of Pediatric Research, University Children´s Hospital Leipzig, University of
226 Leipzig Medical Center, Leipzig, Germany, [123]Department of Epidemiology and Biostatistics, School of
227 Public Health, Tongji Medical College, Huazhong University of Science and Technology, Wuhan,
228 China, [124]Genome Institute of Singapore, Agency for Science, Technology and Research, Singapore,
229 Singapore, [125]Department of Preventive Medicine, Keck School of Medicine of USC, Los Angeles, CA,
230 USA, [126]USC Diabetes and Obesity Research Institute, Keck School of Medicine of USC, Los Angeles,
231 CA, USA, [127]Department of Community Medicine, Faculty of Health Sciences, UIT the Arctic
232 University of Norway, Tromsø, Norway, [128]MRC Unit for Lifelong Health & Ageing at UCL, London,
233 UK, [129]Department of Public Health, Amsterdam Public Health Research Institute, Amsterdam
234 Universities Medical Center, Amsterdam, The Netherlands, [130]Department of Clinical Epidemiology,
235 Biostatistics, and Bioinformatics, Amsterdam Public Health Research Institute, Amsterdam
236 Universities Medical Center, Amsterdam, The Netherlands, [131]Department of Public Health and
237 Primary Care, School of Clinical Medicine, University of Cambridge, Cambridge, UK, [132]Department of
238 Nutrition, Exercise, and Sports, Faculty of Science, University of Copenhagen, Copenhagen,
239 Denmark, [133]Department of Internal Medicine, University of Groningen, University Medical Center
240 Groningen, Groningen, The Netherlands, [134]Department of Medical Biotechnology and Translational
241 Medicine, University of Milan, Milan, Italy, [135]Centro Cardiologico Monzino, IRCCS, Milan, Italy,
242 [136]Diabetes and Obesity Research Institute, Cedars-Sinai Medical Center, Los Angeles, CA, USA,
243 [137]Department of Epidemiology and Prevention, Division of Public Health Sciences, Wake Forest
244 School of Medicine, Winston-Salem, NC, USA, [138]Medical Department III – Endocrinology,
245 Nephrology, Rheumatology, University of Leipzig Medical Center, Leipzig, Germany, [139]Medical
246 Genomics and Metabolic Genetics Branch, National Human Genome Research Institute, National
247 Institues of Health, Bethesda, MD, USA, [140]Department for Prevention and Care of Diabetes, Faculty
248 of Medicine Carl Gustav Carus, Technische Universität Dresden, Dresden, Germany, [141]Department
249 of Biochemistry, Wake Forest School of Medicine, Winston-Salem, NC, USA, [142]Centre for Genomic
250 and Experimental Medicine, Institute of Genetics & Molecular Medicine, University of Edinburgh,
251 Western General Hospital, Edinburgh, UK, [143]Usher Institute, University of Edinburgh, Edinburgh, UK,
252 [144]Department of Internal Medicine, National Taiwan University Hospital, Taipei, Taiwan, [145]Graduate
253 Institute of Medical Genomics and Proteomics, National Taiwan University, Taipei, Taiwan,
254 [146]Institute of Biomedical Sciences, Academia Sinica, Taipei, Taiwan, [147]Department of Nutrition,
255 Gillings School of Global Public Health, University of North Carolina, Chapel Hill, NC, USA,

256 [148]Department of Population Science and Experimental Medicine, Institute of Cardiovascular Science,
257 University College London, London, UK, [149]Department of Nutrition and Dietetics, School of Health
258 Science and Education, Harokopio University of Athens, Athens, Greece, [150]Department of Medicine
259 Solna, Cardiovascular medicine, Stockholm, Sweden, [151]Department of Epidemiology, Shanghai
260 Cancer Institute, Shanghai, China, [152]Division of Cardiovascular Medicine, Radcliffe Department of
261 Medicine, University of Oxford, Oxford, UK, [153]Department of Psychiatry, Interdisciplinary Center
262 Psychopathy and Emotion Regulation, University of Groningen, University Medical Center
263 Groningen, Groningen, The Netherlands, [154]Institute for Clinical Diabetology, German Diabetes
264 Center, Leibniz Center for Diabetes Research at Heinrich Heine University Düsseldorf, Düsseldorf,
265 Germany, [155]Division of Endocrinology and Diabetology, Medical Faculty, Heinrich Heine University
266 Düsseldorf, Düsseldorf, Germany, [156]German Center for Diabetes Research (DZD), Düsseldorf,
267 Germany, [157]Internal Medicine, Endocrine & Metabolism, Tri-Service General Hospital, Taipei,
268 Taiwan, [158]School of Medicine, National Defense Medical Center, Taipei, Taiwan, [159]Internal
269 Medicine, Endocrinology, Diabetes & Metabolism, Diabetes and Metabolism Research Center, The
270 Ohio State University Wexner Medical Center, Columbus, OH, USA, [160]Department of Environmental
271 and Preventive Medicine, Jichi Medical University School of Medicine, Shimotsuke, Japan,
272 [161]Department of Anti-aging Medicine, Ehime University Graduate School of Medicine, Toon, Japan,
273 [162]National Institute of Public Health, University of Southern Denmark, Odense, Denmark,
274 [163]Department of Medicine, Endocrinology, Diabetes & Metabolism, Johns Hopkins University School
275 of Medicine, Baltimore, MD, USA, [164]Clinical Diabetes, Endocrinology & Metabolism, Translational
276 Research & Cellular Therapeutics, Beckman Research Institute of the City of Hope, Duarte, CA, USA,
277 [165]Department of Clinical Gene Therapy, Osaka University Graduate School of Medicine, Suita, Japan,
278 [166]Department of Geriatric and General Medicine, Osaka University Graduate School of Medicine,
279 Suita, Japan, [167]Department of Public Health, University of Split School of Medicine, Split, Croatia,
280 [168]Institute of Biomedicine, Bioinformatics Center, Univeristy of Eastern Finland, Kuopio, Finland,
281 [169]Department of Medicine, University of Eastern Finland and Kuopio University Hospital, Kuopio,
282 Finland, [170]MRC Integrative Epidemiology Unit, University of Bristol, Bristol, Bristol, UK, [171]USC-Office
283 of Population Studies Foundation, University of San Carlos, Cebu City, Philippines, [172]Department of
284 Anthropology, Sociology and History, University of San Carlos, Cebu City, Philippines, [173]State Key
285 Laboratory of Oncogene and Related Genes & Department of Epidemiology, Shanghai Cancer
286 Institute, Renji Hospital, Shanghai Jiaotong University School of Medicine, Shanghai, China,
287 [174]Groningen, The Netherlands, [175]Internal Medicine, Endocrine & Metabolism, Taichung Veterans
288 General Hospital, Taichung, Taiwan, [176]Center for Geriatrics and Gerontology,, Taichung Veterans
289 General Hospital, Taichung, Taiwan, [177]National Defense Medical Center, National Yang-Ming
290 University, Taipei, Taiwan, [178]Diabetes Prevention Unit, National Institute for Health and Welfare,
291 Helsinki, Finland, [179]Center for Clinical Research and Prevention, Bispebjerg and Frederiksberg
292 Hospital, Copenhagen, Denmark, [180]Department of Clinical Medicine, Faculty of Health and Medical
293 Sciences, University of Copenhagen, Copenhagen, Denmark, [181]Yong Loo Lin School of Medicine,
294 National University of Singapore and National University Health System, Singapore, Singapore,
295 [182]Department of Medicine, University of Texas Health Sciences Center, San Antonio, TX, USA,
296 [183]Department of Internal Medicine, Aichi Gakuin University School of Dentistry, Nagoya, Japan,
297 [184]Department of Diabetes, Diabetes, & Nutritional Sciences, James Black Centre, King's College
298 London, London, UK, [185]Department of Functional Pathology, Shimane University School of
299 Medicine, Izumo, Japan, [186]Department of Medicine and Pharmacology, New York Medical College
300 School of Medicine, Valhalla, NY, USA, [187]Oxford NIHR Biomedical Research Centre, Oxford University
301 Hospitals NHS Foundation Trust, Oxford, UK, [188]Colorado School of Public Health, University of
302 Colorado Anschutz Medical Campus, Aurora, CO, USA, [189]Department of Geriatric Medicine and
303 Neurology, Ehime University Graduate School of Medicine, Toon, Japan, [190]Institute of Epidemiology,
304 Helmholtz Zentrum München Research Center for Environmental Health, Neuherberg, Bavaria,
305 Germany, [191]Institute for Medical Information Processing, Biometry, and Epidemiology, Ludwig-
306 Maximilians University Munich, Munich, Bavaria, Germany, [192]Gen-info LtD, Zagreb, Croatia,

307 [193]Department of Epidemiology and Population Health, Albert Einstein College of Medicine, Bronx,
308 NY, USA, [194]Genetics and Pharmacogenomics, Merck Sharp & Dohme Corp., Kenilworth, NJ, USA,
309 [195]Department of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA, USA,
310 [196]Department of Internal Medicine, Erasmus Medical Center, Rotterdam, The Netherlands, [197]Centre
311 for Global Health, The Usher Institute, University of Edinburgh, Edinburgh, UK, [198]Ophthalmology &
312 Visual Sciences Academic Clinical Program (Eye ACP), Duke-NUS Medical School, Singapore,
313 Singapore, [199]BHF Glasgow Cardiovascular Research Centre, Institute of Cardiovascular and Medical
314 Sciences, University of Glasgow, Glasgow, UK, [200]Department of Experimental Diabetology, German
315 Institute of Human Nutrition Potsdam-Rehbruecke, Nuthetal, Germany, [201]German Center for
316 Diabetes Research (DZD e.V.), Neuherberg, Germany, [202]Department of Genetics, Shanghai-MOST
317 Key Laboratory of Health and Disease Genomics, Chinese National Human Genome Center at
318 Shanghai (CHGC) and Shanghai Academy of Science & Technology (SAST), Shanghai, China, [203]Sarepta
319 Therapeutics, Cambridge, Massachusetts, USA, [204]Medical Research Council Human Genetics Unit,
320 Institute for Genetics and Cancer, University of Edinburgh, Edinburgh, UK, [205]Department of
321 Laboratory Medicine and Pathology, University of Minnesota, Minneapolis, MN, USA, [206]Department
322 of Nutrition, Harvard T.H. Chan School of Public Health, Boston, MA, USA, [207]Institute of Clinical
323 Medicine, Internal Medicine, University of Eastern Finland, Kuopio, Finland, [208]Department of
324 Medicine, Bornholms Hospital, Rønne, Denmark, [209]Department of Internal Medicine, Division of
325 Endocrinology, Leiden University Medical Center, Leiden, The Netherlands, [210]Laboratory for
326 Experimental Vascular Medicine, Leiden University Medical Center, Leiden, The Netherlands,
327 [211]Department of Human Genetics, Leiden University Medical Center, Leiden, The Netherlands,
328 [212]Department of Human Biology, University of Split School of Medicine, Split, Croatia, [213]Carolina
329 Population Center, University of North Carolina, Chapel Hill, NC, USA, [214]Department of
330 Endocrinology and Metabolism, Instituto Nacional de Ciencias Medicas y Nutricion, Mexico City,
331 Mexico, [215]Unidad de Investigación de Enfermedades Metabólicas, Instituto Nacional de Ciencias
332 Médicas y Nutrición and Tec Salud, Mexico City, Mexico, [216]Instituto Tecnológico y de Estudios
333 Superiores de Monterrey Tec Salud, Mexico City, Mexico, [217]Department of Medical Genomics,
334 Pfizer/University of Granada/Andalusian Government Center for Genomics and Oncological
335 Research (GENYO), Granada, Spain, [218]Institute for Environmental Medicine, Chronic Inflammatory
336 Diseases, Karolinska Institutet, Solna, Sweden, [219]Department of Genetics, Division of Statistical
337 Genomics, Washington University School of Medicine, St. Louis, MO, USA, [220]Clinical and Health
338 Services Research, National Institute on Minority Health and Health Disparities, Bethesda, MD, USA,
339 [221]Department of Medicine, General Internal Medicine, Johns Hopkins University School of Medicine,
340 Baltimore, MD, USA, [222]Medical School, Royal Perth Hospital Unit, University of Western Australia,
341 Perth, WA, Australia, [223]Department of Integrative Biomedical Sciences, University of Cape Town,
342 Cape Town, South Africa, [224]Aberdeen Centre for Health Data Science, 1:042 Polwarth Building,,
343 School of Medicine, Medical, Science and Nutrition, University of Aberdeen, Foresterhill, Aberdeen,
344 UK, [225]Human Genetics Center, School of Public Health, The University of Texas Health Science
345 Center at Houston, Houston, TX, USA, [226]Human Genome Sequencing Center, Baylor College of
346 Medicine, Houston, TX, USA, [227]Division of Endocrinology and Diabetes, Graduate School of
347 Molecular Endocrinology and Diabetes, University of Ulm, Ulm, Baden-Württemberg, Germany,
348 [228]LKC School of Medicine, Nanyang Technological University, Singapore and Imperial College
349 London, UK, Singapore, Singapore, [229]Hasso Plattner Institute for Digital Health at Mount Sinai, Icahn
350 School of Medicine at Mount Sinai, New York, NY, USA, [230]Digital Health Center, Hasso Plattner
351 Institut, University Potsdam, Potsdam, Germany, [231]Department of Medicine, Keck School of
352 Medicine of USC, Los Angeles, CA, USA, [232]Department of Physiology and Neuroscience, Keck School
353 of Medicine of USC, Los Angeles, CA, USA, [233]INSERM UMR 1283 / CNRS UMR 8199, European
354 Institute for Diabetes (EGID), Université de Lille, Lille, France, [234]INSERM UMR 1283 / CNRS UMR
355 8199, European Institute for Diabetes (EGID), Institut Pasteur de Lille, Lille, France, [235]Imperial
356 College Healthcare NHS Trust, Imperial College London, London, UK, [236]MRC-PHE Centre for
357 Environment and Health, Imperial College London, London, UK, [237]Harvard Medical School, Boston,

358 MA, USA, [238]Department of Medicine, Jackson Heart Study, University of Mississippi Medical Center,
359 Jackson, MS, USA, [239]Department of Medicine, Faculty of Medicine, University of Kelaniya, Ragama,
360 Sri Lanka, [240]Department of Nutrition and Dietetics, School of Health Science and Education,
361 Harokopio University of Athens, Kallithea, Greece, [241]Department of Clinical Sciences, Lund
362 University, Malmö, Sweden, [242]Laboratory of Epidemiology and Population Sciences, National
363 Institute on Aging Intramural Research Program, National Institutes of Health, Baltimore, MD, USA,
364 [243]CNR Institute of Clinical Physiology, Pisa, Italy, [244]Intramural Research Program, National Institute
365 of Aging, Baltimore, MD, USA, [245]Diabetes Unit and Center for Genomic Medicine, Massachusetts
366 General Hospital, Boston, MA, USA, [246]Department of Medicine, Harvard Medical School, Boston,
367 MA, USA, [247]Department of Public Health and Clinical Medicine, Umeå University, Umeå, Sweeden,
368 [248]Department of Genomics of Common Disease, Imperial College London, London, UK,
369 [249]Department of Medicine, Cardiovascular medicine, Karolinska Institutet, Stockholm, Sweden,
370 [250]Department of Medicine, Division of Endocrinology, Diabetes & Metabolism, Cedars-Sinai Medical
371 Center, Los Angeles, CA, USA, [251]Diabetes Centre, Lund University, Sweden, [252]Finnish Instituter of
372 Molecular Medicine, Helsinki University, Helsinki, Finland, [253]Faculty of Medicine, School of health
373 sciences, University of Iceland, Reykjavik, Iceland, [254]Department of Epidemiology, Cardiovascular
374 Health Research Unit, University of Washington, Seattle, WA, USA, [255]Department of Medicine,
375 Division of Cardiovascular Medicine, Stanford University School of Medicine, Stanford University,
376 Stanford, CA, USA, [256]Division of Epidemiology and Community Health, University of Minnesota,
377 Minneapolis, MN, USA, [257]Department of Epidemiology and Biostatistics, MRC-PHE Centre for
378 Environment and Health, School of Public Health, Imperial College London, London, UK, [258]Center for
379 Life Course Health Research, Faculty of Medicine, University of Oulu, Oulu, Finland, [259]Unit of
380 Primary Health Care, Oulu Univerisity Hospital, OYS, Oulu, Finland, [260]Department of Life Sciences,
381 College of Health and Life Sciences, Brunel University London, London, UK, [261]Department of
382 Ophthalmology, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany, [262]Beijing
383 Institute of Ophthalmology, Beijing Ophthalmology and Visual Science Key Lab, Beijing Tongren Eye
384 Center, Beijing Tongren Hospital, Capital Medical University, Beijing, China, [263]Institute of Molecular
385 and Clinical Ophthalmology Basel IOB, Basel, Switzerland, [264]Netherlands Heart Institute, Utrecht,
386 The Netherlands, [265]MRC/UVRI and LSHTM (Uganda Research Unit), Entebbe, Uganda, [266]Faculty of
387 Medicine, Institute of Health Sciences, University of Oulu, Oulu, Finland, [267]Unit of General Practice,
388 Oulu University Hospital, Oulu, Finland, [268]Department of Epidemiology and Public Health, UCL,
389 London, UK, [269]Department of Public Health Solutions, Finnish Institute for Health and Welfare,
390 Helsinki, Finland, [270]Department of Medicine, University of Helsinki and Helsinki University Central
391 Hospital, Helsinki, Finland, [271]Minerva Foundation Institute for Medical Research, Helsinki, Finland,
392 [272]National Heart and Lung Institute, Imperial College London, London, UK, [273]IFB Adiposity Diseases,
393 University of Leipzig Medical Center, Leipzig, Germany, [274]Institute for Social and Economic Research,
394 University of Essex, Colchester, UK, [275]University Institute of Primary Care and Public Health, Division
395 of Biostatistics, University of Lausanne, Lausanne, Switzerland, [276]Institute of Biomedicine, School of
396 Medicine, University of Eastern Finland, Finland, [277]Department of Clinical Physiology and Nuclear
397 Medicine, Kuopio University Hospital, Kuopio, Finland, [278]Foundation for Research in Health Exercise
398 and Nutrition, Kuopio Research Institute of Exercise Medicine, Kuopio, Finland, [279]Institute of
399 Environmental Medicine, Cardiovascular and Nutritional Epidemiology, Karolinska Institutet,
400 Stockholm, Sweden, [280]Department of Medical Sciences, Uppsala, Sweden, [281]Big Data Institute,
401 Nuffield Department of Medicine, University of Oxford, Oxford, UK, [282]Nuffield Department of
402 Women's and Reproductive Health, University of Oxford, Oxford, UK, [283]Department of Medical
403 Epidemiology and Biostatistics and the Swedish Twin Registry, Karolinska Institutet, Stockholm,
404 Sweden, [284]Department of Public Health and Primary Care, Leiden University Medical Center, Leiden,
405 The Netherlands, [285]Institute of Cardiovascular and Medical Sciences, University of Glasgow,
406 Glasgow, UK, [286]Division of Population Health and Genomics, School of Medicine, University of
407 Dundee, Ninewells Hospital and Medical School, Dundee, UK, [287]Centre for Cognitive Ageing and
408 Cognitive Epidemiology, University of Edinburgh, Edinburgh, UK, [288]Department of Health Services,

409 Cardiovascular Health Research Unit, University of Washington, Seattle, WA, USA, [289]Department of
410 Epidemiology, Tulane University School of Public Health and Tropical Medicine, New Orleans, LA,
411 USA, [290]Department of Pediatrics, Genetic and Genomic medicine, University of California, Irvine,
412 Irvine, CA, USA, [291]Havard Medical School, Boston, MA, USA, [292]Tampere, Finnish Diabetes
413 Association, Tampere, Finland, [293]Pirkanmaa Hospital District, Tampere, Finland, [294]Department of
414 Medicine, University of Cambridge, Cambridge, UK, [295]South Karelia Central Hospital, Lappeenranta,
415 Finland, [296]Department of Psychology, University of Miami, Miami, FL, USA, [297]Paul Langerhans
416 Institute Dresden of the Helmholtz Center Munich, University Hospital and Faculty of Medicine,
417 Dresden, Germany, [298]Division of Population Health and Genomics, Ninewells Hospital and Medical
418 School, University of Dundee, Dundee, UK, [299]Division of Sleep and Circadian Disorders, Brigham and
419 Women's Hospital, Boston, MA, USA, [300]Department of Public Health, Section of Epidemiology,
420 Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark,
421 [301]Departent of Molecular and Cellular Therapeutics, Royal College of Surgeons in Ireland, Dublin,
422 Ireland, [302]Department of Aging and Health, Guy’s and St Thomas’ Foundation Trust, London, UK,
423 [303]Cardiovascular and Metabolic Disease Signature Research Program, Duke-NUS Medical School,
424 Singapore, Singapore, [304]Department of Public Health Solutions, National Institute for Health and
425 Welfare, Helsinki, Finland, [305]Department of Public Health, University of Helsinki, Helsinki, Finland,
426 [306]Saudi Diabetes Research Group, King Abdulaziz University, Jeddah, Saudi Arabia, [307]Department of
427 Genomic Medicine and Environmental Toxicology, Instituto de Investigaciones Biomedicas,
428 Universidad Nacional Autonoma de Mexico, Mexico City, Mexico, [308]Department of Public Health
429 and Clinical Nutrition, University of Eastern Finland, Finland, [309]Department of Medicine, Internal
430 Medicine, Lausanne University Hospital (CHUV), Lausanne, Switzerland, [310]Department of Public
431 Health Sciences, Wake Forest School of Medicine, Winston-Salem, NC, USA, [311]Faculty of Medical
432 Sciences, Newcastle University, Newcastle upon Tyne, UK, [312]Beijing Tongren Eye Center, Beijing Key
433 Laboratory of Intraocular Tumor Diagnosis and Treatment, Beijing Ophthalmology & Visual Sciences
434 Key Lab, Beijing Tongren Hospital, Capital Medical University, Beijing, China, China, [313]Department of
435 Public Health, Faculty of Medicine, University of Kelaniya, Ragama, Sri Lanka, [314]Department of
436 Research and Evaluation, Kaiser Permanente of Southern California, Pasadena, CA, USA, [315]Institute
437 for Molecular Bioscience, The University of Queensland, Queensland, Australia, [316]Kurume University
438 School of Medicine, Japan, [317]Wellcome Sanger Institute, Hinxton, UK, [318]TUM School of Medicine,
439 Technical University of Munich and Klinikum Rechts der Isar, Munich, Germany, [319]Department of
440 Pediatrics, Division of Endocrinology, Stanford School of Medicine, Stanford, CA, USA, [320]Wellcome
441 Centre for Human Genetics, Nuffield Department of Medicine, University of Oxford, Oxford, UK,
442 [321]Department of Medicine, Division of General Internal Medicine, Massachusetts General Hospital,
443 Boston, MA, USA, [322]Department of Medicine, General Internal Medicine, Massachusetts General
444 Hospital, Boston, MA, USA, [323]Department of Medicine, Diabetes Unit and Endocrine Unit,
445 Massachusetts General Hospital, Boston, MA, USA, [324]Department of Human Genetics, University of
446 Michigan, Ann Arbor, MI, USA, [325]Centre for Genetics and Genomics Versus Arthritis, Division of
447 Musculoskeletal and Dermatological Sciences, The University of Manchester, Manchester, UK,
448 [326]Centre for Musculoskeletal Research, Division of Musculoskeletal and Dermatological Sciences,
449 The University of Manchester, Manchester, UK, [327]Department of Biostatistics, University of
450 Liverpool, Liverpool, UK
451
452 [#] Denote shared authorship contributions
453 @ Current address: Genentech, South San Francisco, CA
454
455 ***Corresponding author:** Inês Barroso, Exeter Centre of Excellence for Diabetes Research (EXCEED),
456 Exeter Medical School, University of Exeter, Exeter, UK, +44 1392 408221,
457 ines.barroso@exeter.ac.uk
458
459

460
461 **Abstract**

462 Glycemic traits are used to diagnose and monitor type 2 diabetes, and cardiometabolic health. To
463 date, most genetic studies of glycemic traits have focused on individuals of European ancestry. Here,
464 we aggregated genome-wide association studies in up to 281,416 individuals without diabetes (30%
465 non-European ancestry) with fasting glucose, 2h-glucose post-challenge, glycated hemoglobin, and
466 fasting insulin data. Trans-ancestry and single-ancestry meta-analyses identified 242 loci (99 novel;
467 $P<5\times10^{-8}$), 80% with no significant evidence of between-ancestry heterogeneity. Analyses restricted
468 to European ancestry individuals with equivalent sample size would have led to 24 fewer new loci.
469 Compared to single-ancestry, equivalent sized trans-ancestry fine-mapping reduced the number of
470 estimated variants in 99% credible sets by a median of 37.5%. Genomic feature, gene-expression
471 and gene-set analyses revealed distinct biological signatures for each trait, highlighting different
472 underlying biological pathways. Our results increase understanding of diabetes pathophysiology by
473 use of trans-ancestry studies for improved power and resolution.

474 Fasting glucose (FG), 2h-glucose post-challenge (2hGlu), and glycated hemoglobin (HbA1c) are
475 glycemic traits used to diagnose diabetes[1]. In addition, HbA1c is the most commonly used biomarker
476 to monitor glucose control in patients with diabetes. Fasting insulin (FI) reflects a combination of
477 insulin secretion and insulin resistance, both components of type 2 diabetes (T2D), and insulin
478 clearance[2]. Collectively, all four glycemic traits are useful to better understand T2D
479 pathophysiology[3-5] and cardiometabolic outcomes[6].
480
481 To date, genome-wide association studies (GWAS) and analysis of Metabochip and exome arrays
482 have identified >120 loci associated with glycemic traits in individuals without diabetes[7-15]. However,
483 despite considerable differences in the prevalence of T2D risk factors across ancestries[16-18], most
484 glycemic trait GWAS have insufficient representation of individuals of non-European ancestry.
485 Additionally, they have limited resolution for fine-mapping of causal variants and for effector
486 transcript identification. Here, we present large-scale trans-ancestry meta-analyses of GWAS for four
487 glycemic traits in individuals without diabetes. We aimed to identify additional glycemic trait-
488 associated loci; investigate the portability of loci and genetic scores across ancestries; leverage
489 differences in effect allele frequency (EAF), effect size, and linkage disequilibrium (LD) across diverse
490 populations to conduct fine-mapping and aid causal variant/effector transcript identification; and
491 compare the genetic architecture of glycemic traits to further identify the cell-types and target
492 tissues most influenced by these traits which inform T2D pathophysiology.
493
494 **Results**
495 **Study design and definitions**
496 To identify loci associated with glycemic traits FG, 2hGlu, FI, and HbA1c, we aggregated GWAS in up
497 to 281,416 individuals without diabetes, ~30% of whom were of non-European ancestry [13% East
498 Asian, 7% Hispanic, 6% African-American, 3% South Asian, and 2% sub-Saharan African (Ugandan
499 data only available for HbA1c)]. Each cohort imputed data to the 1000 Genomes Project reference
500 panel[19] (phase 1 v3, March 2012, or later; **Methods, Supplementary Table 1, Extended Data Figure
501 1, Supplementary Note**). Up to ~49.3 million variants were directly genotyped or imputed, with
502 between 38.6 million (2hGlu) and 43.5 million variants (HbA1c) available for analysis after exclusions
503 based on minor allele count (MAC < 3) and imputation quality (imputation r^2 or INFO score <0.40) in
504 each cohort. FG, 2hGlu and FI analyses were adjusted for BMI[15] but for simplicity they are
505 abbreviated as FG, 2hGlu and FI (**Methods**).
506
507 We first performed trait-specific fixed-effect meta-analyses *within* each ancestry using METAL[20]
508 (**Methods**). We defined “single-ancestry lead” variants as the strongest trait-associated variants
509 (P<5x10^{-8}) within a 1Mb region in an ancestry (**Table 1**). Within each ancestry and each autosome,
510 we used approximate conditional analyses in GCTA[21,22], to identify “single-ancestry index variants”
511 (P<5x10^{-8}) that exert conditionally distinct effects on the trait (**Table 1**, **Methods**, **Supplementary
512 Note**). This approach identified 124 FG, 15 2hGlu, 48 FI and 139 HbA1c variants that were significant
513 in at least one ancestry (**Supplementary Table 2**).
514
515 Next, we conducted trait-specific *trans-ancestry* meta-analyses using MANTRA (**Methods,
516 Supplementary Table 1, Supplementary Note**) to identify genome-wide significant “trans-ancestry
517 lead variants”, defined as the most significant trait-associated variant across all ancestries ($\log_{10}$
518 Bayes Factor [BF] >6, equivalent to P<5x10^{-8})[23] (**Table 1**, **Methods**). Here, we present trans-ancestry
519 results as our primary results (**Supplementary Table 2**).
520
521 Causal variants are expected to affect related glycemic traits and may be shared across ancestries.
522 Therefore, we combined all single-ancestry lead variants, single-ancestry index variants, and/or
523 trans-ancestry lead variants (for any trait) mapping within 500Kb of each other, into a single “trans-
524 ancestry locus” bounded by 500Kb flanking sequences (**Table 1, Extended Data Figure 2**). As defined,

525 a trans-ancestry locus may contain multiple causal variants affecting one or more glycemic traits,
526 exerting their effect in one or more ancestry.
527
528 **Glycemic trait locus discovery**
529 Trans-ancestry meta-analyses identified 235 trans-ancestry loci, of which 59 contained lead variants
530 for more than one trait. In addition, we identified seven "single-ancestry loci" that did not contain
531 any trans-ancestry lead variants (**Table 1**, **Supplementary Table 2**). Of the 242 combined loci, 99
532 (including 6 of the 7 single-ancestry) had not been previously associated with any of the four
533 glycemic traits or with T2D, at the time of analysis (**Figure 1**, **Supplementary Table 3, Supplementary**
534 **note**). However, based on recent East Asian and trans-ancestry T2D GWAS meta-analyses[23-27], the
535 lead variants at 27/99 novel glycemic trait loci have strong evidence of association with T2D ($P<10^{-4}$;
536 13 loci with $P<5x10^{-8}$), suggesting they are also important in T2D pathophysiology (**Supplementary**
537 **Tables 2 and 4**).
538
539 Of the six single-ancestry novel loci, three were unique to non-European ancestry individuals
540 (**Supplementary Table 3**). An African American association for FI (lead variant rs12056334) near
541 *LOC100128993* (an uncharacterized RNA gene; **Supplementary Note**), an African American
542 association for FG (lead variant rs61909476) near *ETS1* and a Hispanic association for FG (lead
543 variant rs12315677) within *PIK3C2G* (**Supplementary Table 3**). Despite broadly similar EAF across
544 ancestries, rs61909476 was significantly associated with FG only in African American individuals (EAF
545 ~7%, b=0.0812 mmol/l, SE=0.01 mmol/l, $P=3.9\times10^{-8}$ vs EAF 10-17%, b=0-0.002 mmol/l, se=0.003-
546 0.017 mmol/l, P=0.44-0.95 in all other ancestries, **Supplementary table 2, Supplementary note**). The
547 nearest gene, *ETS1*, encodes a transcription factor that is expressed in mouse pancreatic β-cells, and
548 its overexpression decreases glucose-stimulated insulin secretion in mouse islets[28]. Located within
549 the *PIK3C2G* gene, rs12315677 has an 84% EAF in Hispanic (70-94% in other ancestries) and is
550 significantly associated with FG in this ancestry alone (b=0.0387 mmol/l, SE=0.0075 mmol/l,
551 $P=4.0\times10^{-8}$ vs b=-0.0128-0.010 mmol/l, SE=0.003-0.018 mmol/l, P=0.14-0.76 in all other ancestries,
552 **Supplementary note**). In mice, deletion of *Pik3c2g* leads to a phenotype characterized by reduced
553 glycogen storage in the liver, hyperlipidemia, adiposity, and insulin resistance with increasing age, or
554 after a high fat diet[29]. Instances of similar EAFs but differing effect sizes between populations, could
555 be due to genotype-by-environment or other epistatic effects. Alternatively, lower imputation
556 accuracy in smaller sample sizes could deflate effect sizes, although imputation quality for these
557 variants was good (average r^2=0.81). Finally, the variants detected here may be in LD with ancestry-
558 specific causal variants not interrogated here that differ in frequency across ancestries. However, we
559 could not find evidence of rarer alleles in the cognate populations from the 1000G project
560 (**Supplementary Table 5**). The final three single-ancestry loci were identified in individuals of
561 European ancestry (**Supplementary note**).
562
563 Next, by rescaling the standard errors of allelic effect sizes to artificially boost the sample size of the
564 European meta-analysis to match that of trans-ancestry meta-analysis, we determined that 21 of the
565 novel trans-ancestry loci would not have been discovered with an equivalent sample size comprised
566 exclusively of European ancestry individuals (**Supplementary note**). Their discovery was due to the
567 higher EAF and/or larger effect size in non-European ancestry populations. In particular, two loci
568 (near *LINC0088*5 and *MIR4278*) contain East Asian and African American single-ancestry lead
569 variants, respectively, suggesting that these specific ancestries may be driving the trans-ancestry
570 discovery (**Supplementary Tables 2-3**). Combined with the three single-ancestry non-European loci
571 described above, our results show that 24% (24/99) of novel loci were discovered due to the
572 contribution of non-European ancestry participants, strengthening the argument for expanding
573 genetic studies in diverse populations.
574
575 **Allelic architecture of glycemic traits**

576 Single-ancestry and trans-ancestry results combined increased the number of established loci for FG
577 to 102 (182 signals, 53 novel loci), FI to 66 (95 signals, 49 novel loci), 2hGlu to 21 (28 signals, 11
578 novel loci), and HbA1c to 127 (218 signals, 62 novel loci) (**Supplementary Table 2**), with significant
579 overlap across traits (**Extended Data Figure 3**). We also detected (*P*<0.05 or $\log_{10}$BF>0) the vast
580 majority (~90%) of previously established glycemic signals, 70-88% of which attained genome-wide
581 significance (**Supplementary Note**, **Supplementary Table 6**). Given that analyses for FG, FI, and
582 2hGlu were performed adjusted for BMI, we confirmed that collider bias did not influence >98% of
583 signals discovered (**Supplementary note**)[31]. As expected, given the greater power due to increased
584 sample sizes, new association signals tended to have smaller effect sizes and/or EAFs in European
585 ancestry individuals compared to established signals (**Extended Data Figure 4**).
586
587 **Characterization of lead variants across ancestries**
588 To better understand the transferability of trans-ancestry lead variants across ancestries, we
589 investigated the pairwise EAF correlation and the pairwise summarized heterogeneity of effect sizes
590 between ancestries[32] (**Methods**, **Supplementary Note**). Consistent with population history and
591 evolution, these results demonstrated considerable EAF correlation (ρ^2>0.70) between European
592 and Hispanic, European and South Asian, and Hispanic and South Asian populations, consistent
593 across all four traits, and between African Americans and Ugandans for HbA1c (**Extended Data**
594 **Figure 5**). Despite significant EAF correlations, some pairwise comparisons exhibited strong evidence
595 for effect size heterogeneity between ancestries that was less consistent between traits (**Extended**
596 **Data Figure 5).** However, sensitivity analyses demonstrated that, across all comparisons, the
597 evidence for heterogeneity is driven by a small number of variants, with between 81.5% (for HbA1c)
598 and 85.7% of trans-ancestry lead variants (for FG) showing no evidence for trans-ancestry
599 heterogeneity (*P*>0.05) (**Supplementary Note**).
600
601 **Trait variance explained by associated loci**
602 The trait variance explained by genome-wide significant loci was assessed using the single-ancestry
603 variants only or a combination of single-ancestry and trans-ancestry variants (**Supplementary Table**
604 **7**) with betas extracted from the relevant single-ancestry meta-analysis results (**Methods**). The
605 variance explained was assessed by linear regression in a subset of the contributing cohorts
606 (**Methods**, **Supplementary Tables 8-11**). In general, the approach that explained the most variance
607 was to begin with the trans-ancestry lead variants that had *P*<0.1 in the relevant single-ancestry
608 meta-analysis, then add in all single-ancestry variants that were not in LD with the trans-ancestry
609 variants (LD r^2<0.1) (List C, **Supplementary Tables 8-11**, **Figure 2**). Using this approach, the mean
610 variance in the trait distribution explained was between 0.7% (2hGlu in EUR) and 6% (HbA1c in AA).
611 The European-based estimates explained more variance relative to previous estimates of 2.8% for
612 FG and 1.7% for HbA1c[33] (**Supplementary Note**).
613
614 **Transferability of EUR ancestry-derived polygenic scores**
615 To investigate the transferability of polygenic scores across ancestries we used the PRS-CSauto
616 software[34] to first build polygenic scores for each glycemic trait based on European ancestry data.
617 However, the training set for 2hGlu was too small so this trait was excluded. To build the polygenic
618 scores (PGS), for each trait we first removed five of the largest European cohorts from the European
619 ancestry meta-analysis**.** These five cohorts were meta-analyzed and used as our European ancestry
620 test dataset, for each trait. The remaining European ancestry cohorts were also meta-analyzed and
621 used as the training dataset, from which we derived a PGS for each trait (**Methods**). We used PRS-
622 CSauto to revise the effect size estimates for the variants in the score (obtained from the training
623 European datasets) based on the LD of the test population. PRS-CSauto does not have LD reference
624 panels for South Asian or Hispanic ancestry and as such we were unable to test the transferability of
625 the PGS into those populations. The “gtx” package[35] (**Methods**) was used to obtain the R^2 for each
626 test population (**Figure 3, Supplementary Table 12**). Consistent with other complex traits[36], the

627 European ancestry-derived PGS had greater predictive power into test data of European ancestry
628 than other ancestry groups.
629

630 **Fine-mapping**

631 We fine-mapped, 231 trans-ancestry and six single-ancestry autosomal loci (**Supplementary Table 2,**
632 **Supplementary note**). Using FINEMAP with ancestry-specific LD and an average LD matrix across
633 ancestries, we conducted fine-mapping both within (161 loci with single-ancestry lead variants) and
634 across ancestries (231 loci) for each trait (**Methods**). Because 59 of the 231 trans-ancestry loci were
635 associated with more than one trait, we conducted trans-ancestry fine-mapping for a total of 305
636 locus-trait associations. Of these 305 locus-trait combinations, FINEMAP estimated the presence of a
637 single causal variant at 186 loci (61%), while multiple distinct causal variants were implicated at 126
638 loci (39%), for a total of 464 causal variants (**Figure 4A**).
639

640 *Credible sets for causal variants*

641 At each locus, we next constructed credible sets (CS) for each causal variant that account for >=99%
642 of the posterior probability of association (PPA). We identified 21 locus-trait associations (at 19 loci)
643 for which the 99% CS included a single variant, and we highlight four examples (**Methods,**
644 **Supplementary Note**, **Figure 4B, Supplementary Table 13**).
645

646 At *MTNR1B* and *SIX3* we identified, respectively, rs10830963 (PPA>0.999, for both HbA1c and FG)
647 and rs12712928 (PPA=0.997, for FG) as the likely causal variants. At both loci previous studies
648 confirm these variants affect transcriptional activity[37,38,39] (**Supplementary note**). At a locus near
649 *PFKM* associated with HbA1c, trans-ancestry fine-mapping identified rs12819124 (PPA>0.999) as the
650 likely causal variant. This variant has been previously associated with mean corpuscular
651 hemoglobin[40], suggesting an effect on HbA1c via the red blood cell (RBC, **Supplementary note**). At
652 *HBB*, we identifed rs334 (PPA>0.999; Glu7Val) as the likely causal variant associated with HbA1c.
653 rs334 is a causal variant of sickle cell anemia[41], previously associated with urinary albumin-to-
654 creatinine ratio in Caribbean Hispanic individuals[42], severe malaria in a Tanzanian study population[43],
655 hematocrit and mean corpuscular volume in Hispanic/Latino populations[44], and RBC distribution in
656 Ugandan individuals[45], all pointing to a variant effect on HbA1c via non-glycemic pathways.
657

658 The remaining locus-trait associations with a single variant in the 99% CS (**Supplementary Table 13**)
659 point to variants that could be prioritized for functional follow-up to elucidate impact on glycemic
660 trait physiology.
661

662 At an additional 156 locus-trait associations trans-ancestry fine-mapping identified 99% CS with 50
663 or fewer variants (**Figure 4B, Supplementary Table 13**). Consistent with the potential for >1 causal
664 variant in a locus, 74 locus-trait associations contained 88 variants with PPA>0.90 that are strong
665 candidate causal variants (**Supplementary Table 14**). For example, 10 are coding variants including
666 several missense such as the *HBB* Glu7Val mentioned above, *GCKR* Leu446Pro, *RREB1* Asp1771Asn,
667 *G6PC2* Pro324Ser, *GLP1R* Ala316Thr, and *TMPRSS6* Val736Ala**,** each of which have been proposed or
668 shown to affect gene function[12,46-50]. We additionally identified *AMPD3* Val311Leu (PPA=0.989) and
669 *TMC6* Trp125Arg (PPA>0.999) variants associated with HbA1c which were previously detected in an
670 exome array analysis but had not been fine-mapped with certainty due to the absence of backbone
671 GWAS data[30]. Our fine-mapping now suggest these variants are likely causal and identify their
672 cognate genes as effector transcripts.
673

674 Finally, we evaluated the resolution obtained in the trans-ancestry versus single-ancestry fine-
675 mapping (**Methods, Supplementary Note**). We compared the number of variants in 99% CS across
676 98 locus-trait associations which, as suggested by FINEMAP, had a single causal variant in both trans-
677 ancestry and single-ancestry analyses. Fine-mapping within and across ancestries was conducted

678 using the same set of variants. At 8 of 98 locus-trait associations single-ancestry fine-mapping
679 identified a single variant in the CS. In addition, at 72 of the 98 locus-trait associations, the number
680 of variants in the 99% CS was smaller in the trans-ancestry fine-mapping (**Figure 4C**), which likely
681 reflects the larger sample size and differences in LD structure, EAFs, and effect sizes across diverse
682 populations. To quantify the estimated improvement in fine-mapping resolution attributable to the
683 multi-ancestry GWAS, we then compared 99% CS sizes from the trans-ancestry fine-mapping to
684 single-ancestry-specific data emulating the same total sample size by rescaling the standard errors
685 (**Methods**). Of the 72 locus-trait associations with estimated improved fine-mapping in trans-
686 ancestry analysis, resolution at 38 (53%) was improved because of the larger sample size in the
687 trans-ancestry fine-mapping analysis (**Figure 4C**), and this estimated improved resolution would
688 likely have been obtained in a European-only fine-mapping effort with equivalent sample size.
689 However, at 34 (47%) loci, the inclusion of samples from multiple diverse populations yielded the
690 estimated improved resolution. On average, ancestry differences led to a reduction in the median
691 number of variants in the 99% CS from 24 to 15 variants (37.5% median reduction; **Figure 4C**),
692 demonstrating the value of conducting fine-mapping across ancestries.

693

694 **HbA1c Signal Classification**

695 HbA1c-associated variants can exert their effects on HbA1c levels through both glycemic and non-
696 glycemic pathways [7,51] and their correct classification can affect T2D diagnostic accuracy[7,52]. Using
697 prior association results for other glycemic, RBC, and iron traits, and a fuzzy clustering approach we
698 classified variants into their most likely mode of action (**Methods**, **Supplementary note**). Of the 218
699 HbA1c-associated variants, 27 (12%) could not be characterized due to missing data and 23 (11%)
700 could not be classified into a "known" class (**Supplementary note**). The remaining signals were
701 classified as principally: a) glycemic (n=53; 24%), b) affecting iron levels/metabolism (n=12; 6%), or c)
702 RBC traits (n=103; 47%). A genetic risk score (GRS) composed of all HbA1c-associated signals was
703 strongly associated with T2D risk (OR=2.4, 95% CI 2.3-2.5, $P=2.7\times10^{-298}$). However, when using
704 partitioned GRSs composed of these different classes of variants (**Methods**), we found the T2D
705 association was mainly driven by variants influencing HbA1c through glycemic pathways (OR=2.6,
706 95% CI 2.5-2.8, $P=2.3\times10^{-250}$), with weaker evidence of association (despite the larger number of
707 variants in the GRS) and a more modest risk (OR=1.4, 95% CI 1.2-1.7, $P=4.7\times10^{-4}$) imparted by signals
708 in the mature RBC cluster that were not glycemic (i.e. where those specific variants had $P>0.05$ for
709 FI, 2hGlu and FG) (**Extended Data Figure 6, Supplementary note**). This contrasts our previous finding
710 where we found no significant association between a risk score of non-glycemic variants and T2D[7].
711 Our current results could be partly driven by T2D cases being diagnosed based on HbA1c levels that
712 may be influenced by the non-glycemic signals, or by glycemic effects not captured by FI, 2hGlu or
713 FG measures.
714

715 **Biological signatures of glycemic trait associated loci**

716 To better understand distinct and shared biological signatures underlying variant-trait associations,
717 we conducted genomic feature enrichment, eQTL co-localization, and tissue and gene-set
718 enrichment analyses across all four traits.
719

720 **Epigenomic landscape of trait-associated variants**
721 We explored the genomic context underlying glycemic trait loci by computing overlap enrichment
722 for annotations such as coding, conserved regions, and super enhancers merged across multiple cell
723 types[53-55] using the GREGOR tool[56]. We observed that FG, FI and HbA1c signals (**Supplementary**
724 **Table 7**) were significantly ($P<8.4\times10^{-4}$, Bonferroni threshold for 59 annotations) enriched in
725 evolutionarily conserved regions (**Fig 5A, Extended Data Figure 7, Supplementary Table 15**).

726
727 We then considered epigenomic landscapes defined in individual cell/tissue types. Previously,
728 stretch enhancers (StrE, enhancer chromatin states ≥3kb in length) in pancreatic islets were shown
729 to be highly cell-specific and strongly enriched with T2D risk signals[57]. Considering StrEs across 31
730 cell-types[39], FG and 2hGlu signals showed the highest enrichment in islets (FG: fold-
731 enrichment=4.70, P=2.7x10^{-24}; 2hGlu: fold-enrichment=5.51, P=3.6x10^{-4} **Figure 5A, Supplementary**
732 **Table 16**), highlighting the importance of islets for these traits. FI signals were enriched in skeletal
733 muscle (fold-enrichment=3.17, P=7.8x10^{-6}) and adipose StrEs (fold-enrichment=3.27, P=1.8x10^{-7})
734 consistent with these tissues as targets of insulin action (**Figure 5A**). StrEs in individual cell types
735 showed higher enrichment than super enhancers merged across cell types, highlighting the
736 importance of cell-specific analyses (**Figure 5A**). HbA1c signals were enriched in StrEs of multiple cell
737 types and tissues, but have the strongest enrichment in K562 leukemia derived cells (fold-
738 enrichment=3.24, P=1.2x10^{-7}, **Figure 5A**). Among the "hard" glycemic and red blood cell (mature +
739 reticulocyte) HbA1c signals, glycemic signals were enriched in islet StrEs (fold-enrichment=3.96,
740 P=3.7x10^{-16}) while red blood cell signals were enriched in K562 StrEs (fold-enrichment=7.5,
741 P=2.08x10^{-14}, **Figure 5B, Supplementary Table 17**). These analyses suggest that these glycemic trait-
742 associated variants influence the function of tissue-specific enhancers.
743
744 Independent analyses with fGWAS[58] and GARFIELD[59] yielded consistent results (**Extended Data**
745 **Figures 8 and 9, Supplementary Tables 16 and 18**). Notably, FI signals at a lenient threshold of P<10^{-}
746 ^{5} were enriched in liver StrEs using GARFIELD (odds ratio=1.92, P=1.7x10^{-4}) (**Extended Data Figure**
747 **9A**). This suggests that liver regulatory annotations are relevant for FI GWAS signals, but that we lack
748 power to detect significant enrichment using the genome-wide significant loci and the current set of
749 reference annotations.
750
751 We next explored the 27 loci driving the FI enrichment in adipose and skeletal muscle, 11 of which
752 overlapped StrEs in both tissues (**Figure 5C**). At the *COL4A2* locus, variants within an intronic region
753 overlap StrEs in adipose tissue, skeletal muscle, and a human skeletal muscle myoblast (HSMM) cell
754 line that are not shared across other cell/tissue types. Among these, rs9555695 (in the 99% CS) also
755 overlaps accessible chromatin regions in adipose (**Figure 5D**). At a narrow signal with no proxy
756 variants (LD r^2>0.7 in Europeans), the lead trans-ancestry variant rs62271373 (PPA = 0.94) located in
757 an intergenic region ~25kb from the *LINC01214* gene overlaps StrEs specific to adipose and HSMM
758 and an active enhancer chromatin state in skeletal muscle (**Figure 5E**). Collectively, the tissue-
759 specific epigenomic signatures at GWAS signals provide an opportunity to nominate tissues where
760 these variants are likely to be active. This map may help future efforts to deconvolute GWAS signals
761 into tissue-specific disease pathology.
762
763 **Co-localization of GWAS and eQTLs**
764 Among the 99 novel glycemic trait loci, we identified co-localized eQTLs at 34 loci in blood,
765 pancreatic islets, subcutaneous or visceral adipose, skeletal muscle, or liver, providing suggestive
766 evidence of causal genes (**Supplementary Table 19**). The co-localized eQTLs include several genes
767 previously reported at glycemic trait loci: *ADCY5*, *CAMK1D, IRS1*, *JAZF1*, and *KLF14*[60-62]. For some
768 additional loci, the co-localized genes have prior evidence for a role in glycemic regulation. For
769 example, the lead trans-ancestry variant and likely causal variant, rs1799815 (PPA=0.993),
770 associated with FI is the strongest variant associated with expression of *INSR*, encoding the insulin
771 receptor, in subcutaneous adipose from METSIM (P=2x10^{-9}) and GTEx (P=5x10^{-6}). The A allele at
772 rs1799815 is associated with higher FI and lower expression of *INSR*, consistent with the relationship
773 between insulin resistance and reduced INSR function[63]. In a second example, rs841572, the trans-
774 ancestry lead variant associated with FG, has the highest PPA (PPA=0.535) among the 20 variants in
775 the 99% CS and is in strong LD (r^2=0.87) with the lead eQTL variant (rs841576, also in the 99% CS)
776 associated with *SLC2A1* expression in blood (eQTLGen P=1x10^{-8}). *SLC2A1*, also known as *GLUT1*,

777 encodes the major glucose transporter in brain, placenta, and erythrocytes, and is responsible for
778 glucose entry into the brain[64]. rs841572-A is associated with lower FG and lower *SLC2A1* expression.
779 While rare missense variants in *SLC2A1* are an established cause of seizures and epilepsy[65], our data
780 suggest that *SLC2A1* variants also affect plasma glucose levels within a population. These co-
781 localized signals provide possible regulatory mechanisms for variant effects on genes to influence
782 glycemic traits.
783
784 The co-localized eQTLs also provide new insights into the mechanisms at glycemic trait loci. For
785 example, rs9884482 (in the 99% CS) is associated with FI and *TET2* expression in subcutaneous
786 adipose (P=2x10^{-20}); rs9884482 is in high LD (r^2=0.96 in Europeans) with the lead *TET2* eQTL variant
787 (rs974801). *TET2* encodes a DNA-demethylase that can affect transcriptional repression [66]. Adipose
788 Tet2 expression is reduced in diet-induced insulin resistance in mice[67], and knockdown of Tet2
789 blocked adipogenesis[67,68]. Consistently, in human adipose tissue, rs9884482-C was associated with
790 lower *TET2* expression and higher FI. In a second example, rs617948 is associated with HbA1c (in the
791 99% CS) and is the lead variant associated with *C2CD2L* expression in blood (eQTLGen P=3x10^{-96}).
792 *C2CD2L*, also known as *TMEM24,* regulates pulsatile insulin secretion and facilitates release of
793 insulin pool reserves[69,70]. rs617948-G was associated with higher HbA1c and lower *C2CD2L*, providing
794 evidence for a role of this insulin secretion protein in glucose homeostasis. Our HbA1c "soft"
795 clustering assigned this signal to both the "unknown" (0.51 probability) and "reticulocyte" (0.42
796 probability) clusters. rs617948 is strongly associated with HbA1c (P<6.8x10^{-8}), but not with FG, FI or
797 2hGlu (P>0.05, **Supplementary Table 20, Supplementary Note**). This suggests an effect of this
798 variant on reticulocyte biology, and on insulin secretion, potentially influencing HbA1c levels through
799 different tissues, and providing a plausible explanation for the classification as "unknown".
800
801 **Tissue Expression**
802 Consistent with effector transcript expression analysis using GTEx data[30], we found significant
803 differences in tissue expression across the glycemic trait signals. FG signals were enriched for genes
804 expressed in the pancreas (FDR<0.05), while there were an insufficient number of significant
805 associations in 2hGlu to identify enrichment for any tissue or cell type at FDR<0.2 threshold. FI
806 signals were enriched for connective tissue and cells (which includes adipose tissue), endocrine
807 glands, blood cells, and muscles (FDR<0.2) and HbA1c signals were significantly enriched for genes
808 expressed in the pancreas, hemic, and immune system (FDR<0.05) (**Figure 6, Supplementary Table**
809 **21**). Consistent with previous analysis[30], FI-enrichment for connective tissue was driven by adipose
810 tissue (subcutaneous and visceral), while the newly described enrichment with endocrine glands was
811 driven by the adrenal glands and cortex (**Supplementary Table 21**). Beyond enrichment for genes
812 expressed in glycemic-related tissues, HbA1c signals were enriched with genes expressed in blood,
813 consistent with the role of RBC in this trait and our previous results[30].
814
815 The association between FI signals and genes expressed in adrenal glands is notable, suggesting a
816 possible direct role for these genes in insulin resistance. These genes might influence cortisol levels,
817 which could contribute to insulin resistance and FI levels through impaired insulin receptor signaling
818 in peripheral tissues, as well as influencing body fat distribution, stimulate lipolysis, and other
819 indirect mechanisms[71,72].
820
821
822 **Gene-set Analyses**
823 Next, we performed gene-set analysis using DEPICT (**Methods**). In keeping with previous results[30],
824 we found distinct gene-sets enriched (FDR<0.05) for each glycemic trait except 2hGlu, which had
825 insufficient associations to have power in this analysis. FG-associated variants highlighted gene-sets
826 involved in metabolism and gene-sets involved in general cellular function such as "cytoplasmic
827 vesicle membrane" and "circadian clock"" (**Figure 7A**). In contrast, in addition to metabolism-related

828 gene-sets, FI-associated variants highlighted pathways related to growth, cancer and reproduction
829 (**Figure 7B**). This is consistent with the role of insulin as a mitogenic hormone, and with
830 epidemiological links between insulin and certain types of cancer[73] and reproductive disorders such
831 as polycystic ovary syndrome[74]. HbA1c-associated variants highlighted many gene-sets (**Figure 7C**),
832 including those linked to metabolism and hematopoiesis, again recapitulating our postulated effects
833 of variants on glucose and RBC biology. Additional pathways from HbA1c-associated variants also
834 highlighted previous "CREBP PPI" and lipid biology related to T2D[75] and HbA1c[76], respectively, and
835 potential new biology through which variants may influence HbA1c.
836
837 **Discussion**
838 Here we describe a large glycemic trait meta-analysis of GWAS for which 30% of the population was
839 composed of East Asian, Hispanic, African-American, South Asian and sub-Saharan African
840 participants. This effort identified 242 loci (235 trans-ancestry and seven single-ancestry), which
841 jointly explain between 0.7% (2hGlu in European ancestry individuals) and 6% (HbA1c in African
842 American ancestry individuals) of the variance in glycemic traits in any given ancestry. While
843 114/242 loci are associated with T2D ($P<10^{-4}$; 83 loci with $P<5x10^{-8}$, **Supplementary Table 4**),
844 absence of strong evidence of association at the remaining loci ($P≥10^{-4}$) suggests that for alleles
845 more frequent than 5% we can exclude T2D ORs≥1.07 with 80% power (alpha=$5x10^{-8}$; and ORs≥1.05
846 for alpha=10^{-4}) given a current study of 228,499 T2D cases and 1,178,783 controls[27]. We identified
847 486 signals associated with glycemic traits, of which eight have MAF<1%, and 45 have 1%<=MAF<5%
848 in all ancestries, highlighting that 89% of signals identified are common in at least one ancestry
849 studied.
850
851 A key aim of our study was to evaluate the added advantage of including population diversity in
852 genetic discovery and fine-mapping efforts. Beyond the larger sample size included in the trans-
853 ancestry meta-analysis, we were able to estimate the contribution of non-European ancestry data in
854 locus discovery and fine-mapping resolution. We found that 24 of the 99 newly discovered loci owe
855 their discovery to the inclusion of East Asian, Hispanic, African-American, South Asian and sub-
856 Saharan African participant data, due to differences in EAF and effect sizes across ancestries.
857
858 Comparison of 295 trans-ancestry lead variants (315 locus-trait associations) across ancestries
859 demonstrated that between 81.5% (HbA1c) and 85.7% (FG) of the trans-ancestry lead variants had
860 no evidence of trans-ancestry heterogeneity in allelic effects ($P>0.05$).
861
862 Given sample size and power limitations, genome-wide significant trait-associated variants in a
863 single-ancestry explain only a modest proportion of trait variance in that ancestry (**Figure 2**). We
864 demonstrate that trans-ancestry lead variants explain more trait variance than the ancestry-specific
865 variants (**Figure 2**). This shows that even though some trans-ancestry lead variants are not genome-
866 wide significant in all ancestries, they contribute to the genetic architecture of the trait in most
867 ancestries.
868
869 We evaluated for the first time the transferability of European ancestry-derived glycemic trait PGS
870 into other ancestries. Consistent with other traits[36,77,78], we confirm that European ancestry-derived
871 PGS perform much worse when the test dataset is from a different ancestry. Each trait-specific PGS
872 improves trait variance explained by between 3.5-fold (HbA1c) and 6-fold (FG) in the European
873 dataset (**Figure 3, Supplementary Table 12**) compared to a score built only from trans-ancestry lead
874 variants and European index variants (**Figure 2, Supplementary tables 9-12**).
875
876 Despite development of approaches to derive polygenic risk scores[79], we note the difficulty in using
877 summary level data to build a PGS in one ancestry and then apply it in test datasets of different
878 ancestry. While PRS-CSauto[34] is able to use summary level data, revision of the effect size estimates

879 to account for LD required reference panels that matched the ancestry of the test dataset. However,
880 the current software lacks appropriate reference panels for many ancestries, precluding its broad
881 application. Future developments of trans-ancestry PGS are required for improved cross-ancestry
882 performance.
883
884 We show that fine-mapping resolution is improved in trans-ancestry, compared to single-ancestry
885 fine-mapping efforts. In ~50% of our loci, we showed that the improvement was due to differences
886 in EAF, effect size, or LD structure between ancestries, and not just due to the overall increased
887 sample size available for trans-ancestry fine-mapping. By performing trans-ancestry fine-mapping,
888 and co-localizing GWAS signals with eQTL signals and coding variants, we identified new candidate
889 causal genes. Altogether, these results motivate continued expansion of genetic and genomic efforts
890 in diverse populations to improve understanding of these traits in groups disproportionally affected
891 by T2D.
892
893 Given data on four different glycemic traits and their utility to diagnose and monitor T2D and
894 metabolic health, we also sought to characterize biological features underlying these traits. We
895 show that despite significant sharing of loci across the four traits, each trait is also characterized by
896 unique features based on StrE, gene expression and gene-set signatures. Combining genetic data
897 from these traits with T2D data will further elucidate pathways driving normal physiology and
898 pathophysiology, and help further develop useful predictive scores for disease classification and
899 management[4,5].
900
901
902 **References**
903
904 1 in *Use of Glycated Haemoglobin (HbA1c) in the Diagnosis of Diabetes Mellitus:*
905 *Abbreviated Report of a WHO Consultation* (World Health Organization
906 Copyright © World Health Organization 2011., 2011).
907 2 Goodarzi, M. O. *et al.* Fasting insulin reflects heterogeneous physiological processes:
908 role of insulin clearance. *American journal of physiology. Endocrinology and*
909 *metabolism* **301**, E402-408, doi:10.1152/ajpendo.00013.2011 (2011).
910 3 Dimas, A. S. *et al.* Impact of type 2 diabetes susceptibility variants on quantitative
911 glycemic traits reveals mechanistic heterogeneity. *Diabetes* **63**, 2158-2171,
912 doi:10.2337/db13-0949 (2014).
913 4 Udler, M. S. *et al.* Type 2 diabetes genetic loci informed by multi-trait associations
914 point to disease mechanisms and subtypes: A soft clustering analysis. *PLoS medicine*
915 **15**, e1002654, doi:10.1371/journal.pmed.1002654 (2018).
916 5 Udler, M. S., McCarthy, M. I., Florez, J. C. & Mahajan, A. Genetic Risk Scores for
917 Diabetes Diagnosis and Precision Medicine. *Endocrine reviews* **40**, 1500-1520,
918 doi:10.1210/er.2019-00088 (2019).
919 6 Sarwar, N. *et al.* Diabetes mellitus, fasting blood glucose concentration, and risk of
920 vascular disease: a collaborative meta-analysis of 102 prospective studies. *Lancet*
921 **375**, 2215-2222, doi:10.1016/s0140-6736(10)60484-9 (2010).
922 7 Wheeler, E. *et al.* Impact of common genetic determinants of Hemoglobin A1c on
923 type 2 diabetes risk and diagnosis in ancestrally diverse populations: A transethnic
924 genome-wide meta-analysis. *PLoS medicine* **14**, e1002383,
925 doi:10.1371/journal.pmed.1002383 (2017).

926 8 Dupuis, J. *et al.* New genetic loci implicated in fasting glucose homeostasis and their
927 impact on type 2 diabetes risk. *Nature genetics* **42**, 105-116, doi:10.1038/ng.520
928 (2010).
929 9 Manning, A. K. *et al.* A genome-wide approach accounting for body mass index
930 identifies genetic variants influencing fasting glycemic traits and insulin resistance.
931 *Nature genetics* **44**, 659-669, doi:10.1038/ng.2274 (2012).
932 10 Walford, G. A. *et al.* Genome-Wide Association Study of the Modified Stumvoll
933 Insulin Sensitivity Index Identifies BCL2 and FAM19A2 as Novel Insulin Sensitivity
934 Loci. *Diabetes* **65**, 3200-3211, doi:10.2337/db16-0199 (2016).
935 11 Horikoshi, M. *et al.* Discovery and Fine-Mapping of Glycaemic and Obesity-Related
936 Trait Loci Using High-Density Imputation. *PLoS genetics* **11**, e1005230,
937 doi:10.1371/journal.pgen.1005230 (2015).
938 12 Mahajan, A. *et al.* Identification and functional characterization of G6PC2 coding
939 variants influencing glycemic traits define an effector transcript at the G6PC2-
940 ABCB11 locus. *PLoS genetics* **11**, e1004876, doi:10.1371/journal.pgen.1004876
941 (2015).
942 13 Hwang, J. Y. *et al.* Genome-wide association meta-analysis identifies novel variants
943 associated with fasting plasma glucose in East Asians. *Diabetes* **64**, 291-298,
944 doi:10.2337/db14-0563 (2015).
945 14 Chen, P. *et al.* Multiple nonglycemic genomic loci are newly associated with blood
946 level of glycated hemoglobin in East Asians. *Diabetes* **63**, 2551-2562,
947 doi:10.2337/db13-1815 (2014).
948 15 Scott, R. A. *et al.* Large-scale association analyses identify new loci influencing
949 glycemic traits and provide insight into the underlying biological pathways. *Nature*
950 *genetics* **44**, 991-1005, doi:10.1038/ng.2385 (2012).
951 16 Spanakis, E. K. & Golden, S. H. Race/ethnic difference in diabetes and diabetic
952 complications. *Current diabetes reports* **13**, 814-823, doi:10.1007/s11892-013-0421-
953 9 (2013).
954 17 Tillin, T. *et al.* Insulin resistance and truncal obesity as important determinants of the
955 greater incidence of diabetes in Indian Asians and African Caribbeans compared with
956 Europeans: the Southall And Brent REvisited (SABRE) cohort. *Diabetes care* **36**, 383-
957 393, doi:10.2337/dc12-0544 (2013).
958 18 Whincup, P. H. *et al.* Early emergence of ethnic differences in type 2 diabetes
959 precursors in the UK: the Child Heart and Health Study in England (CHASE Study).
960 *PLoS medicine* **7**, e1000263, doi:10.1371/journal.pmed.1000263 (2010).
961 19 Auton, A. *et al.* A global reference for human genetic variation. *Nature* **526**, 68-74,
962 doi:10.1038/nature15393 (2015).
963 20 Willer, C. J., Li, Y. & Abecasis, G. R. METAL: fast and efficient meta-analysis of
964 genomewide association scans. *Bioinformatics (Oxford, England)* **26**, 2190-2191,
965 doi:10.1093/bioinformatics/btq340 (2010).
966 21 Yang, J., Lee, S. H., Goddard, M. E. & Visscher, P. M. GCTA: a tool for genome-wide
967 complex trait analysis. *American journal of human genetics* **88**, 76-82,
968 doi:10.1016/j.ajhg.2010.11.011 (2011).
969 22 Yang, J. *et al.* Conditional and joint multiple-SNP analysis of GWAS summary statistics
970 identifies additional variants influencing complex traits. *Nature genetics* **44**, 369-375,
971 s361-363, doi:10.1038/ng.2213 (2012).

972 23 Genome-wide association study of 14,000 cases of seven common diseases and
973 3,000 shared controls. *Nature* **447**, 661-678, doi:10.1038/nature05911 (2007).
974 24 Mahajan, A. *et al.* Trans-ancestry genetic study of type 2 diabetes highlights the
975 power of diverse populations for discovery and translation. *medRxiv*,
976 2020.2009.2022.20198937, doi:10.1101/2020.09.22.20198937 (2020).
977 25 Mahajan, A. *et al.* Fine-mapping type 2 diabetes loci to single-variant resolution
978 using high-density imputation and islet-specific epigenome maps. *Nature genetics*
979 **50**, 1505-1513, doi:10.1038/s41588-018-0241-6 (2018).
980 26 Spracklen, C. N. *et al.* Identification of type 2 diabetes loci in 433,540 East Asian
981 individuals. *Nature* **582**, 240-245, doi:10.1038/s41586-020-2263-3 (2020).
982 27 Vujkovic, M. *et al.* Discovery of 318 new risk loci for type 2 diabetes and related
983 vascular outcomes among 1.4 million participants in a multi-ancestry meta-analysis.
984 *Nature genetics* **52**, 680-691, doi:10.1038/s41588-020-0637-y (2020).
985 28 Luo, Y. *et al.* Transcription factor Ets1 regulates expression of thioredoxin-interacting
986 protein and inhibits insulin secretion in pancreatic beta-cells. *PloS one* **9**, e99049,
987 doi:10.1371/journal.pone.0099049 (2014).
988 29 Braccini, L. *et al.* PI3K-C2gamma is a Rab5 effector selectively controlling endosomal
989 Akt2 activation downstream of insulin signalling. *Nature communications* **6**, 7400,
990 doi:10.1038/ncomms8400 (2015).
991 30 Ng, N. H. J. *et al.* Tissue-Specific Alteration of Metabolic Pathways Influences
992 Glycemic Regulation. *bioRxiv*, 790618, doi:10.1101/790618 (2019).
993 31 Aschard, H., Vilhjalmsson, B. J., Joshi, A. D., Price, A. L. & Kraft, P. Adjusting for
994 heritable covariates can bias effect estimates in genome-wide association studies.
995 *American journal of human genetics* **96**, 329-339, doi:10.1016/j.ajhg.2014.12.021
996 (2015).
997 32 Lee, J. J. *et al.* Gene discovery and polygenic prediction from a genome-wide
998 association study of educational attainment in 1.1 million individuals. *Nature*
999 *genetics* **50**, 1112-1121, doi:10.1038/s41588-018-0147-3 (2018).
1000 33 Nolte, I. M. *et al.* Missing heritability: is the gap closing? An analysis of 32 complex
1001 traits in the Lifelines Cohort Study. *European journal of human genetics : EJHG* **25**,
1002 877-885, doi:10.1038/ejhg.2017.50 (2017).
1003 34 Ge, T., Chen, C. Y., Ni, Y., Feng, Y. A. & Smoller, J. W. Polygenic prediction via
1004 Bayesian regression and continuous shrinkage priors. *Nature communications* **10**,
1005 1776, doi:10.1038/s41467-019-09718-5 (2019).
1006 35 Dastani, Z. *et al.* Novel loci for adiponectin levels and their influence on type 2
1007 diabetes and metabolic traits: a multi-ethnic meta-analysis of 45,891 individuals.
1008 *PLoS genetics* **8**, e1002607, doi:10.1371/journal.pgen.1002607 (2012).
1009 36 Martin, A. R. *et al.* Clinical use of current polygenic risk scores may exacerbate health
1010 disparities. *Nature genetics* **51**, 584-591, doi:10.1038/s41588-019-0379-x (2019).
1011 37 Gaulton, K. J. *et al.* Genetic fine mapping and genomic annotation defines causal
1012 mechanisms at type 2 diabetes susceptibility loci. *Nature genetics* **47**, 1415-1425,
1013 doi:10.1038/ng.3437 (2015).
1014 38 Spracklen, C. N. *et al.* Identification and functional analysis of glycemic trait loci in
1015 the China Health and Nutrition Survey. *PLoS genetics* **14**, e1007275,
1016 doi:10.1371/journal.pgen.1007275 (2018).

1017 39 Varshney, A. *et al.* Genetic regulatory signatures underlying islet gene expression
1018 and type 2 diabetes. *Proceedings of the National Academy of Sciences of the United*
1019 *States of America* **114**, 2301-2306, doi:10.1073/pnas.1621192114 (2017).
1020 40 Kichaev, G. *et al.* Leveraging Polygenic Functional Enrichment to Improve GWAS
1021 Power. *American journal of human genetics* **104**, 65-75,
1022 doi:10.1016/j.ajhg.2018.11.008 (2019).
1023 41 Shriner, D. & Rotimi, C. N. Whole-Genome-Sequence-Based Haplotypes Reveal Single
1024 Origin of the Sickle Allele during the Holocene Wet Phase. *American journal of*
1025 *human genetics* **102**, 547-556, doi:10.1016/j.ajhg.2018.02.003 (2018).
1026 42 Kramer, H. J. *et al.* African Ancestry-Specific Alleles and Kidney Disease Risk in
1027 Hispanics/Latinos. *Journal of the American Society of Nephrology : JASN* **28**, 915-922,
1028 doi:10.1681/asn.2016030357 (2017).
1029 43 Ravenhall, M. *et al.* Novel genetic polymorphisms associated with severe malaria and
1030 under selective pressure in North-eastern Tanzania. *PLoS genetics* **14**, e1007172,
1031 doi:10.1371/journal.pgen.1007172 (2018).
1032 44 Hodonsky, C. J. *et al.* Genome-wide association study of red blood cell traits in
1033 Hispanics/Latinos: The Hispanic Community Health Study/Study of Latinos. *PLoS*
1034 *genetics* **13**, e1006760, doi:10.1371/journal.pgen.1006760 (2017).
1035 45 Gurdasani, D. *et al.* Uganda Genome Resource Enables Insights into Population
1036 History and Genomic Discovery in Africa. *Cell* **179**, 984-1002.e1036,
1037 doi:10.1016/j.cell.2019.10.004 (2019).
1038 46 Rees, M. G. *et al.* Cellular characterisation of the GCKR P446L variant associated with
1039 type 2 diabetes risk. *Diabetologia* **55**, 114-122, doi:10.1007/s00125-011-2348-5
1040 (2012).
1041 47 Bonomo, J. A. *et al.* The ras responsive transcription factor RREB1 is a novel
1042 candidate gene for type 2 diabetes associated end-stage kidney disease. *Human*
1043 *molecular genetics* **23**, 6441-6447, doi:10.1093/hmg/ddu362 (2014).
1044 48 Wessel, J. *et al.* Low-frequency and rare exome chip variants associate with fasting
1045 glucose and type 2 diabetes susceptibility. *Nature communications* **6**, 5897,
1046 doi:10.1038/ncomms6897 (2015).
1047 49 Scott, R. A. *et al.* A genomic approach to therapeutic target validation identifies a
1048 glucose-lowering GLP1R variant protective for coronary heart disease. *Science*
1049 *translational medicine* **8**, 341ra376, doi:10.1126/scitranslmed.aad3744 (2016).
1050 50 Nai, A. *et al.* TMPRSS6 rs855791 modulates hepcidin transcription in vitro and serum
1051 hepcidin levels in normal individuals. *Blood* **118**, 4459-4462, doi:10.1182/blood-
1052 2011-06-364034 (2011).
1053 51 Soranzo, N. *et al.* Common variants at 10 genomic loci influence hemoglobin A(1)(C)
1054 levels via glycemic and nonglycemic pathways. *Diabetes* **59**, 3229-3239,
1055 doi:10.2337/db10-0502 (2010).
1056 52 Sarnowski, C. *et al.* Impact of Rare and Common Genetic Variants on Diabetes
1057 Diagnosis by Hemoglobin A1c in Multi-Ancestry Cohorts: The Trans-Omics for
1058 Precision Medicine Program. *American journal of human genetics* **105**, 706-718,
1059 doi:10.1016/j.ajhg.2019.08.010 (2019).
1060 53 Kundaje, A. *et al.* Integrative analysis of 111 reference human epigenomes. *Nature*
1061 **518**, 317-330, doi:10.1038/nature14248 (2015).

1062 54 Nagel, M. *et al.* Meta-analysis of genome-wide association studies for neuroticism in
1063 449,484 individuals identifies novel genetic loci and pathways. *Nature genetics* **50**,
1064 920-927, doi:10.1038/s41588-018-0151-7 (2018).
1065 55 Savage, J. E. *et al.* Genome-wide association meta-analysis in 269,867 individuals
1066 identifies new genetic and functional links to intelligence. *Nature genetics* **50**, 912-
1067 919, doi:10.1038/s41588-018-0152-6 (2018).
1068 56 Schmidt, E. M. *et al.* GREGOR: evaluating global enrichment of trait-associated
1069 variants in epigenomic features using a systematic, data-driven approach.
1070 *Bioinformatics (Oxford, England)* **31**, 2601-2606, doi:10.1093/bioinformatics/btv201
1071 (2015).
1072 57 Parker, S. C. *et al.* Chromatin stretch enhancer states drive cell-specific gene
1073 regulation and harbor human disease risk variants. *Proceedings of the National*
1074 *Academy of Sciences of the United States of America* **110**, 17921-17926,
1075 doi:10.1073/pnas.1317023110 (2013).
1076 58 Pickrell, J. K. Joint analysis of functional genomic data and genome-wide association
1077 studies of 18 human traits. *American journal of human genetics* **94**, 559-573,
1078 doi:10.1016/j.ajhg.2014.03.004 (2014).
1079 59 Iotchkova, V. *et al.* GARFIELD classifies disease-relevant genomic features through
1080 integration of functional annotations with association signals. *Nature genetics* **51**,
1081 343-353, doi:10.1038/s41588-018-0322-6 (2019).
1082 60 van de Bunt, M. *et al.* Transcript Expression Data from Human Islets Links Regulatory
1083 Signals from Genome-Wide Association Studies for Type 2 Diabetes and Glycemic
1084 Traits to Their Downstream Effectors. *PLoS genetics* **11**, e1005694,
1085 doi:10.1371/journal.pgen.1005694 (2015).
1086 61 Civelek, M. *et al.* Genetic Regulation of Adipose Gene Expression and Cardio-
1087 Metabolic Traits. *American journal of human genetics* **100**, 428-443,
1088 doi:10.1016/j.ajhg.2017.01.027 (2017).
1089 62 Scott, L. J. *et al.* The genetic regulatory signature of type 2 diabetes in human
1090 skeletal muscle. *Nature communications* **7**, 11764, doi:10.1038/ncomms11764
1091 (2016).
1092 63 Ben Harouch, S., Klar, A. & Falik Zaccai, T. C. in *GeneReviews((R))* (eds M. P. Adam *et*
1093 *al.*) (University of Washington, Seattle
1094 University of Washington, Seattle. GeneReviews is a registered trademark of the University
1095 of Washington, Seattle. All rights reserved., 1993).
1096 64 Agus, D. B. *et al.* Vitamin C crosses the blood-brain barrier in the oxidized form
1097 through the glucose transporters. *The Journal of clinical investigation* **100**, 2842-
1098 2848, doi:10.1172/jci119832 (1997).
1099 65 Wolking, S. *et al.* Focal epilepsy in glucose transporter type 1 (Glut1) defects: case
1100 reports and a review of literature. *Journal of neurology* **261**, 1881-1886,
1101 doi:10.1007/s00415-014-7433-5 (2014).
1102 66 Guallar, D. *et al.* RNA-dependent chromatin targeting of TET2 for endogenous
1103 retrovirus control in pluripotent stem cells. *Nature genetics* **50**, 443-451,
1104 doi:10.1038/s41588-018-0060-9 (2018).
1105 67 Bian, F. *et al.* TET2 facilitates PPARgamma agonist-mediated gene regulation and
1106 insulin sensitization in adipocytes. *Metabolism: clinical and experimental* **89**, 39-47,
1107 doi:10.1016/j.metabol.2018.08.006 (2018).

1108 68 Yoo, Y. *et al.* TET-mediated hydroxymethylcytosine at the Ppargamma locus is
1109 required for initiation of adipogenic differentiation. *International journal of obesity*
1110 *(2005)* **41**, 652-659, doi:10.1038/ijo.2017.8 (2017).
1111 69 Lees, J. A. *et al.* Lipid transport by TMEM24 at ER-plasma membrane contacts
1112 regulates pulsatile insulin secretion. *Science (New York, N.Y.)* **355**,
1113 doi:10.1126/science.aah6171 (2017).
1114 70 Pottekat, A. *et al.* Insulin biosynthetic interaction network component, TMEM24,
1115 facilitates insulin reserve pool release. *Cell reports* **4**, 921-930,
1116 doi:10.1016/j.celrep.2013.07.050 (2013).
1117 71 Androulakis, II *et al.* Patients with apparently nonfunctioning adrenal incidentalomas
1118 may be at increased cardiovascular risk due to excessive cortisol secretion. *The*
1119 *Journal of clinical endocrinology and metabolism* **99**, 2754-2762,
1120 doi:10.1210/jc.2013-4064 (2014).
1121 72 Altieri, B. *et al.* Adrenocortical tumors and insulin resistance: What is the first step?
1122 *International journal of cancer* **138**, 2785-2794, doi:10.1002/ijc.29950 (2016).
1123 73 Johansson, M. *et al.* The influence of obesity-related factors in the etiology of renal
1124 cell carcinoma-A mendelian randomization study. *PLoS medicine* **16**, e1002724,
1125 doi:10.1371/journal.pmed.1002724 (2019).
1126 74 Diamanti-Kandarakis, E. & Dunaif, A. Insulin resistance and the polycystic ovary
1127 syndrome revisited: an update on mechanisms and implications. *Endocrine reviews*
1128 **33**, 981-1030, doi:10.1210/er.2011-1034 (2012).
1129 75 Morris, A. P. *et al.* Large-scale association analysis provides insights into the genetic
1130 architecture and pathophysiology of type 2 diabetes. *Nature genetics* **44**, 981-990,
1131 doi:10.1038/ng.2383 (2012).
1132 76 Leong, A. *et al.* Mendelian Randomization Analysis of Hemoglobin A(1c) as a Risk
1133 Factor for Coronary Artery Disease. *Diabetes care* **42**, 1202-1208, doi:10.2337/dc18-
1134 1712 (2019).
1135 77 Duncan, L. *et al.* Analysis of polygenic risk score usage and performance in diverse
1136 human populations. *Nature communications* **10**, 3328, doi:10.1038/s41467-019-
1137 11112-0 (2019).
1138 78 Mostafavi, H. *et al.* Variable prediction accuracy of polygenic scores within an
1139 ancestry group. *eLife* **9**, doi:10.7554/eLife.48376 (2020).
1140 79 Choi, S. W., Mak, T. S. & O'Reilly, P. F. Tutorial: a guide to performing polygenic risk
1141 score analyses. *Nature protocols* **15**, 2759-2772, doi:10.1038/s41596-020-0353-1
1142 (2020).
1143 80 D'Orazio, P. *et al.* Approved IFCC recommendation on reporting results for blood
1144 glucose (abbreviated). *Clinical chemistry* **51**, 1573-1576,
1145 doi:10.1373/clinchem.2005.051979 (2005).
1146 81 Voight, B. F. *et al.* The metabochip, a custom genotyping array for genetic studies of
1147 metabolic, cardiovascular, and anthropometric traits. *PLoS genetics* **8**, e1002793,
1148 doi:10.1371/journal.pgen.1002793 (2012).
1149 82 Abecasis, G. R. *et al.* An integrated map of genetic variation from 1,092 human
1150 genomes. *Nature* **491**, 56-65, doi:10.1038/nature11632 (2012).
1151 83 Li, Y., Willer, C. J., Ding, J., Scheet, P. & Abecasis, G. R. MaCH: using sequence and
1152 genotype data to estimate haplotypes and unobserved genotypes. *Genetic*
1153 *epidemiology* **34**, 816-834, doi:10.1002/gepi.20533 (2010).

1154 84 Pei, Y. F., Zhang, L., Li, J. & Deng, H. W. Analyses and comparison of imputation-
1155 based association methods. *PloS one* **5**, e10827, doi:10.1371/journal.pone.0010827
1156 (2010).
1157 85 Winkler, T. W. *et al.* Quality control and conduct of genome-wide association meta-
1158 analyses. *Nature protocols* **9**, 1192-1212, doi:10.1038/nprot.2014.071 (2014).
1159 86 Devlin, B. & Roeder, K. Genomic control for association studies. *Biometrics* **55**, 997-
1160 1004 (1999).
1161 87 Morris, A. P. Transethnic meta-analysis of genomewide association studies. *Genetic*
1162 *epidemiology* **35**, 809-822, doi:10.1002/gepi.20630 (2011).
1163 88 Bulik-Sullivan, B. K. *et al.* LD Score regression distinguishes confounding from
1164 polygenicity in genome-wide association studies. *Nature genetics* **47**, 291-295,
1165 doi:10.1038/ng.3211 (2015).
1166 89 Dastani, Z. *et al.* Novel loci for adiponectin levels and their influence on type 2
1167 diabetes and metabolic traits: a multi-ethnic meta-analysis of 45,891 individuals.
1168 *PLoS genetics* **8**, e1002607, doi:10.1371/journal.pgen.1002607 (2012).
1169 90 Benner, C. *et al.* FINEMAP: efficient variable selection using summary data from
1170 genome-wide association studies. *Bioinformatics (Oxford, England)* **32**, 1493-1501,
1171 doi:10.1093/bioinformatics/btw018 (2016).
1172 91 Astle, W. J. *et al.* The Allelic Landscape of Human Blood Cell Trait Variation and Links
1173 to Common Complex Disease. *Cell* **167**, 1415-1429.e1419,
1174 doi:10.1016/j.cell.2016.10.042 (2016).
1175 92 Canela-Xandri, O., Rawlik, K. & Tenesa, A. An atlas of genetic associations in UK
1176 Biobank. *Nature genetics* **50**, 1593-1599, doi:10.1038/s41588-018-0248-z (2018).
1177 93 Benyamin, B. *et al.* Novel loci affecting iron homeostasis and their effects in
1178 individuals at risk for hemochromatosis. *Nature communications* **5**, 4926,
1179 doi:10.1038/ncomms5926 (2014).
1180 94 Binesh, N. & Rezghi, M. Fuzzy clustering in community detection based on
1181 nonnegative matrix factoriztion with two novel evaluation criteria. *Applied Soft*
1182 *Computing* **69**, 689-703 (2018).
1183 95 Scott, R. A. *et al.* An Expanded Genome-Wide Association Study of Type 2 Diabetes in
1184 Europeans. *Diabetes* **66**, 2888-2902, doi:10.2337/db16-1253 (2017).
1185 96 Ernst, J. *et al.* Mapping and analysis of chromatin state dynamics in nine human cell
1186 types. *Nature* **473**, 43-49, doi:10.1038/nature09906 (2011).
1187 97 Mikkelsen, T. S. *et al.* Comparative epigenomic analysis of murine and human
1188 adipogenesis. *Cell* **143**, 156-169, doi:10.1016/j.cell.2010.09.006 (2010).
1189 98 Ernst, J. & Kellis, M. ChromHMM: automating chromatin-state discovery and
1190 characterization. *Nature methods* **9**, 215-216, doi:10.1038/nmeth.1906 (2012).
1191 99 Battle, A., Brown, C. D., Engelhardt, B. E. & Montgomery, S. B. Genetic effects on
1192 gene expression across human tissues. *Nature* **550**, 204-213,
1193 doi:10.1038/nature24277 (2017).
1194 100 Zhernakova, D. V. *et al.* Identification of context-dependent expression quantitative
1195 trait loci in whole blood. *Nature genetics* **49**, 139-145, doi:10.1038/ng.3737 (2017).
1196 101 Westra, H. J. *et al.* Systematic identification of trans eQTLs as putative drivers of
1197 known disease associations. *Nature genetics* **45**, 1238-1243, doi:10.1038/ng.2756
1198 (2013).

1199 102 Joehanes, R. *et al.* Integrated genome-wide analysis of expression quantitative trait
1200 loci aids interpretation of genomic association studies. *Genome biology* **18**, 16,
1201 doi:10.1186/s13059-016-1142-6 (2017).
1202 103 Pers, T. H. *et al.* Biological interpretation of genome-wide association studies using
1203 predicted gene functions. *Nature communications* **6**, 5890,
1204 doi:10.1038/ncomms6890 (2015).
1205
1206
1207 **Acknowledgments**
1208 The authors thank all investigators, staff members, and study participants for their contribution to
1209 all participating studies. The funders had no role in study design, data collection, analysis, decision to
1210 publish, or preparation of the manuscript. The authors received no specific funding for this work. A
1211 full list of funding and individual and study acknowledgments appears in the **Supplementary Note**.
1212
1213 **Author contributions**
1214 Project coordination: I.B.
1215 Writing group: J.C., C.N.S, G.M., A.V., L.J.C, S.C.J.P., K.L.M., C.L., E.W., A.P.M., I.B.
1216 Central analysis group: J.C., C.N.S, G.M., A.V., L.J.C, J.L., S.W., Y.W., X.Z., M.H., T.S.B., R.M., J.W., A.P.,
1217 R.L., K.H.K.C., J.Y., M.D.A, A.Y.C., A.C., J.H., S.H., M.A.K., T.L., W.M., H.M-M., A.N., S.C.N., K.N., C.K.R.,
1218 D.R., R.R., D.R., C.S., X.S., L.S., I.D.S., C.A.W., Y.W., P.W., W.Z., J.I.R., A.L.G., M.I.M., J.D., J.B.M., R.A.S.,
1219 I.P., A.L., CT.L., S.C.J.P., K.L.M., C.L., E.W., A.P.M., I.B.
1220 Cohort analysts: T.S.A., E.VR.A., L.F.B., J.A.B., N.P.B., C.P.C., B.E.C., J.C., X.C., L.C., C.C., B.H.C., K.C.,
1221 Y.C., H.G.d., G.E.D., A.D., Q.D., J.E., S.A.F., J.G., F.G., J.G., S.G., Y.H., F.P.H., J.H., Y.H., T.H., A.H., M.H.,
1222 R.A.J., T.K., K.A.K., Y.K., M.E.K., I.K.K., S.L., L.A.L., C.D.L., M.L., M.L., S.L., J.L., M.L., J.L., V.L., M.M.,
1223 C.M., M.E.M., A.N., M.N., D.N., R.N., G.P., M.P., L.R., L.J.R., S.S.R., N.R.R., R.R., K.R., S.S., R.S., K.E.S.,
1224 B.S., K.S., A.V.S., L.S., T.S., R.J.S., F.T., J.T., S.T., E.v., P.J.v., N.V., M.V., H.W., C.W., N.W., H.R.W.,
1225 W.W., T.W., A.W., A.R.W., T.X., M.Z., J.Z., W.Z.
1226 Cohort genotyping and phenotyping:
1227 N.A., Z.A., A.A., S.JL.B., D.B., M.B., R.N.B., A.B., M.B., L.L.B., S.R.B., D.W.B., Q.C., A.C., H.C., Y.C.,
1228 E.J.C.d., A.D., S.D., G.E., A.F., M.F., C.F., Y.G., A.P.G., A.G., S.H., C.A.H., C.H., A.A.H., C.H., W.A.H., S.I.,
1229 M.I., M.ArfanI., W.CraigJ., M.E.J., P.K.J., R.R.K., F.R.K., T.K., C.K., W.K., I.K., T.K., J.K., K.L., K.L., D.A.L.,
1230 N.R.L., R.N.L., H.L., S.L., J.L., A.L., J.L., C.L., T.M., F.M., G.M., S.M., S.M., T.N., G.N.N., J.L.N., M.N.,
1231 M.J.N., J.M.N., Y.O., A.P., P.A.P., O.P., Q.Q., D.R., D.F.R., A.R., F.R., K.R., I.R., C.S., K.S., N.S., A.S., J.S.,
1232 H.M.S., K.D.T., T.M.T., B.T., P.RHJ.T., E.T., M.Y.T., A.U., R.M.v., D.v., A.v., J.V.V., J.V., H.V., T.W., K.W.,
1233 T.Z.
1234 Cohort oversight and/or principal investigator:
1235 G.R.A., L.S.A., C.AlbertoA., M.E.A., P.A., L.A., D.M.B., L.J.B., S.B., H.B., C.B., M.B., E.B., B.O.B.,
1236 K.B., D.I.B., E.P.B., T.A.B., M.C., M.J.C., J.C.C., D.I.C., Y.C., C.C., F.S.C., A.C., F.C., H.d., G.D.,
1237 S.E., M.K.E., E.F., L.F., J.C.F., P.W.F., T.M.F., P.F., B.G., M.O.G., P.G., H.G., N.G., S.G., L.G.,
1238 V.G., X.G., A.H., T.H., C.H., S.R.H., B.L.H., W.H., E.I., P.S.J., M.J., J.B.J., J.WouterJ., P.K., R.K.,
1239 S.L.R.K., N.K., S.M.K., B.K., M.K., H.A.K., J.S.K., A.K., P.K., D.K., M.K., Z.K., M.L., T.A.L., L.J.L.,
1240 K.L., H.L., X.L., L.L., C.L., S.L., R.J.F.L., P.KE.M., A.M., A.M., D.O.M., T.A.M., P.B.M., I.N., J.R.O.,
1241 A.J.O., K.K.O., S.P., C.N.A.P., N.D.P., O.P., C.E.P., D.J.P., P.P.P., M.A.P., B.M.P., L.Q., L.J.R.,
1242 R.R., S.R., P.M.R., F.R.R., T.E.S., M.S., J.S.,N.S., P.S., L.J.S., E.S., P.S., X.S., P.ElineS., K.S.S.,
1243 B.H.S., H.S., T.S., T.I.A.S., T.D.S., A.S., C.J.S., M.S., L.S., Y.T., E.T., N.J.T., A.T., J.T., T.T., M.U.,
1244 P.v., C.v., , P.V., T.GM.V., L.E.W., M.W., Y.X.W., N.J.W., R.M.W., H.W., W.B.W., A.R.W., G.W.,
1245 J.F.W., T.W., J.W., A.H.X., L.R.Y., L.Y., M.Y., E.Z., W.Z., A.B.Z., J.I.R., A.L.G., M.I.M., J.D., J.B.M.,
1246 R.A.S., I.P., A.L., C.L., S.CJ.P., K.L.M., C.L., E.W., A.P.M., I.B.
1247
1248 **Competing interests statement**

A. Astrup is the recipient of honoraria as speaker for a wide range of Danish and international concerns and receives royalties from textbooks, and from popular diet and cookery books. A. Astrup is also co-inventor of a number of patents, including Methods of inducing weight loss, treating obesity and preventing weight gain (licensee Gelesis, USA) and Biomarkers for predicting degree of weight loss (licensee Nestec SA, CH), owned by the University of Copenhagen, in accordance with Danish law. I. Barroso and spouse own stock in GlaxoSmithKline and Incyte Corporation. B.H. Chen is now an employee of Life Epigenetics, Inc.; all work was completed prior to employment at Life Epigenetics. A.Y. Chu is now an employee of Merck & Co.; all work was completed prior to employment by Merck & Co. J.C. Florez has received consulting honoraria from Janssen. J. Gayan is now an employee of F. Hoffmann-La Roche Ltd, and owns stock of Roche and GlaxoSmithKline. A.L. Gloyn has received honoraria from Merck and Novo Nordisk. As of June 2019, ALG discloses that her spouse is an employee of Genentech and hold stock options in Roche. E. Ingelsson is now an employee of GSK; all work was completed prior to his employment by GSK. W. März has received grants and/or personal fees from the following companies/corporations: Siemens Healthineers, Aegerion Pharmaceuticals, AMGEN, Astrazeneca, Sanofi, Alexion Pharmaceuticals, BASF, Abbott Diagnostics Numares AG, Berlin-Chemie, Akzea Therapeutics, Bayer Vital GmbH , bestbion dx GmbH, Boehringer Ingelheim Pharma GmbH Co KG, Immundiagnostik GmbH, Merck Chemicals GmbH, MSD Sharp and Dohme GmbH, Novartis Pharma GmbH, Olink Proteomics, and Synlab Holding Deutschland GmbH. M.I. McCarthy has served on advisory panels for Pfizer, NovoNordisk, Zoe Global and received honoraria from Merck, Pfizer, NovoNordisk and Eli Lilly. He holds stock options in Zoe Global and has received research funding from Abbvie, Astra Zeneca, Boehringer Ingelheim, Eli Lilly, Janssen, Merck, NovoNordisk, Pfizer, Roche, Sanofi Aventis, Servier, Takeda. He is now an employee of Genentech and a holder of Roche stock. J.B. Meigs has consulted for Quest Diagnostics, Inc., who manufacturers of an HbA1c assay. M.E. Montasser has received grant funding from Regeneron Pharmaceutials. M.E. Montasser is also an inventor on a patent that was published by the United States Patent and Trademark Office on December 6, 2018 under Publication Number US 2018-0346888, and international patent application that was published on December 13, 2018 under Publication Number WO-2018/226560; all work was completed before these COI arose, and are unrelated to this work. D. Mook-Kanamori is a part-time clinical research consultant for Metabolon. J.L. Nadler is a member of the Scientific Advisory Board for Veralox Therapeutics Inc. C.N.A. Palmer has received research support from GlaxoSmithKline and AstraZeneca unrelated to this project. B.M. Psaty serves on the Steering Committee of the Yale Open Data Access Project funded by Johnson & Johnson. N. Sattar has consulted for Astrazeneca, Boehringer Ingelheim, Eli Lilly, Novo Nordisk, Napp and Sanofi and received grant support from Boehringer Ingelheim. R.A. Scott is an employee and shareholder of GlaxoSmithKline. T. Spector is the founder of Zoe Global Ltd. J. Tuomilehto receives research support from Bayer, is a consultant for Eli Lily, and holds stock in Orion Pharma and Aktivolabs Ltd.

Figure Legends

Figure 1 - Summary of all 242 loci identified in this study. 235 trans-ancestry loci are shown in orange (novel) or black (established) along with seven single-ancestry loci (blue) represented by nearest gene. Each locus is mapped to corresponding chromosome (outer segment). Each set of rows shows the results from the trans-ancestry analysis (orange) and each of the ancestries: European (purple), African American (tan), East Asian (grey), South Asian (green), Hispanic (yellow), sub-Saharan African (Ugandan-pink). Loci with a corresponding type 2 diabetes signal are represented by red circles in the middle of the plot.

Figure 2 – Trait variance explained by associated loci. The boxplots show the maximum, first quartile, median, third quartile and minimum of trait variance explained when using a genetic score with single-ancestry lead and index variants (EUR, AA, EAS, HISP and SAS) or a combination of individual trait trans-ancestry lead variants and single-ancestry lead and index variants (TA+EUR,

1300 TA+AA, TA+EAS, TA+HISP and TA+SAS). Variance explained for each trait (FG, FI and HbA1c) in each
1301 ancestry is shown on different panels and in different colors. Data points represent the variance
1302 explained in individual cohorts used in this analysis. R^2 was estimated in 1 to 11 cohorts with sample
1303 sizes ranging from 489 to 9,758 (**Supplementary Tables 8-11**).

1304 **Figure 3 – Transferability of PGS across ancestries.** For each trait, the barplots represent trait
1305 variance explained when using a European ancestry-derived PGS in European, East Asian and African
1306 American test datasets. Variance explained (the height of each bar) for each trait (FG, FI and HbA1c)
1307 in each ancestry is shown on different panels and in different colors.
1308
1309 **Figure 4 - Trans-ancestry fine-mapping.** A) Number of plausible causal variants at each locus-trait
1310 association derived from FINEMAP. B) Number of variants within each 99% credible set. Twenty-one
1311 locus-trait associations at 19 loci were mapped to a single variant in the 99% credible set. C) Fine-
1312 mapping resolution. For each of the 98 locus-trait associations with a predicted single causal variant
1313 in both trans-ancestry and single-ancestry analyses, the number of variants included in the 99%
1314 credible set in the single-ancestry fine-mapping (x axis; logarithmic scale) is plotted against those in
1315 the trans-ancestry fine-mapping (y axis; logarithmic scale). Trans-ancestry and single-ancestry fine-
1316 mapping were based on the same set of variants. After removing eight locus-trait associations with
1317 one variant in the 99% credible sets in both trans-ancestry and single-ancestry analyses, there were
1318 18 locus-trait associations (in grey) where trans-ancestry fine-mapping did not improve the
1319 resolution of fine-mapping results (i.e. number of variants in the 99% credible set did not decrease).
1320 Of the 72 locus-trait associations with improved trans-ancestry fine-mapping resolution (blue and
1321 red) further analyses in European fine-mapping emulating the total sample size in trans-ancestry
1322 fine-mapping demonstrated that 34 locus-trait associations (in red) were improved because of both
1323 total sample size and differences across ancestries, while 38 locus-trait associations (in blue) were
1324 only improved due to increased sample size in the original trans-ancestry fine-mapping analysis.

1325 **Figure 5 - Epigenomic landscape of trait-associated variants**. A: Enrichment of GWAS variants to
1326 overlap genomic regions including 'Static Annotations' which are common or 'static' across cell types
1327 and 'Stretch Enhancers' which are identified in each tissue/cell type. The numbers of signals for each
1328 trait are indicated in parentheses. Enrichment was calculated using GREGOR [56]. One-sided test for
1329 significance (red) is determined after Bonferroni correction to account for 59 total annotations
1330 tested for each trait; nominal significance (P<0.05) is indicated in yellow. B: Enrichment for HbA1c
1331 GWAS signals partitioned into "hard" Glycemic and Red Blood Cell cluster (signals from "hard"
1332 mature Red Blood Cell and reticulocyte clusters together) to overlap annotations including StrEs in
1333 Islets and the blood-derived leukemia cell line K562, respectively (additional partitioned results in
1334 **Supplementary Table 17**). C: Individual FI GWAS signals that drive enrichment in Adipose and
1335 Skeletal Muscle StrEs. D, E: Genome browser shots of FI GWAS signals – intronic region of the
1336 *COL4A2* gene (D) and an inter-genic region ~25kb from *LINC01214* gene (E) showing GWAS SNPs
1337 (lead and LD r^2>0.8 proxies), ATAC-seq signal tracks and chromatin state annotations in different
1338 tissues/cell types.

1339 **Figure 6 - Tissues and cell types significantly enriched for genes within glycemic-associated loci.**
1340 Top panel FG-associated loci, middle panel FI-associated loci, bottom panel Hba1c-associated loci.
1341 FDR thresholds are shown in red (q<0.05), orange (q<0.2), black (q≥0.2).

1342 **Figure 7 - Gene-set enrichment analyses.** Results from affinity-propagation clustering of significantly
1343 enriched gene-sets (FDR<0.05) identified by DEPICT for A) FG, B) FI, and C) HbA1c. Each node is a
1344 meta gene-set which is represented by an exemplar gene-set within the meta gene-set. For example,
1345 in B. "chronic myeloid leukemia " is an exemplar gene-set representing a much broader meta gene-
1346 set relating to cancer and represented in the zoomed in section on the right. Similarities between
1347 the meta gene-sets are represented by Pearson correlation coefficients (r>0.3). The nodes are

1348 colored according to the minimum gene-set enrichment p-value for gene-sets in that meta gene-set..
1349 PPI=protein-protein interaction network.
1350

1351 **Tables**
1352
1353 **Table 1 – Glossary of terms -** This study combined analyses of trait-associations across multiple correlated
1354 glycemic traits and across multiple ancestries, which has presented challenges in our ability to apply commonly
1355 used terms with clarity. For this reason, we define below terms often used in the field with variable meaning,
1356 as well as definitions of new terms used in this study.
1357

Term	Definition
EA (Effect allele)	The effect allele was that defined by METAL based on trans-ancestry FG results and aligned such that the same allele was kept as the effect allele across all ancestries and traits, irrespective of its allele frequency or effect size for that particular ancestry and trait, in this way the effect allele is not necessarily the trait-increasing allele.
Single-ancestry lead variant	Variant with the smallest p-value amongst all with $P < 5x10^{-8}$, within a 1Mb region, based on analysis of a single trait in a single ancestry.
Single-ancestry index variants	Variants identified by GCTA analysis of each autosome, and that appear to exert conditionally distinct effects on a given trait in a given ancestry ($P < 5x10^{-8}$). As defined, these include the single-ancestry lead variants.
Trans-ancestry lead variant	Variant identified by trans-ethnic meta-analysis of a given trait that has the strongest association for that trait ($log_{10}BF > 6$, which is broadly equivalent to $P < 5x10^{-8}$) within a 1Mb region.
Single-ancestry locus	1Mb region centred on a single-ancestry lead variant which does not contain a lead variant identified in the trans-ancestry meta-analysis (i.e., does not contain a trans-ancestry lead variant).
Signal	Conditionally independent association between a trait and a set of variants in LD with each other and which is noted by the corresponding index variant.
Trans-ancestry locus	A genomic interval that contains trans-ancestry trait-specific lead variants, with/out additional single-ancestry index variants, for one or more traits. This region is defined by starting at the telomere of each chromosome and selecting the first single-ancestry index variant or trans-ancestry lead variant for any trait. If other trans-ancestry lead variants or single-ancestry index variants mapped within 500kb of the first signal, then they were merged into the same locus. This process was repeated until there were no more signals within 500kb of the previous variant. A 500kb interval was added to the beginning of the first signal, and the end of the last signal to establish the final boundary of the trans-ancestry locus (**Extended Data Figure 2**). As defined, a trans-ancestry locus may not have a single lead trans-ancestry variant, but may instead contain multiple trans-ancestry lead variants, one for each trait.

1358
1359
1360
1361
1362
1363
1364
1365
1366

1367 **Online Methods**
1368 ***Study design and participants***
1369 This study included trait data from four glycemic traits: fasting glucose (FG), fasting insulin (FI), 2hr
1370 post-challenge glucose (2hGlu), and glycated hemoglobin (HbA1c). The total number of contributing
1371 cohorts ranged from 41 (2hGlu) to 131 (FG), and the maximum sample size for each trait ranged
1372 from 85,916 (2hGlu) to 281,416 (FG) (**Supplementary Table 1)**. Ancestry was initially defined at the
1373 cohort level, but within each cohort ancestry was confirmed with genetic data with ancestry outliers
1374 removed (**Supplementary Table 1**). Overall, European ancestry (EUR) participants dominated the
1375 sample size for all traits, representing between 68.0% (HbA1c) to 73.8% (2hGlu) of the overall
1376 sample size. African Americans (AA) represented between 1.7% (2hGlu) to 5.9% (FG) of participants;
1377 individuals of Hispanic ancestry (HISP) represented between 6.8% (FG) to 14.6% (2hGlu) of
1378 participants; individuals of East-Asian ancestry (EAS) represented between 9.9% (2hGlu) to 15.4%
1379 (HbA1c) of participants; and South-Asian ancestry (SAS) individuals represented between 0% (no
1380 contribution to 2hGlu) to 4.4% (HbA1c) of participants. Data from Ugandan participants were only
1381 available for the HbA1c analysis and represented 2% of participants.
1382
1383 ***Phenotypes***
1384 Analyses included data for FG and 2hGlu measured in mmol/l, FI measured in pmol/l, and HbA1c in
1385 % [where possible, studies reported HbA1c as a National Glycohemoglobin Standardization Program
1386 (NGSP) percent]. Similar to previous MAGIC efforts[7], individuals were excluded if they had type 1 or
1387 type 2 diabetes (defined by physician diagnosis); reported use of diabetes-relevant medication(s); or
1388 had a FG ≥7 mmol/L, 2hGlu ≥11.1mmol/L, or HbA1c ≥ 6.5%, as detailed in **Supplementary Table 1**.
1389 2hGlu measures were obtained 120 minutes after a glucose challenge in an oral glucose tolerance
1390 test (OGTT). Measures for FG and FI taken from whole blood were corrected to plasma level using
1391 the correction factor 1.13[80].
1392
1393 ***Genotyping, quality control, and imputation***
1394 Each participating cohort performed study-level quality control, imputation, and association
1395 analyses following a shared analysis plan. Cohorts were genotyped using commercially available
1396 genome-wide arrays or the Illumina CardioMetabochip (Metabochip) array (**Supplementary Table**
1397 **1)**[81]. Prior to imputation, each cohort performed stringent sample and variant quality control (QC) to
1398 ensure only high-quality variants were kept in the genotype scaffold for imputation. Sample quality
1399 control checks included removing samples with low call rate < 95%, extreme heterozygosity, sex
1400 mismatch with X chromosome variants, duplicates, first- or second-degree relatives (unless by
1401 design), or ancestry outliers. Following sample QC, cohorts applied variant QC thresholds for call rate
1402 (< 95%), Hardy-Weinberg Equilibrium (HWE) $P < 1x10^{-6}$, and minor allele frequency (MAF). Full
1403 details of QC thresholds and exclusions by participating cohort are available in **Supplementary Table**
1404 **1.**
1405
1406 Imputation was performed up to the 1000 Genomes Project phase 1 (v3) cosmopolitan reference
1407 panel[82], with a small number of cohorts imputing up to the 1000 Genomes phase 3 panel[19] or
1408 population-specific reference panels (**Supplementary Table 1).**
1409
1410 ***Study level association analyses***
1411 Each of the glycemic traits (FG, natural log FI, and 2hGlu) were regressed on BMI (except HbA1c),
1412 study-specific covariates, and principal components (unless implementing a linear mixed model).
1413 Analyses for FG, FI, and 2hGlu were adjusted for BMI as we had previously shown this did not
1414 materially affect results for FG and 2hGlu but improved our ability to detect FI-associated loci[15]. For
1415 simplicity, we refer to the traits as FG, FI and 2hGlu. For a discussion on collider bias see
1416 **Supplementary Note section 2c**. Both the raw and rank-based inverse normal transformed residuals
1417 from the regression were tested for association with genetic variants using SNPTEST[23] or

1418 Mach2Qtl[83,84]. Poorly imputed variants, defined as imputation $r^2 < 0.4$ or INFO score < 0.4, were
1419 excluded from downstream analyses (**Supplementary Table 1)**. Following study level QC,
1420 approximately 12,229,036 variants (GWAS cohorts) and 1,999,204 variants (Metabochip cohorts)
1421 were available for analysis (**Supplementary Table 1)**.
1422
1423 ***Centralized quality control***
1424 Each contributing cohort shared their summary statistic results with the central analysis group who
1425 performed additional QC using EasyQC[85]. Allele frequency estimates were compared to estimates
1426 from 1000Gp1 reference panel[82], and variants were excluded from downstream analyses if there
1427 was a minor allele frequency difference > 0.2 for AA, EUR, HISP, and EAS populations against AFR,
1428 EUR, MXL, and ASN populations from 1000 Genomes Phase 1, respectively, or a minor allele
1429 frequency difference > 0.4 for SAS against EUR populations. At this stage, additional variants were
1430 excluded from each cohort file if they met one of the following criteria: were tri-allelic; had a minor
1431 allele count (MAC) < 3; demonstrated a standard error of the effect size ≥ 10; or were missing an
1432 effect estimate, standard error, or imputation quality. All data that survived QC (approximately
1433 12,186,053 variants from GWAS cohorts and 1,998,657 variants from Metabochip cohorts) were
1434 available for downstream meta-analyses.
1435
1436 ***Single-ancestry meta-analyses***
1437 Single-ancestry meta-analyses were performed within each ancestry group using the fixed-effects
1438 inverse variance meta-analysis implemented in METAL[20]. We applied a double-genomic control (GC)
1439 correction[15,86] to both the study-specific GWAS results and the single-ancestry meta-analysis results.
1440 Study-specific Metabochip results were GC-corrected using 4,973 SNPs included on the Metabochip
1441 array for replication of associations with QT-interval, a phenotype not correlated with our glycemic
1442 traits[15].
1443
1444 ***Identification of single-ancestry index variants***
1445 To identify distinct association index variants across each chromosome within each ancestry (**Table**
1446 **1**), we performed approximate conditional analyses implemented in GCTA[21] using the --cojo-slct
1447 option (autosomes) and distance-based clumping (X chromosome). Linkage disequilibrium (LD)
1448 correlations for GCTA were estimated from a representative cohort from each ancestry: WGHS
1449 (EUR); CHNS (EAS); SINDI (SAS); BioMe (AA); SOL (HISP) and Uganda (for itself). The results from
1450 GCTA were comparable when using alternative cohorts for the LD reference. For any index variant
1451 with a QC flag which caused reason for concern, we performed manual inspection of forest plots to
1452 decide whether the signal was likely to be real (**Supplementary note**). Among 335 single-ancestry
1453 index variants across all traits, this manual inspection was done for 40 signals of which 32 passed
1454 and 8 failed after inspection. Thus, a total of 327 single-ancestry index variants passed and 8 failed.
1455
1456 ***Trans-ancestry meta-analyses***
1457 To leverage power across all ancestries, we also conducted trait-specific trans-ancestry meta-
1458 analysis by combining the single-ancestry meta-analysis results using MANTRA (**Supplementary**
1459 **note**)[87].We defined $\log_{10}$Bayes' Factor (BF) > 6 as genome-wide significant, approximately
1460 comparable to $P < 5 \times 10^{-8}$.
1461
1462 ***Manual curation of trans-ancestry lead variants***
1463 To ensure trans-ancestry lead variants were robust, we performed manual inspection of forest plots
1464 by at least two authors, for any variants with flags indicating possible QC issues (**Supplementary**
1465 **note**). Of 463 trans-ancestry lead variants across all traits, 184 passed without inspection, 131
1466 passed after inspection, and 148 failed after inspection.
1467
1468 ***Comparison of TA lead variants across ancestries***

1469 For each pair of ancestries, we calculated Pearson's correlation in EAFs for each trans-ancestry lead
1470 variant. The pairwise summarized heterogeneity of effect sizes between ancestries was then tested
1471 using the joint F-test of heterogeneity[32]. The test statistic is the sum of Cochran Q-statistics for
1472 heterogeneity across all trans-ancestry signals. Under the null hypothesis, the statistics follows the χ^2
1473 distribution with n degrees of freedom, where n is the number of the trans-ancestry lead variants.
1474
1475 ***LD-pruned variant lists***
1476 Several downstream analyses (for example, genomic feature enrichment, genetic scores, and
1477 estimation of variance explained by associated variants) require independent LD-pruned variants
1478 (r^2<0.1) to avoid double-counting variants which might otherwise be in LD with each other and that
1479 do not provide additional "independent" evidence. Therefore, for these analyses we generated
1480 different lists of either TA or single-ancestry LD pruned (r^2<0.1) variants, keeping in each case the
1481 variant with the strongest evidence of association (**Supplementary Table 7**). Subsequently, we
1482 combined TA and single-ancestry variant lists and conducted further LD pruning. For some analyses,
1483 we took the TA pruned variant list and added single-ancestry signals if the LD r^2<0.1, while for others
1484 we started with the single-ancestry pruned lists and supplemented with TA lead variants if the LD
1485 r^2<0.1. One exception was the list used for eQTL co-localizations, which included all single-ancestry
1486 European signals (without LD pruning) and supplemented with any additional TA lead variants
1487 (starting from the variants with the most significant P-values) in EUR LD r^2<0.1 with any of the
1488 variants already in list, and that reached at least $P<1\times10^{-5}$ in the European ancestry meta-analysis.
1489
1490 **Trait variance explained by associated loci**
1491 To determine how much of the phenotypic variance of each trait could be explained by the
1492 corresponding trait-associated loci, variants were combined in a series of weighted genetic scores
1493 (GS). The analysis was performed in a subset of the cohorts included in the discovery GWAS (with
1494 representation from each ancestry) and in a smaller number of independent cohorts (European
1495 ancestry only). Up to three different GS were derived per trait (and for each ancestry) in order to
1496 evaluate the potential for the trans-ancestry meta-GWAS identified loci to provide additional
1497 information above and beyond that contributed by the ancestry-specific meta-analysis results. These
1498 GS comprised: List A - single-ancestry signals; List B - single-ancestry signals plus trans-ancestry
1499 signals; and List C - trans-ancestry signals plus single-ancestry signals (**Supplementary Table 7**). In
1500 the case of the European ancestry cohorts that contributed to the GWAS, we employed the method
1501 of Nolte *et al.*[33] to adjust the effect sizes (betas) from the GWAS for the contribution of that cohort,
1502 providing sets of cohort-specific effect sizes that were then used to generate the GS. The association
1503 between each GS and its corresponding trait was tested by linear regression and the adjusted R^2
1504 from the model extracted as an estimate of the variance explained.
1505
1506
1507 ***Transferability of polygenic scores (PGS) across ancestries***
1508 We used the PRS-CSauto[34] software to first build European ancestry-derived PGS for each glycemic
1509 trait (FG, FI, 2hGlu, HbA1c) on the basis of summary statistics. However, PRS-CSauto does not
1510 perform well when the training dataset is relatively small and the genetic architecture is sparse[34].
1511 Consequently, 2hGlu was excluded from this analysis. For each trait, to obtain European ancestry
1512 training and test datasets, we first removed all cohorts only genotyped on the Metabochip which
1513 were not included in this analysis. From the remaining cohorts we then removed five of the largest
1514 European cohorts contributing to the respective European ancestry meta-analysis. For each trait,
1515 these five cohorts were meta-analyzed and used as the European ancestry test dataset.
1516 Subsequently, the remaining European ancestry cohorts were also meta-analyzed and used as the
1517 European ancestry training dataset. For each of the other ancestries, cohorts only genotyped on the
1518 Metabochip were also removed, and the remaining cohorts were meta-analyzed, and used as the
1519 non-European ancestry test datasets. Variants with MAF<0.05 or missing in over half of the

1520 individuals in the training dataset were removed[34,88]. The PGS for each trait was built using PRS-
1521 CSauto with default settings[34] with the effect size estimates based on the European training dataset
1522 being revised based on an LD reference panel matching the test dataset. The proportion of the trait
1523 variance explained by the European ancestry-derived PGS (R^2) was estimated using the R package
1524 "gtx"[89] based on the revised effect sizes and summary statistics from the test dataset for each
1525 ancestry.
1526
1527
1528 **Fine-mapping**
1529 Of the 242 loci identified in this study, 237 were autosomal loci which we took forward for fine-
1530 mapping (**Supplementary Table 2)**. We used the Bayesian fine-mapping method FINEMAP[90] (version
1531 1.1) to refine association signals and attempt to identify likely causal variants at each locus.
1532 FINEMAP estimates the maximum number of causal variants at each locus, calculates the posterior
1533 probability of each variant being causal, and proposes the most likely configuration of causal
1534 variants. The posterior probabilities of the configurations in each locus were used to construct 99%
1535 credible sets.
1536
1537 We performed both single-ancestry and trans-ancestry fine-mapping. In both analyses, only data
1538 from cohorts genotyped on GWAS arrays were used, and analyses were limited to trans-ancestry
1539 lead variants and other single-ancestry lead variants present in at least 90% of the samples for each
1540 trait. For the single-ancestry fine-mapping, FINEMAP estimates the number of causal variants in a
1541 region up to a maximum number, which we set to be two plus the number of distinct signals
1542 identified from the GCTA signal selection. FINEMAP uses single-ancestry and trait-specific z-scores
1543 from the fixed-effect meta-analysis in METAL[20] and an ancestry-specific LD reference, which we
1544 created from a subset of cohorts (combined sample size > 30% of the sample size for that ancestry),
1545 weighting each cohort by sample size. In the trans-ancestry fine-mapping, FINEMAP was similarly
1546 used to estimate the number of causal variants starting with two, and trait-specific z-scores and LD
1547 maps were generated from the sample size weighted average of those used in the single-ancestry
1548 fine-mapping. The maximum number of causal variants was iteratively increased by one until it was
1549 larger than the number of causal variants supported by data (Bayes factor), which was the estimated
1550 maximum number of causal variants used in the final run of fine-mapping analysis.
1551
1552 To compare fine-mapping results obtained from the single-ancestry and trans-ancestry efforts,
1553 analyses were limited to fine-mapping regions with evidence for a single likely causal variant in both,
1554 enabling a straightforward comparison of credible sets (**Supplementary note**). To ensure any
1555 difference in the fine-mapping results was not driven by different sets of variants being present in
1556 the different analyses, we repeated the single-ancestry fine-mapping limited to the same set of
1557 variants used in the trans-ancestry fine-mapping. The fine-mapping resolution was assessed based
1558 on comparisons of the 99% credible sets in terms of number of variants included in the set, and
1559 length of the region. To assess whether the improvement in the trans-ancestry fine-mapping was
1560 due to differences in LD, increased sample size, or both, we repeated the trans-ancestry fine-
1561 mapping mimicking the sample size present in the single-ancestry fine-mapping by dividing the
1562 standard errors by the square root of the sample size ratio and compared the results with those
1563 from the single-ancestry fine-mapping.
1564
1565 **Functional Annotation of trait-associated variants**
1566
1567 ***HbA1c signal classification***
1568 There were 218 HbA1c-associated signals from either the single-ancestry (i.e. all GCTA-signals from
1569 any ancestry) or trans-ancestry meta-analyses. To classify these signals in terms of their likely mode
1570 of action (i.e., glycemic, erythrocytic, or other [7]), we examined association summary statistics for the

1571 lead variants at the 218 signals in other large European datasets for 19 additional traits: three
1572 glycemic traits from this study (FG, 2hGlu and FI); seven mature red blood cell (RBC) traits[91,92] (red
1573 blood cell count, mean corpuscular volume, hematocrit, mean corpuscular hemoglobin, mean
1574 corpuscular hemoglobin concentration, hemoglobin concentration and red cell distribution width);
1575 five reticulocyte traits (reticulocyte count, reticulocyte fraction of red cells, immature fraction of
1576 reticulocytes, high light scatter reticulocyte count and high light scatter percentage of red cells)[91,92],
1577 and four iron traits (serum iron, transferrin, transferrin saturation and ferritin)[93]. Of the 218 HbA1c
1578 signals, data were available for the lead (n=183) or proxy (European LD $r^2 > 0.8$, n = 8) variants at 191
1579 signals.
1580
1581 The additional traits were clustered using hierarchical clustering to ensure biologically related traits
1582 would cluster together (**Supplementary note**). We then used a non-negative matrix factorization
1583 (NMF)[94] process to cluster the HbA1c signals. Each cluster was labelled as glycemic, reticulocyte,
1584 mature RBC, or iron related based on the strength of association of signals in the cluster to the
1585 glycemic, reticulocyte, mature RBC and iron traits (**Supplementary note**). To verify that our cluster
1586 naming was correct, we used HbA1c association results conditioned on either FG or iron traits, or
1587 type 2 diabetes association results (**Supplementary note**).
1588
1589 ***HbA1c genetic risk scores (GRSs) and T2D risk***
1590 We constructed GRS for each cluster of HbA1c-associated signals (based on hard clustering) and
1591 tested the association of each cluster with T2D risk using samples from the UK Biobank. Pairs of
1592 HbA1c signals in LD (EUR $r^2>0.10$) were LD pruned by removing the signal with the less significant *P*-
1593 value of association with HbA1c. The GRS for each cluster was calculated based on the logarithm of
1594 odds ratios from the latest T2D study summary statistics[95] and UK Biobank genotypes imputed to the
1595 Haplotype Reference Consortium[19]. From 487,409 UK Biobank samples (age between 46 and 82
1596 years, and 55% female), we excluded participants for the following reasons: 373 with mismatched
1597 sex; 9 not used in the kinship calculation; 78,365 non-European ancestry individuals; and 138,504
1598 with missing T2D status, age, or sex information. We further removed 26,896 related participants
1599 (kinship > 0.088, preferentially removing individuals with the largest number of relatives and
1600 controls where a T2D case was related to a control). T2D cases were defined by: (i) a history of
1601 diabetes without metformin or insulin treatment, (ii) self-reported diagnosis of T2D, or (iii) diagnosis
1602 of T2D in a national registry (N = 17,022, age between 47 and 79 years, and 36% female). Controls
1603 were participants without a history of T2D (N = 226,240, age between 46 and 82 years, and 56%
1604 female). We tested for association between each GRS and T2D using logistic regression including
1605 covariates for age, sex, and the first five principal components. Significance of association was
1606 evaluated by a bootstrap approach to incorporate the variance of each HbA1c associated signal in
1607 the T2D summary data. To do this, we generated the GRS of each cluster 200 times by resampling
1608 the logarithm of odds ratio of each signal with T2D. For each non-glycemic class that had a GRS
1609 significantly associated with T2D, we performed sensitivity analyses to evaluate whether the
1610 association was driven from variants that also belonged to a glycemic cluster when using a soft
1611 clustering approach (the signals were classified as also glycemic in the soft clustering or had an
1612 association $P \leq 0.05$ with any of the three glycemic traits).
1613
1614 ***Chromatin states***
1615 To identify genetic variants within association signals that overlapped predicted chromatin states,
1616 we used a previously published, 13 chromatin state model that included 31 diverse tissues, including
1617 pancreatic islets, skeletal muscle, adipose, and liver[39]. Briefly, this model was generated from
1618 cell/tissue ChIP-seq data for H3K27ac, H3K27me3, H3K36me3, H3K4me1, and H3K4me3, and input
1619 control from a diverse set of publicly available data[53,57,96,97] using the ChromHMM program[98]. As
1620 reported previously[39], StrEs were defined as contiguous enhancer chromatin state (Active Enhancer
1621 1 and 2, Genic Enhancer and Weak Enhancer) segments longer than 3kb[57].

1622 ***Enrichment of genetic variants in genomic features***
1623 We used GREGOR (version 1.2.1) to calculate the enrichment of GWAS variants overlapping static
1624 and StrEs[56]. For calculating the enrichment of glycemic trait-associated variants in these annotations,
1625 we used the filtered list of trait-associated variants as described above (**Supplementary Table 7**) as
1626 input. For calculating the enrichment of sub-classified HbA1c variants, we included the list of loci
1627 characterized as Glycemic, another list of loci characterized as Reticulocyte or mature Red Blood
1628 Cell, collectively representing the red blood cell fraction, along with lists of iron related or
1629 unclassified loci (**Supplementary Table 17**). We used the following parameters in GREGOR
1630 enrichment analyses: European r^2 threshold (for inclusion of variants in LD with the lead variant) =
1631 0.8, LD window size = 1 Mb, and minimum neighbour number = 500.
1632
1633 We used fGWAS (version 0.3.6)[58] to calculate enrichment of glycemic trait-associated variants in
1634 static and StrE annotations using summary level GWAS results. We used the default fGWAS
1635 parameters for enrichment analyses for individual annotations for each trait. For each annotation,
1636 the model provided the natural log of maximum likelihood estimate of the enrichment parameter.
1637 Annotations were considered as significantly enriched if the log2 (parameter estimate) and
1638 respective 95% confidence intervals were above zero or significantly depleted if the log2 (parameter
1639 estimate) and respective 95% confidence intervals were below zero.
1640
1641 We tested enrichment of trait-associated variants in static and StrE annotations with GARFIELD
1642 (v2)[59]. We formatted annotation overlap files as required by the tool; prepared input data at two
1643 GWAS thresholds - of $1x10^{-5}$ and a more stringent $1x10^{-8}$ by pruning and clumping with default
1644 parameters (garfield-prep-chr script). We calculated enrichment in each individual annotation using
1645 garfield-test.R with –c option set to 0. We also calculated the effective number of annotations using
1646 the garfield-Meff-Padj.R script. We used the effective number of annotations for each trait to obtain
1647 Bonferroni corrected significance thresholds for enrichment for each trait.
1648
1649 ***eQTL analyses***
1650 To aid in the identification of candidate casual genes at the European-only and trans-ancestry
1651 association signals, we examined whether any of the lead variants associated with glycemic traits
1652 (**Supplementary Table 7**) were also associated with expression level (FDR < 5%) of nearby transcripts
1653 located within 1 Mb in existing eQTL data sets of blood, subcutaneous adipose, visceral adipose,
1654 skeletal muscle, and pancreatic islet samples[60,61,99-102]. LD was estimated from the collected cohort
1655 pairwise LD information, where available, else from the European samples in 1000G phase 3. GWAS
1656 and eQTL signals likely co-localize when the GWAS variant and the variant most strongly associated
1657 with the expression level of the corresponding transcript (eSNP) exhibit high pairwise LD ($r^2 > 0.8$;
1658 1000 Genomes Phase 3, EUR). At these signals, we conducted reciprocal conditional analyses to test
1659 association between the GWAS variant and transcript level when the eSNP was also included in the
1660 model, and vice versa. We report GWAS and eQTL signals as co-localized if the association for the
1661 eSNP was not significant (FDR ≥ 5%) when conditioned on the GWAS variant; we also report signals
1662 from the eQTLGen whole blood meta-analysis data that meet only the LD threshold because
1663 conditional analysis was not possible.
1664
1665 ***Tissue and gene-set analysis***
1666 We performed enrichment analysis using DEPICT (Data-driven Expression-Prioritized Integration for
1667 Complex Traits) version 3, specifically developed for 1000 Genomes Project imputed meta-analysis
1668 data[103] to identify cell types and tissues in which genes at trait-associated variants were strongly
1669 expressed, and to detect enrichment of gene-sets or pathways. DEPICT data included human gene
1670 expression data for 19,987 genes in 10,968 reconstituted gene sets, and 209 tissues/cell types.
1671 Because gene expression data in DEPICT is based on European samples and LD, we selected trait-
1672 associated variants with $P<10^{-5}$ in the European meta-analysis and tested for enrichment of signals in

1673 each reconstituted gene-set, and each tissue or cell type. Enrichment results with a false discovery
1674 rate (FDR)<0.05 were considered significant. We ran DEPICT based on association results for all traits
1675 among: (i) cohorts with genome-wide data, or (ii) all cohorts (genome-wide and Metabochip
1676 cohorts). Because results were broadly consistent between the two approaches, we present results
1677 from the analysis that contained all cohorts as it had greater statistical power.
1678
1679 ***Statistics and reproducibility***
1680
1681 *Sample size*
1682 No statistical method was used to predetermine sample size. We aimed to bring together the largest
1683 possible sample size with GWAS data from individuals of diverse ancestries (European, Hispanic,
1684 African American, East Asian, South Asian and sub-Saharan African) without diabetes and with data
1685 for one or more of the following traits: fasting glucose, fasting insulin, 2hr post-challenge glucose,
1686 and glycated hemoglobin. The sample sizes were 281,416 (FG), 213,650 (FI), 215,977 (HbA1c) and
1687 85,916 (2hGlu) (**Supplementary Table 1**).Our sample size was sufficiently powered to detect
1688 common variant associations with each of the glycemic traits and was able to detect associations at
1689 242 loci.
1690
1691 *Randomization/ Blinding*
1692 This is a study of continuous traits therefore there were no experiments to randomize and there was
1693 no “outcome” to which investigators needed to be blinded to.
1694
1695 *Data exclusions*
1696 Prior to conducting this study, we identified reasons for which data should be excluded from the
1697 analysis at either the cohort or summary level; these exclusions are as follows. Sample quality
1698 control checks included removing samples with low call rate < 95%, extreme heterozygosity, sex
1699 mismatch with X chromosome variants, duplicates, first- or second-degree relatives (unless by
1700 design), or ancestry outliers. Following sample QC, cohorts applied variant QC thresholds for call rate
1701 (< 95%), Hardy-Weinberg Equilibrium (HWE) P < 1x10-6, and minor allele frequency (MAF). Full
1702 details of QC thresholds and exclusions by participating cohort are available in **Supplementary Table**
1703 **1.** Each contributing cohort shared their summary statistic results with the central analysis group
1704 who performed additional QC using EasyQC. Allele frequency estimates were compared to estimates
1705 from 1000Gp1 reference panel, and variants were excluded from downstream analyses if there was
1706 a minor allele frequency difference > 0.2 for AA, EUR, HISP, and EAS populations against AFR, EUR,
1707 MXL, and ASN populations from 1000 Genomes Phase 1, respectively, or a minor allele frequency
1708 difference > 0.4 for SAS against EUR populations. At this stage, additional variants were excluded
1709 from each cohort file if they met one of the following criteria: were tri-allelic; had a minor allele
1710 count (MAC) < 3; demonstrated a standard error of the effect size ≥ 10; imputation r2 < 0.4 or INFO
1711 score < 0.4; or were missing an effect estimate, standard error, or imputation quality.
1712
1713
1714 ***Data Availability***
1715 Ancestry-specific and overall meta-analysis summary level results are available through the MAGIC
1716 website (https://www.magicinvestigators.org/). Summary statistics are also available through the
1717 GWAS catalogue (https://www.ebi.ac.uk/gwas/) with the following accession codes: GCST90002225,
1718 GCST90002226, GCST90002227, GCST90002228, GCST90002229, GCST90002230, GCST90002231,
1719 GCST90002232, GCST90002233, GCST90002234, GCST90002235, GCST90002236, GCST90002237,
1720 GCST90002238, GCST90002239, GCST90002240, GCST90002241, GCST90002242, GCST90002243,
1721 GCST90002244, GCST90002245, GCST90002246, GCST90002247, and GCST90002248.
1722
1723

1724 ***Code availability***
1725 Source code implementing methods described in the paper are publicly available on
1726 https://zenodo.org/badge/latestdoi/346687844.
1727
1728 **References for Methods**
1729
1730 80 D'Orazio, P. *et al.* Approved IFCC recommendation on reporting results for blood
1731 glucose (abbreviated). *Clinical chemistry* **51**, 1573-1576,
1732 doi:10.1373/clinchem.2005.051979 (2005).
1733 81 Voight, B. F. *et al.* The metabochip, a custom genotyping array for genetic studies of
1734 metabolic, cardiovascular, and anthropometric traits. *PLoS genetics* **8**, e1002793,
1735 doi:10.1371/journal.pgen.1002793 (2012).
1736 82 Abecasis, G. R. *et al.* An integrated map of genetic variation from 1,092 human
1737 genomes. *Nature* **491**, 56-65, doi:10.1038/nature11632 (2012).
1738 83 Li, Y., Willer, C. J., Ding, J., Scheet, P. & Abecasis, G. R. MaCH: using sequence and
1739 genotype data to estimate haplotypes and unobserved genotypes. *Genetic*
1740 *epidemiology* **34**, 816-834, doi:10.1002/gepi.20533 (2010).
1741 84 Pei, Y. F., Zhang, L., Li, J. & Deng, H. W. Analyses and comparison of imputation-
1742 based association methods. *PloS one* **5**, e10827, doi:10.1371/journal.pone.0010827
1743 (2010).
1744 85 Winkler, T. W. *et al.* Quality control and conduct of genome-wide association meta-
1745 analyses. *Nature protocols* **9**, 1192-1212, doi:10.1038/nprot.2014.071 (2014).
1746 86 Devlin, B. & Roeder, K. Genomic control for association studies. *Biometrics* **55**, 997-
1747 1004 (1999).
1748 87 Morris, A. P. Transethnic meta-analysis of genomewide association studies. *Genetic*
1749 *epidemiology* **35**, 809-822, doi:10.1002/gepi.20630 (2011).
1750 88 Bulik-Sullivan, B. K. *et al.* LD Score regression distinguishes confounding from
1751 polygenicity in genome-wide association studies. *Nature genetics* **47**, 291-295,
1752 doi:10.1038/ng.3211 (2015).
1753 89 Dastani, Z. *et al.* Novel loci for adiponectin levels and their influence on type 2
1754 diabetes and metabolic traits: a multi-ethnic meta-analysis of 45,891 individuals.
1755 *PLoS genetics* **8**, e1002607, doi:10.1371/journal.pgen.1002607 (2012).
1756 90 Benner, C. *et al.* FINEMAP: efficient variable selection using summary data from
1757 genome-wide association studies. *Bioinformatics (Oxford, England)* **32**, 1493-1501,
1758 doi:10.1093/bioinformatics/btw018 (2016).
1759 91 Astle, W. J. *et al.* The Allelic Landscape of Human Blood Cell Trait Variation and Links
1760 to Common Complex Disease. *Cell* **167**, 1415-1429.e1419,
1761 doi:10.1016/j.cell.2016.10.042 (2016).
1762 92 Canela-Xandri, O., Rawlik, K. & Tenesa, A. An atlas of genetic associations in UK
1763 Biobank. *Nature genetics* **50**, 1593-1599, doi:10.1038/s41588-018-0248-z (2018).
1764 93 Benyamin, B. *et al.* Novel loci affecting iron homeostasis and their effects in
1765 individuals at risk for hemochromatosis. *Nature communications* **5**, 4926,
1766 doi:10.1038/ncomms5926 (2014).
1767 94 Binesh, N. & Rezghi, M. Fuzzy clustering in community detection based on
1768 nonnegative matrix factoriztion with two novel evaluation criteria. *Applied Soft*
1769 *Computing* **69**, 689-703 (2018).
1770 95 Scott, R. A. *et al.* An Expanded Genome-Wide Association Study of Type 2 Diabetes in
1771 Europeans. *Diabetes* **66**, 2888-2902, doi:10.2337/db16-1253 (2017).

96 Ernst, J. *et al.* Mapping and analysis of chromatin state dynamics in nine human cell types. *Nature* **473**, 43-49, doi:10.1038/nature09906 (2011).

97 Mikkelsen, T. S. *et al.* Comparative epigenomic analysis of murine and human adipogenesis. *Cell* **143**, 156-169, doi:10.1016/j.cell.2010.09.006 (2010).

98 Ernst, J. & Kellis, M. ChromHMM: automating chromatin-state discovery and characterization. *Nature methods* **9**, 215-216, doi:10.1038/nmeth.1906 (2012).

99 Battle, A., Brown, C. D., Engelhardt, B. E. & Montgomery, S. B. Genetic effects on gene expression across human tissues. *Nature* **550**, 204-213, doi:10.1038/nature24277 (2017).

100 Zhernakova, D. V. *et al.* Identification of context-dependent expression quantitative trait loci in whole blood. *Nature genetics* **49**, 139-145, doi:10.1038/ng.3737 (2017).

101 Westra, H. J. *et al.* Systematic identification of trans eQTLs as putative drivers of known disease associations. *Nature genetics* **45**, 1238-1243, doi:10.1038/ng.2756 (2013).

102 Joehanes, R. *et al.* Integrated genome-wide analysis of expression quantitative trait loci aids interpretation of genomic association studies. *Genome biology* **18**, 16, doi:10.1186/s13059-016-1142-6 (2017).

103 Pers, T. H. *et al.* Biological interpretation of genome-wide association studies using predicted gene functions. *Nature communications* **6**, 5890, doi:10.1038/ncomms6890 (2015).