Genome-Wide Meta-Analysis of 241,258 Adults Accounting for Smoking Behavior Identifies Novel Loci for Obesity Traits

Anne E. Justice*[1], Thomas W Winkler*[2], Mary F Feitosa*[3], Misa Graff*[1], Virginia A Fisher*[4], Kristin Young*[1], Llilda Barata*[3], Xuan Deng[4], Jacek Czajkowski[3], David Hadley[5,6], Julius S. Ngwa[7,4], Tarunveer S Ahluwalia[8], Audrey Y Chu[9,10], Nancy L. Heard-Costa[11,9], Elise Lim[4], Jeremiah Perez[4], John D. Eicher[12], Zoltán Kutalik[13-15], Luting Xue[4], Anubha Mahajan[16], Frida Renström[17,18], Joseph Wu[4], Qibin Qi[19], Shafqat Ahmad[18], Tamuno Alfred[20,21], Najaf Amin[22], Lawrence F Bielak[23], Amelie Bonnefond[24], Jennifer Bragg[25,26], Gemma Cadby[27], Martina Chittani[28], Scott Coggeshall[29], Tanguy Corre[14,15], Nese Direk[30,31], Joel Eriksson[32], Krista Fischer[33], Mathias Gorski[2,34], Marie Neergaard Harder[8], Momoko Horikoshi[16,35], Tao Huang[36,37], Jennifer E Huffman[12,38], Anne U Jackson[26], Johanne Marie Justesen[8], Stavroula Kanoni[39], Leena Kinnunen[40], Marcus E Kleber[41], Pirjo Komulainen[42], Meena Kumari[43,44], Unhee Lim[45], Jian'an Luan[46], Leo-Pekka Lyytikäinen[47,48], Massimo Mangino[49,50], Ani Manichaikul[51], Jonathan Marten[38], Rita PS Middelberg[52], Martina Müller-Nurasyid[53-55], Pau Navarro[38], Louis Pérusse[56,57], Natalia Pervjakova[33,58], Cinzia Sarti[59], Albert Vernon Smith[60,61], Jennifer A Smith[23], Alena Stančáková[62], Rona J Strawbridge[63,64], Heather M Stringham[26], Yun Ju Sung[65], Toshiko Tanaka[66], Alexander Teumer[67], Stella Trompet[68,69], Sander W van der Laan[70], Peter J van der Most[71], Jana V Van Vliet-Ostaptchouk[72], Sailaja L Vedantam[73,74], Niek Verweij[75], Jacqueline M Vink[76,77], Veronique Vitart[38], Ying Wu[78], Loic Yengo[24], Weihua Zhang[79,80], Jing Hua Zhao[46], Martina E Zimmermann[2], Niha Zubair[81], Gonçalo R Abecasis[26], Linda S Adair[82], Saima Afaq[79,80], Uzma Afzal[79,80], Stephan JL Bakker[83], Traci M Bartz[29,84], John Beilby[85-87], Richard N Bergman[88], Sven Bergmann[14,15], Reiner Biffar[89], John Blangero[90], Eric Boerwinkle[91], Lori L Bonnycastle[92], Erwin Bottinger[20,93], Daniele Braga[28], Brendan M Buckley[94], Steve Buyske[95,96], Harry Campbell[97], John C Chambers[80,79,98], Francis S Collins[92], Joanne E Curran[90], Gert J de Borst[99], Anton JM de Craen[+69], Eco JC de Geus[76,100], George Dedoussis[101], Graciela E Delgado[41], Hester M den Ruijter[70], Gudny Eiriksdottir[60], Anna Eriksson[32], Tõnu Esko[33], Jessica D Faul[102], Ian Ford[103], Terrence Forrester[104], Karl Gertow[63,64], Bruna Gigante[105], Nicola Glorioso[106], Jian Gong[81], Harald Grallert[107-109], Tanja B Grammer[41], Niels Grarup[8], Saskia Haitjema[70], Göran Hallmans[110], Anders Hamsten[63,64], Torben Hansen[8], Tamara B Harris[111], Catharina A Hartman[112], Maija Hassinen[42], Nicholas D Hastie[38], Andrew C Heath[113], Dena Hernandez[114], Lucia Hindorff[115], Lynne J Hocking[116,117], Mette Hollensted[8], Oddgeir L Holmen[118], Georg Homuth[119], Jouke Jan Hottenga[76], Jie Huang[120], Joseph Hung[121,122], Nina Hutri-Kähönen[123,124], Erik Ingelsson[125-127], Alan L James[85,128,121], John-Olov Jansson[129], Marjo-Riitta Jarvelin[130-133], Min A Jhun[23], Marit E Jørgensen[134], Markus Juonala[135,136], Mika Kähönen[137,138], Magnus Karlsson[139], Heikki A. Koistinen[40,140,141], Ivana Kolcic[142], Genovefa Kolovou[143], Charles Kooperberg[81], Bernhard K Krämer[41], Johanna Kuusisto[144], Kirsti Kvaløy[145], Timo A Lakka[146,42], Claudia Langenberg[46], Lenore J Launer[111], Karin Leander[105], Nanette R Lee[147,148], Lars Lind[149], Cecilia M Lindgren[150,16], Allan Linneberg[151-153], Stephane Lobbens[24], Marie Loh[79], Mattias Lorentzon[32], Robert Luben[154], Gitta Lubke[155], Anja Ludolph-Donislawski[53,156], Sara Lupoli[28], Pamela AF Madden[113], Reija Männikkö[42], Pedro Marques-Vidal[157], Nicholas G Martin[52], Colin A McKenzie[104], Barbara McKnight[29,84,158], Dan Mellström[32], Cristina Menni[49], Grant W Montgomery[159], AW (Bill) Musk[85,160,161], Narisu Narisu[92], Matthias Nauck[162], Ilja M Nolte[71], Albertine J Oldehinkel[112], Matthias Olden[2], Ken K Ong[46], Sandosh Padmanabhan[163,117], Patricia A Peyser[23], Charlotta Pisinger[164,165], David J Porteous[166,117], Olli T Raitakari[167,168], Tuomo Rankinen[169], DC Rao[65,113,170], Laura J Rasmussen-Torvik[171], Rajesh Rawal[107,108], Treva Rice[65,113], Paul M Ridker[10,172], Lynda M Rose[10], Stephanie Rosse[81], Igor Rudan[97], Serena Sanna[173], Mark A Sarzynski[169], Naveed Sattar[174], Kai Savonen[42], David Schlessinger[175], Salome Scholtens[71], Claudia Schurmann[20,21], Robert A Scott[46], Bengt Sennblad[63,64,176], Marten A Siemelink[70], Günther Silbernagel[177], P Eline Slagboom[178], Harold Snieder[71], Jan A Staessen[179,180], David J Stott[181], Morris A Swertz[182], Amy J Swift[92], Kent D Taylor[183,184], Bamidele O Tayo[185], Barbara Thorand[108,109], Dorothee Thuillier[24], Jaakko Tuomilehto[186], Andre G Uitterlinden[187,30], Liesbeth Vandenput[32], Marie-

Claude Vohl[57,188], Henry Völzke[67], Judith M Vonk[71], Gérard Waeber[157], Melanie Waldenberger[107,108], RGJ Westendorp[189], Sarah Wild[97], Gonneke Willemsen[76], Bruce HR Wolffenbuttel[72], Andrew Wong[190], Alan F Wright[38], Wei Zhao[23], M. Carola Zillikens[187], Damiano Baldassarre[191,192], Beverley Balkau[193], Stefania Bandinelli[194], Carsten A Böger[34], Dorret I Boomsma[76], Claude Bouchard[169], Marcel Bruinenberg[195], Daniel I Chasman[10,196], Yii-Der Ida Chen[197], Peter S Chines[92], Richard S Cooper[185], Francesco Cucca[173,198], Daniele Cusi[199], Ulf de Faire[105], Luigi Ferrucci[66], Paul W Franks[18,200,37], Philippe Froguel[24,201], Penny Gordon-Larsen[82,202], Hans-Jörgen Grabe[203,204], Vilmundur Gudnason[60,61], Christopher A Haiman[205], Caroline Hayward[38,117], Kristian Hveem[145], Andrew D. Johnson[12], J Wouter Jukema[68,206,207], Sharon LR Kardia[23], Mika Kivimaki[44], Jaspal S Kooner[208,80,98], Diana Kuh[190], Markku Laakso[144], Terho Lehtimäki[47,48], Loic Le Marchand[45], Winfried März[209,210], Mark I McCarthy[35,16,211], Andres Metspalu[33], Andrew P Morris[212,16], Claes Ohlsson[32], Lyle J Palmer[213], Gerard Pasterkamp[70,214], Oluf Pedersen[8], Annette Peters[108,109], Ulrike Peters[81], Ozren Polasek[142,97], Bruce M Psaty[215-217], Lu Qi[37], Rainer Rauramaa[42,218], Blair H Smith[219,117], Thorkild IA Sørensen[8,220,221], Konstantin Strauch[53,156], Henning Tiemeier[222], Elena Tremoli[191,192], Pim van der Harst[75,182,223], Henrik Vestergaard[8,134], Peter Vollenweider[157], Nicholas J Wareham[46], David R Weir[102], John B Whitfield[52], James F Wilson[224,38], Inês Barroso[225-227], Michael Boehnke[26], Panagiotis Deloukas[39,225,228], Caroline S Fox[9], Joel N Hirschhorn[73,229,74], David J Hunter[230,37,231,74], Tim D Spector[49], David P Strachan [5,232], Cornelia M van Duijn[22,233,234], Iris M Heid[2,235], Karen L Mohlke[78], Jonathan Marchini[236], Ruth JF Loos[†20,21,46,237,238], Tuomas O. Kilpeläinen[†8,46,239], Ching-Ti Liu[†4], Ingrid B Borecki[†3], Kari E North[†1], L. Adrienne Cupples[†4,9]

*These authors contributed equally to this work.

[†]These authors jointly supervised this work.

[+] Anton JM de Craen recently passed away while this work was in process.

Affiliations

1. Department of Epidemiology, University of North Carolina, Chapel Hill, NC 27599.
2. Department of Genetic Epidemiology, Institute of Epidemiology and Preventive Medicine, University of Regensburg, D-93053 Regensburg, Germany.
3. Division of Statistical Genomics, Department of Genetics, Washington University School of Medicine; St. Louis, MO, 63108 USA.
4. Department of Biostatistics, Boston University School of Public Health, Boston, MA 02118.
5. Population Health Research Institute, St. George's, University of London, London, SW17 0RE, UK.
6. TransMed Systems, Inc., Cupertino, CA 95014.
7. Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health, Baltimore MD.
8. The Novo Nordisk Foundation Center for Basic Metabolic Research, Section of Metabolic Genetics, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark.
9. NHLBI Framingham Heart Study, Framingham, MA, 01702 USA.
10. Division of Preventive Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA USA.
11. Department of Neurology, Boston University School of Medicine, Boston, MA, 02118, USA.
12. Population Sciences Branch, National Heart, Lung, and Blood Institute, National Institutes of Health, The Framingham Heart Study, Framingham, MA, USA.
13. Institute of Social and Preventive Medicine (IUMSP), Centre Hospitalier Universitaire Vaudois (CHUV), Lausanne, Switzerland.
14. Department of Medical Genetics, University of Lausanne, Lausanne, Switzerland.
15. Swiss instititute of Bioinformatics.
16. Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, OX3 7BN, UK.
17. Department of Biobank Research, Umeå University, Umeå, Sweden.
18. Department of Clinical Sciences, Genetic and Molecular Epidemiology Unit, Lund University, Skåne University Hospital Malmö, SE-205 02, Malmö, Sweden.
19. Department of Epidemiology and Population Health, Albert Einstein College of Medicine, Bronx, NY, USA;.
20. The Charles Bronfman Institute for Personalized Medicine, Icahn School of Medicine at Mount Sinai, New York, NY, USA.
21. The Genetics of Obesity and Related Metabolic Traits Program, Icahn School of Medicine at Mount Sinai, New York, NY, USA.
22. Genetic Epidemiology Unit, Department of Epidemiology, Erasmus University Medical Center, Rotterdam, 3015GE, The Netherlands.
23. Department of Epidemiology, School of Public Health, University of Michigan, Ann Arbor, MI, USA.
24. University of Lille, CNRS, Institut Pasteur of Lille, UMR 8199 - EGID, Lille, France.
25. Internal Medicine - Nephrology, University of Michigan, Ann Arbor, Michigan, USA.
26. Department of Biostatistics and Center for Statistical Genetics, University of Michigan, Ann Arbor, MI 48109, USA.
27. Centre for Genetic Origins of Health and Disease, University of Western Australia, Crawley, Australia.
28. Dept. Health Sciences, University of Milan, Via A. Di Rudiní, 8 20142, Milano, Italy.
29. Department of Biostatistics, University of Washington, Seattle, WA 98195.
30. Department of Epidemiology, Erasmus Medical Center, Rotterdam, Netherlands.
31. Department of Psychiatry, Dokuz Eylul University, Izmir, Turkey.

32. Centre for Bone and Arthritis Research, Department of Internal Medicine and Clinical Nutrition, University of Gothenburg, Sweden.
33. Estonian Genome Center, University of Tartu, Tartu 51010, Estonia.
34. Department of Nephrology, University Hospital Regensburg, Regensburg, Germany.
35. Oxford Centre for Diabetes, Endocrinology and Metabolism, University of Oxford, Churchill Hospital, Oxford, OX3 7LJ, UK.
36. Epidemiology Domain, Saw Swee Hock School of Public Health, National University of Singapore, Singapore 117549.
37. Department of Nutrition, Harvard School of Public Health, Boston, MA 02115, USA.
38. MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, Scotland.
39. William Harvey Research Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, UK.
40. Department of Health, National Institute for Health and Welfare, Helsinki, FI-00271 Finland.
41. Vth Department of Medicine, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany.
42. Kuopio Research Institute of Exercise Medicine, Kuopio, Finland.
43. ISER, University of Essex, Colchester, Essex, UK CO43SQ.
44. Department of Epidemiology and Public Health, UCL, London, UK. WC1E 6BT.
45. Epidemiology Program, University of Hawaii Cancer Center, Honolulu, HI 96813, USA.
46. MRC Epidemiology Unit, University of Cambridge School of Clinical Medicine, Institute of Metabolic Science, Cambridge, CB2 0QQ, UK.
47. Department of Clinical Chemistry, Fimlab Laboratories, Tampere 33520, Finland.
48. Department of Clinical Chemistry, University of Tampere School of Medicine, Tampere 33014, Finland.
49. Department of Twin Research and Genetic Epidemiology, King's College London, London, UK.
50. NIHR Biomedical Research Centre at Guy's and St. Thomas' Foundation Trust, London, UK.
51. Center for Public Health Genomics and Biostatistics Section, Department of Public Health Sciences, University of Virginia, Charlottesville, Virginia 22903.
52. Genetic Epidemiology, QIMR Berghofer Medical Research Institute, Brisbane, Australia.
53. Institute of Genetic Epidemiology, Helmholtz Zentrum München - German Research Center for Environmental Health, D-85764 Neuherberg, Germany.
54. Department of Medicine I, Ludwig-Maximilians-Universität, D-81377 Munich, Germany.
55. DZHK (German Centre for Cardiovascular Research), partner site Munich Heart Alliance, Munich, Germany.
56. Department of Kinesiology, Faculty of Medicine, Université Laval, Québec, Canada.
57. Institute of Nutrition and Functional Foods, Université Laval, Québec, Canada.
58. Department of Biotechnology, Institute of Molecular and Cell Biology, University of Tartu, Tartu 51010, Estonia.
59. Department of Social and Health Care, City of Helsinki, Helsinki, Finland.
60. Icelandic Heart Association, Kopavogur, Iceland.
61. Faculty of Medicine, University of Iceland, Reykjavik, Iceland.
62. Department of Medicine, Institute of Clinical Medicine, University of Eastern Finland, 70210 Kuopio, Finland.
63. Cardiovascular Medicine Unit, Department of Medicine Solna, Karolinska Institutet, Stockholm, Sweden.
64. Center for Molecular Medicine, Karolinska University Hospital Solna, Stockholm, Sweden.
65. Division of Biostatistics, Washington University School of Medicine, St Louis, MO.

66. Translational Gerontology Branch, National Institute on Aging, Baltimore MD, USA.
67. Institute for Community Medicine, University Medicine Greifswald, Germany.
68. Department of Cardiology, Leiden University Medical Center, The Netherlands.
69. Department of Gerontology and Geriatrics, Leiden University Medical Center, The Netherlands.
70. Laboratory of Experimental Cardiology, Department of Cardiology, Division Heart & Lungs, UMC Utrecht, the Netherlands.
71. Department of Epidemiology, University of Groningen, University Medical Center Groningen, The Netherlands.
72. Department of Endocrinology, University of Groningen, University Medical Center Groningen, the Netherlands.
73. Divisions of Endocrinology and Genetics and Center for Basic and Translational Obesity Research, Boston Children's Hospital, Boston MA 02115 USA.
74. Broad Institute of Harvard and MIT, Cambridge, MA 02142 USA.
75. Department of Cardiology, University Medical Center Groningen, University of Groningen, the Netherlands.
76. Department of Biological Psychology, Vrije Universiteit, Amsterdam, the Netherlands.
77. NCA, Vrije Universiteit & Vrije Universiteit Medical Center, Amsterdam, the Netherlands.
78. Department of Genetics, University of North Carolina, Chapel Hill, NC 27599 USA.
79. Dept Epidemiology and Biostatistics, School of Public Health, Imperical College London, UK.
80. Cardiology, Ealing Hospital NHS Trust, Middlesex, UK.
81. Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle WA USA.
82. Department of Nutrition, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill NC 27599.
83. Department of Medicine, University Medical Center Groningen, University of Groningen, Groningen, the Netherlands.
84. Cardiovascular Health Research Unit, Department of Medicine, University of Washington, Seattle, WA 98101.
85. Busselton Population Medical Research Institute, Nedlands, WA 6009, Australia.
86. PathWest Laboratory Medicine of WA, Sir Charles Gairdner Hospital, Nedlands, WA 6009, Australia.
87. School of Pathology and Laboraty Medicine, The University of Western Australia, 35 Stirling Hwy, Crawley, WA 6009, Australia.
88. Diabetes and Obesity Research Institute, Cedars-Sinai Medical Center, Los Angeles, California, USA.
89. Clinic for Prosthetic Dentistry, Gerostomatology and Material Science, University Medicine Greifswald, Germany.
90. South Texas Diabetes and Obesity Institute, University of Texas Rio Grande Valley, Brownsville, TX.
91. Human Genetics Center, The University of Texas Health Science Center, PO Box 20186, Houston, TC 77225.
92. Medical Genomics and Metabolic Genetics Branch, National Human Genome Research Institute, NIH, Bethesda, MD 20892, USA.
93. Department of Pharmacology and Systems Therapeutics, Icahn School of Medicine at Mount Sinai, New York, NY, USA.
94. Department of Pharmacology and Therapeutics, University College Cork, Ireland.
95. Department of Genetics, Rutgers University, Piscataway, NJ 08854, USA.
96. Department of Statistics and Biostatistics, Rutgers University, Piscataway, NJ 08854, USA;.

97. Usher Institute for Population Health Sciences and Informatics, The University of Edinburgh, Scotland, UK
98. Imperial College Healthcare NHS Trust, London, UK.
99. Department of Vascular Surgery, Division of Surgical Specialties, UMC Utrecht, the Netherlands.
100. EMGO+ Institute Vrije Universiteit & Vrije Universiteit Medical Center.
101. Department of Nutrition and Dietetics, School of Health Science and Education, Harokopio University, Athens, Greece.
102. Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor, MI, USA.
103. Robertson Center for Biostatistics, University of Glasgow, United Kingdom.
104. Tropical Metabolism Research Unit, Tropical Medicine Research Institute, University of the West Indies, Mona, JMAAW15 Jamaica.
105. Unit of Cardiovascular Epidemiology, Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden.
106. Hypertension and Related Disease Centre, AOU-University of Sassari.
107. Research Unit of Molecular Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, D-85764 Neuherberg, Germany.
108. Institute of Epidemiology II, Helmholtz Zentrum München - German Research Center for Environmental Health, D-85764 Neuherberg, Germany.
109. German Center for Diabetes Research, D-85764 Neuherberg, Germany.
110. Department of Public Health and Clinical Medicine, Section for Nutritional Research, Umeå University, Umeå, Sweden.
111. Laboratory of Epidemiology, Demography, and Biometry, National Institute on Aging, National Institutes of Health, Bethesda, MD.
112. Interdisciplinary Center Psychopathology and Emotion Regulation (ICPE), University of Groningen, University Medical Centre Groningen, Groningen, The Netherlands.
113. Department of Psychiatry, Washington University School of Medicine, St. Louis, MO.
114. Laboratory of Neurogenetics, National Institute on Aging, Bethesda, MD, USA.
115. Division of Genomic Medicine, National Human Genome Research Institute, National Institutes of Health, Bethesda, MD 20892.
116. Institute of Medical Sciences, University of Aberdeen, Foresterhill, Aberdeen, UK, AB25 2ZD.
117. Generation Scotland, Centre for Genomic and Experimental Medicine, University of Edinburgh, Edinburgh, Scotland.
118. St. Olav Hospital, Trondheim University Hospital, Trondheim, Norway.
119. Interfaculty Institute for Genetics and Functional Genomics, University Medicine Greifswald, Germany.
120. Department of Human Genetics, Wellcome Trust Sanger Institute, Hinxton, Cambridge, UK.
121. School of Medicine and Pharmacology, The University of Western Australia, 25 Stirling Hwy, Crawley, WA 6009, Australia.
122. Department of Cardiovascular Medicine, Sir Charles Gairdner Hospital, Nedlands, WA 6009, Australia.
123. Department of Pediatrics, Tampere University Hospital, Tampere 33521, Finland.
124. Department of Pediatrics, University of Tampere School of Medicine, Tampere 33014, Finland.
125. Department of Medical Sciences, Molecular Epidemiology, Uppsala University, Uppsala, 751 85, Sweden.
126. Department of Medicine, Division of Cardiovascular Medicine, Stanford University School of Medicine, Stanford, CA 94305, USA.
127. Science for Life Laboratory, Uppsala University, Uppsala, 750 85, Sweden.

128. Department of Pulmonary Physiology and Sleep Medicine, Sir Charles Gairdner Hospital, Nedlands, WA 6009, Australia.
129. Department of Physiology, Institute of Neuroscience and Physiology, the Sahlgrenska Academy at the University of Gothenburg, Gothenburg, Sweden.
130. Department of Epidemiology and Biostatistics, MRC–PHE Centre for Environment & Health, School of Public Health, Imperial College London, UK
131. Center for Life Course Epidemiology, Faculty of Medicine, P.O.Box 5000, FI-90014 University of Oulu, Finland
132. Biocenter Oulu, University of Oulu, Finland.
133. Unit of Primary Care, Oulu University Hospital, Kajaanintie 50, P.O.Box 20, FI-90220 Oulu, 90029 OYS, Finland.
134. Steno Diabetes Center, Gentofte, Denmark.
135. Department of Medicine, University of Turku, Turku 20520 Finland.
136. Division of Medicine, Turku University Hospital, Turku 20521, Finland.
137. Department of Clinical Physiology, Tampere University Hospital, Tampere 33521, Finland.
138. Department of Clinical Physiology, University of Tampere School of Medicine, Tampere 33014, Finland.
139. Clinical and Molecular Osteoporosis Research Unit, Department of Orthopedics and Clinical Sciences, Skåne University Hospital, Lund University, Malmö, Sweden.
140. Department of Medicine and Abdominal Center: Endocrinology, University of Helsinki and Helsinki University Central Hospital, Helsinki, FI-00029 Finland.
141. Minerva Foundation Institute for Medical Research, Biomedicum 2U, Helsinki, FI-00290 Finland.
142. Department of Public Health, Faculty of Medicine, University of Split, Croatia.
143. Department of Cardiology, Onassis Cardiac Surgery Center, Athens, Greece.
144. Department of Medicine, University of Eastern Finland and Kuopio University Hospital, 70210 Kuopio, Finland.
145. HUNT Research Centre, Department of Public Health and General Practice, Norwegian University of Science and Technology, 7600 Levanger, Norway.
146. Institute of Biomedicine/Physiology, University of Eastern Finland, Kuopio Campus, Finland.
147. USC-Office of Population Studies Foundation, Inc., University of San Carlos, Cebu City 6000, Philippines.
148. Department of Anthropology, Sociology and History, University of San Carlos, Cebu City 6000, Philippines.
149. Department of Medical Sciences, Cardiovascular Epidemiology, Uppsala University, Uppsala 751 85, Sweden.
150. Li Ka Shing Centre for Health Information and Discovery, The Big Data Institute, University of Oxford, Oxford OX3 7BN, UK.
151. Research Centre for Prevention and Health, the Capital Region of Denmark, Copenhagen, Denmark.
152. Department of Clinical Experimental Research, Rigshospitalet, Glostrup, Denmark.
153. Department of Clinical Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark.
154. Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK.
155. Department of Psychology, University of Notre Dame, Notre Dame, USA.
156. Institute of Medical Informatics, Biometry and Epidemiology, Chair of Genetic Epidemiology, Ludwig-Maximilians-Universität, D-81377 Munich, Germany.
157. Department of Medicine, Internal Medicine, Centre Hospitalier Universitaire Vaudois (CHUV), Lausanne, Switzerland.

158. Program in Biostatistics and Biomathematics, Fred Hutchinson Cancer Research Center, Seattle, WA 98109.
159. Molecular Epidemiology, QIMR Berghofer Medical Research Institute, Brisbane, Australia.
160. School of Population Health, The University of Western Australia, 35 Stirling Hwy, Crawley, WA 6009, Australia.
161. Department of Respiratory Medicine, Sir Charles Gairdner Hospital, Nedlands, WA 6009, Australia.
162. Institute of Clinical Chemistry and Laboratory Medicine, University Medicine Greifswald, Germany.
163. Institute of Cardiovascular and Medical Sciences, BHF Glasgow Cardiovascular Research Centre, University of Glasgow, Scotland.
164. Research Center for Prevention and Health, Glostrup Hospital, Glostrup Denmark.
165. Department of Public Health, Faculty of Health Sciences, University of Copenhagen, Denmark.
166. Centre for Genomic and Experimental Medicine, Institute of Genetics and Molecular Medicine, University of Edinburgh.
167. Department of Clinical Physiology and Nuclear Medicine, Turku University Hospital, Turku 20521, Finland.
168. Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, Turku 20520, Finland.
169. Human Genomics Laboratory, Pennington Biomedical Research Center, Baton Rouge, LA.
170. Department of Genetics, Washington University School of Medicine, St. Louis, MO.
171. Department of Preventive Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL USA.
172. Division of Cardiology, Brigham and Women's Hospital, Boston MA USA.
173. Istituto di Ricerca Genetica e Biomedica (IRGB), Consiglio Nazionale Delle Ricerche (CNR), Cittadella Universitaria di Monserrato, 09042, Monserrato, Italy.
174. BHF Glasgow Cardiovascular Research Centre, Faculty of Medicine, Glasgow, United Kingdom.
175. Laboratory of Genetics, National Institute on Aging, National Institutes of Health, Baltimore, MD, USA.
176. Science for Life Laboratory, Karolinska Institutet, Stockholm, Sweden.
177. Division of Angiology, Department of Internal Medicine, Medical University of Graz, Austria.
178. Department of Molecular Epidemiology, Leiden University Medical Center, Leiden , The Netherlands.
179. Research Unit Hypertension and Cardiovascular Epidemiology, Department of Cardiovascular Science , University of Leuven, Campus Sint Rafael, Kapucijnenvoer 35, Leuven; Belgium.
180. R&D VitaK Group, Maastricht University, Brains Unlimited Building, Oxfordlaan 55, Maastricht, The Netherlands.
181. Institute of Cardiovascular and Medical Sciences, Faculty of Medicine, University of Glasgow, United Kingdom.
182. Department of Genetics, University of Groningen, University Medical Center Groningen, the Netherlands.
183. Center for Translational Genomics and Population Sciences, Los Angeles Biomedical Research Institute at Harbor/UCLA Medical Center, Torrance, CA, USA.
184. Department of Pediatrics, University of California Los Angeles, Los Angeles, CA
185. Department of Public Health Sciences, Stritch School of Medicine, Loyola University of Chicago, Maywood, IL 61053 USA.
186. National Institute for Health and Welfare, Helsinki, Finland.
187. Department of Internal Medicine, Erasmus Medical Center, Rotterdam, Netherlands.

188. School of Nutrition, Université Laval, Québec, Canada.
189. Department of Public Health, and Center for Healthy Ageing, University of Copenhagen, Denmark.
190. MRC Unit for Lifelong Health and Ageing at UCL, 33 Bedford Place, London, WC1B 5JU, UK.
191. Dipartimento di Scienze Farmacologiche e Biomolecolari, Università di Milano, Milan , Italy.
192. Centro Cardiologico Monzino, IRCCS, Milan, Italy.
193. Inserm U-1018, CESP, 94807 Villejuif cedex France.
194. Geriatric Unit, Azienda USL Toscana centro, Florence Italy.
195. Lifelines Cohort Study, PO Box 30001, 9700 RB Groningen, the Netherlands.
196. Division of Genetics, Brigham and Women's Hospital, Boston MA USA.
197. Institute for Translational Genomics and Population Sciences, Los Angeles BioMedical Research Institute and Department of Pediatrics, Harbor-UCLA, Torrance, CA 90502, USA.
198. Dipartimento di Scienze Biomediche, Universita' degli Studi di Sassari, Sassari, Italy.
199. Sanipedia srl, Bresso (Milano), Italy and Institute of Biomedical Technologies National Centre of Research Segrate (Milano), Italy.
200. Department of Public Health & Clinical Medicine, Umeå University, Umeå, Sweden.
201. Department of Genomics of Common Disease, Imperial College London, London, UK.
202. Carolina Population Center, University of North Carolina at Chapel Hill, Chapel Hill NC 27516.
203. Department of Psychiatry and Psychotherapy, University Medicine Greifswald, Germany.
204. German Center for Neurodegenerative Diseases (DZNE), Site Rostock/ Greifswald, Germany.
205. Department of Preventive Medicine, Norris Comprehensive Cancer Center, Keck School of Medicine, University of Southern California, Los Angeles, CA 90089.
206. Durrer Center for Cardiogenetic Research, Amsterdam, The Netherlands.
207. Interuniversity Cardiology Institute of the Netherlands, Utrecht, The Netherlands.
208. Faculty of Med, National Heart & Lung Institute, Cardiovascular Science, Hammersmith Campus, Hammersmith Hospital, Hammersmith Campus, Imperial College London, UK.
209. Synlab Academy, Synlab Services GmbH, Mannheim, Germany.
210. Clinical Institute of Medical and Chemical Laboratory Diagnostics, Medical University of Graz, Graz, Austria.
211. Oxford National Institute for Health Research (NIHR) Biomedical Research Centre, Churchill Hospital, Oxford, UK.
212. Department of Biostatistics, University of Liverpool, Liverpool L69 3GL, UK.
213. School of Public Health, University of Adelaide, Adelaide, Australia.
214. Laboratory of Clinical Chemistry and Hematology, Division Laboratories & Pharmacy, UMC Utrecht, the Netherlands.
215. Department of Medicine, University of Washington, Seattle, WA 98195.
216. Department of Epidemiology, University of Washington, Seattle, WA 98101.
217. Group Health Research Institute, Group Health Cooperative, Seattle, WA 98101.
218. Department of Clinical Physiology and Nuclear Medicine, Kuopio University Hospital, Kuopio, Finland.
219. Division of Population Health Sciences,Mackenzie Building, Kirsty Semple Way, Ninewells Hospital and Medical School, Dundee, DD2 4RB.
220. Institute of Preventive Medicine, Bispebjerg and Frederiksberg Hospital (2000 Frederiksberg), The Capital Region, Copenhagen, Denmark.
221. MRC Integrative Epidemiology Unit, Bristol University, Bristol, UK.
222. Department of Psychiatry Erasmus Medical Center, Rotterdam, Netherlands.
223. Durrer Center for Cardiogenetic Research, ICIN-Netherlands Heart Institute, Utrecht, The Netherlands.

224. Usher Institute for Population Health Sciences and Informatics, The University of Edinburgh, Scotland, UK.
225. Wellcome Trust Sanger Institute, Hinxton, Cambridge, UK.
226. NIHR Cambridge Biomedical Research Centre, Level 4, Institute of Metabolic Science Box 289 Addenbrooke's Hospital Cambridge CB2 OQQ, UK.
227. University of Cambridge Metabolic Research Laboratories, Level 4, Institute of Metabolic Science Box 289 Addenbrooke's Hospital Cambridge CB2 OQQ, UK.
228. Princess Al-Jawhara Al-Brahim Centre of Excellence in Research of Hereditary Disorders (PACER-HD), King Abdulaziz University, Jeddah, Saudi Arabia.
229. Department of Genetics, Harvard Medical School, Boston MA 02115 USA.
230. Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA 02115 USA.
231. Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA 02115 USA.
232. Division of Population Health Sciences and Education, St George's, University of London, London, SW17 0RE, UK.
233. Netherlands Genomics Initiative (NGI)-sponsored Netherlands Consortium for Healthy Aging (NCHA).
234. Center for Medical Systems Biology, Leiden, The Netherlands.
235. Institute of Genetic Epidemiology, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg, 85764, Germany.
236. Department of Statistics, University of Oxford, Oxford, UK.
237. Mount Sinai School of Medicine, New York, NY 10029, USA.
238. The Mindich Child Health and Development Institute, Icahn School of Medicine at Mount Sinai, New York, NY, USA.
239. Department of Preventive Medicine, The Icahn School of Medicine at Mount Sinai, New York, NY, 10029, USA.

Author Contributions

LAC, KEN, IBB, TOK, TWW, RJFL and AEJ conceived and designed the study. LAC, RJFL, AEJ and TOK coordinated the collection of genome-wide association and interaction study results from the participating studies. The association and interaction results were contributed by SWvdL, MASi, SH, GJdB, HMDR and GP (AtheroExpress); AVS, TBH, GE, LJL, and VG (AGES study); KEN, MGr, AEJ, KY, EBoe, and PGL (ARIC study); JBW, NGM, RPSM, PAFM, ACH and GWM (AUSTWIN study); DPS (B58C study); GC, LJP, JoH, AWM, ALJ, and JBe (BHS study); CSch, TA, EBot and RJFL (BioMe); TT, DHe, LF (BLSA); BM, TMB, KDT, SC and BMP (CHS); YW, NRL, LSA and KLM (CLHNS study); ZK, PMV, TC, SBe, GWa, and PV (COLAUS study); JMart, IR, and CH (Croatia-Korcula study); VV, IK, and OPo (Croatia-Vis study); LY, AB, DT, SLo, BB and PF (DESIR study); RRau, TAL, PK, MHa, KSa, and RM (DR's EXTRA study); KF, NP, TE, and AMe (EGCUT study); JL, RAS, CL, and NJW (Ely study); CL, JL, RAS and NJW (EPIC), JHZ, RL, RAS, and NJW (EPIC-Norfolk study); NA, MCZ and CMvD (ERF study); IBB, MFF, JC and LB (Family Heart Study); JL, RAS, CL, RJFL and NJW (Fenland study); FX, JW, JSN, VAF, NLHC, CTL, CSF, and LAC (FramHS); MBo, FSC, KLM, and RNB (FUSION study); JT, LK, CSa, and HAK (FUSION2 study); MGo, BKK and CAB (Gendian); DJP, JEH, LJH, SP, CH and BHS (Generation Scotland); LFB, SLRK, MAJ, and PAP (GENOA study); SAh, FR, IB, GHa and PWF (GLACIER study); JE, CO, JOJ, MLor, AE and LV (GOOD study); TSA, THa and TIAS (GOYA study); BOT, CAM, SLV, TF, JNH and RSC (GxE); MHol, MNH, CP, AL and HVe (Health06 study); YJS, TRi, TRa, MASa, DCR and CB (HERITAGE Family Study); JASm, JDF, SLRK, WZhao, and DRW (HRS study); AUJ, KK, OLH, LLB, AJW, and KH (HUNT2 study); MC, DBr, SLu, NGl, JASt and DC (HYPERGENES); RJS, BS, KG, UdF, AH, ET and

DBa (IMPROVE); TT, DHe, and SBa (InCHIANTI study); JMJ, MEJ, NGr, and OPe (Inter99 study); TWW, IMH, MEZ, MMN, MO, ALD, HG, MW, RRaw, BT, AP and KSt (KORA S3 and S4 studies); JVvVO, JMV, SSch, MASw and BHRW (Lifelines); WZhang, MLo, UA, SAf, JCC, and JSK (LOLIPOP study); MEK, GED, TBG, GS, JiH, and WM (LURIC study); UL, CAH, LLeM, SBu, and LH (MEC study); AMan, LJRT and YdIC (MESA study); MLa, JK, AJS, HMS, PSC, and NN (METSIM study); MKa, DM and CO (MrOS); MHor, MRJ and MIM (NFBC66 study); LQ, THu, QQ and DJH (NHS study); DK, KKO, JL, and AW (NSHD study); JMV, GWi, GL, JJH, EJCdG and DIB (NTR study); PN, AFW, NDH, SW, HC and JFW (ORCADES study); AMah, CML, EI, LL and APM (PIVUS); NV, SJLB and PvdH (Prevend); ST, DJS, BMB, AJMdC, IF, RW, PES, NS and JWJ (PROSPER); LP, MCV, JEC, JBl and CB (QFS study); ND, MCZ, AGU and HT (RS1/RS2/RS3 study); JBr, SSa, DS, GRA, and FC (SardiNIA study); RJS, BS, BG, KL, AH and UdF (SCARFSHEEP); AT, RB, GHo, MN, HVö and HJG (SHIP study); BOT, CAM, SLV, TF, JNH and RSC (SPT); SK, GK, GD and PD (THISEAS); PJvdM, IMN, HS, AJO, CAH and MBr (TRAILS study); MMa, CM and TDS (TwinsUK study); AYC, LMR, PMR and DIC (WGHS study); NZ, SR, JG, CK and UP (WHI study); MKu, CL, JL, and MKi (Whitehall study); LPL, NHK, MJ, MKä, OTR and TL (YFS study). TWW, MGr, KY, JC, DHa, JSN, TSA, NLHC, FR, LX, QQ, JW and AEJ cleaned and quality checked the association and interaction results from the participating studies. TWW, KY, VAF, XD, JC, DHa, JSN, TSA, NLHC, LX and AEJ performed the meta-analyses. AYC, AEJ, LLB, MFF, TOK and LAC collected the supplementary information from the participating studies. AEJ, MGr, MFF, KY and VAF organized the supplementary tables. DHa, TWW, and AEJ provided look-up information from the GWAS meta-analysis of BMI, WAISTadjBMI and WHRadjBMI. JMarc provided lookup information from Smoking GWAS meta-analysis. AEJ performed the look-up in the NHGRI-EBI GWAS Catalog. MGr, XD, AEJ and ZK performed the analyses for variance explained by common variants in the SMK and nonSMK groups. MFF, KY, CTL, XD, LB and AEJ reviewed the literature for the identified loci. AEJ, KY, VAF and MG performed approximate conditional analyses. JDE and ADJ carried out the lookups for Expression Quantitative Trait loci. JP, EL and CTL conducted eQTL analyses in the Framingham Heart Study. AEJ, TWW, MFF, MGr, KY, VAF, XD, LB, JMarc, TOK, CTL, JSN, RJFL, KEN and LAC wrote the manuscript.

Corresponding Authors:

Anne E. Justice
Department of Epidemiology, Gillings School of Global Public Health
University of North Carolina at Chapel Hill
Chapel Hill, NC, 27599
USA.
anne.justice@unc.edu

L. Adrienne Cupples
Department of Biostatistics
Boston University School of Public Health
801 Massachusetts Avenue
Boston, MA 02118
USA
adrienne@bu.edu

ABSTRACT

Few genome-wide association studies (GWAS) account for environmental exposures, like tobacco smoking, with possibly large impacts on the overall trait variance when investigating the genetic contribution to obesity-related traits. We used GWAS data, including up to 51,080 current smokers and up to 190,178 nonsmokers (87% European descent) to identify novel loci influencing BMI and central adiposity, measured as waist circumference adjusted for BMI (WCadjBMI) and waist-to-hip ratio adjusted for BMI (WHRadjBMI). Across the three adiposity traits, we identified 23 novel genetic loci, and nine loci with convincing evidence of GxSMK interaction on obesity-related traits. Many of our newly identified loci highlight novel biological functions, including response to oxidative stress, addictive behavior, and regulatory functions, thus emphasizing the importance of accounting for environmental exposures in genetic analyses. Our results suggest that tobacco smoking may increase the genetic susceptibility to overall adiposity, but attenuate the genetic effects on body fat distribution.

Recent genome-wide association studies (GWAS) have described many loci implicated in obesity, body mass index (BMI), and central adiposity. However, most studies have ignored environmental exposures with possibly large impacts on the overall trait variance[1-8]. Variants that exert genetic effects on obesity susceptibility through interactions with environmental exposures often remain undiscovered due to heterogeneous main effects and stringent significance threshold. Thus, looking only for main effects might miss genetic variants that have specific effects in subgroups of the population, such as smokers[9].

It is often noted that currently-smoking individuals display both lower weight and BMI and higher waist circumference (WC) as compared to nonsmokers[10-15]. Individuals who smoke also have the smallest fluctuations in weight compared to those who have never smoked or have stopped smoking over approximately 20 years of follow-up[16-18]. There is also evidence that heavy smokers (>20 cigarettes per day [CPD]) and those that have smoked for more than 20 years are at greater risk for obesity than both non-smokers and light to moderate smokers (<20 CPD)[19-21]. Both men and women gain weight rapidly after cigarette smoking cessation[16-18,22-24], subsequently converging to the average weight fluctuation of nonsmokers within the first few years after cessation[16,23,25]. As such, many people, and especially females, intentionally use smoking as a weight management technique[26-28]. It remains unclear why smoking cessation leads to weight gain and why long-term smokers maintain weight throughout adulthood. It has been suggested that tobacco use suppresses appetite[25,29]. Evidence of smokers taking in a greater number of calories than nonsmokers appears to contradict this idea[25]. Another hypothesis is that cigarette smoking results in increased metabolic rate[25,29]. Identifying genes that influence adiposity and interact with smoking may help us clarify the pathways through which smoking influences weight and central adiposity[25].

Thus, a comprehensive study that evaluates smoking in conjunction with genetic contributions is warranted. Using GWAS data from the Genetic Investigation of Anthropometric Traits (GIANT) Consortium, we aim to identify genetic loci that influence obesity, as assessed by BMI, and central obesity independent of overall body size, as assessed by WC adjusted for BMI (WCadjBMI), and waist-to-hip ratio adjusted for BMI (WHRadjBMI). By accounting for current smoking status in our association analyses, we focus both on genetic variants observed through their main effects and their gene-smoking (GxSMK) interaction effects to gain a better understanding of their action on adiposity-related anthropometric traits.

RESULTS

We meta-analyzed study-specific association results from a total of 57 Hapmap-imputed GWAS and 22 studies that were genotyped with the Illumina Metabochip, including up to 241,258 (87% European descent) individuals (51,080 current smokers and 190,178 nonsmokers) for three adiposity traits (BMI, WCadjBMI, WHRadjBMI) while taking current smoking status into account (**Online Methods, Supplementary Figure 1, Supplementary Tables 1-4**). For primary analyses, we conducted meta-analyses across ancestries and sexes. For secondary analyses, we conducted meta-analyses in European-descent studies alone, and sex-specific meta-analyses (**Tables 1-4**, **Supplementary Tables 5-10**). Using the meta-analysis results, we considered four analytical approaches to evaluate the adjusted, joint and interaction effects of smoking on genetic associations with adiposity traits (**Figure 1, Online Methods**). Approach 1 (SNPadjSMK) examined genetic associations with obesity traits after adjusting for SMK. Approach 2 (SNPjoint) considered the joint impact of main effects adjusted for SMK + interaction effects[30]. Approaches 3 (SNPint)[31] and 4 (SNPscreen) were both aimed at identifying interaction effects. Approach 3 considered interaction effects alone; and Approach 4 followed up the loci identified in Approach 1 to identify additional interaction effects (SNPint). Results from Approaches 1-3 were

considered genome-wide significant (GWS) if a p value < $5x10^{-8}$ while Approach 4 used Bonferroni adjustment after screening. Lead variants >500 kb from previous associations with BMI, WCadjBMI, and WHRadjBMI were considered novel.

In total and across the three adiposity traits, we identified 23 novel associated genetic loci (6 for BMI, 11 for WCadjBMI, 6 for WHRadjBMI) and 10 loci with significant GxSMK interaction effects (2 for BMI, 2 for WCadjBMI, 5 for WHRadjBMI) (**Figure 1, Tables 1-4**, **Supplementary Tables 5-10**). We provide a comprehensive comparison with previously-identified loci[6,7] by trait in the supplementary material (**Supplementary Tables 11-13, Supplementary Note 1**).

Accounting for Current Smoking Status Identifies Novel Associations

For the primary meta-analysis of BMI (combined ancestries and combined sexes), a total of 58 loci reached GWS in Approach 1 (SNPadjSMK) (**Supplementary Table 5, Supplementary Figures 2-3**), including two novel loci near *SOX11*, and *SRRM1P2* (**Table 1**). Three more BMI loci were identified using Approach 2 (SNPjoint), including an additional novel locus near *CCDC93* (**Supplementary Figures 4-5**). For WCadjBMI, we identified 62 GWS loci for Approach 1 (SNPadjSMK) and two additional loci for Approach 2 (SNPjoint), including eight novel loci near *KIF1B*, *HDLBP*, *DOCK3*, *ADAMTS3*, *CDK6*, *GSDMC*, *TMEM38B*, and *ARFGEF2* (**Table 1**, **Supplementary Table 6**, **Supplementary Figures 2-5**). Lead variants at one locus near *PSMB10* identified in Approaches 1 and 2 (rs14178 and rs113090, respectively) are >500 kb from a previously-identified WCadjBMI-associated variant (rs16957304); however, after conditioning on the known variant, our signal is greatly attenuated ($P=3.02x10^{-2}$ and $P=5.22x10^{-3}$), indicating that this finding is not novel. For WHRadjBMI, a total of 32 loci were identified in Approach 1 (SNPadjSMK), including one novel locus near *HLA-C*, with no additional loci identified by accounting for interaction with SMK in Approach 2 (SNPjoint) (**Table 1**, **Supplementary Table 7, Supplementary Figures 2-5**).

For results from our primary meta-analyses, we used GCTA[32] to identify loci that harbor multiple independent SNPs associated with any of the outcomes (**Online Methods, Supplementary Tables 14-16**). Conditional analyses revealed no secondary signals within 500 kb of our novel lead SNPs for any of our three traits. Additionally, we performed conditional association analyses to determine if novel variants in the current analysis were independent of previously identified GWAS loci associated with related traits of interest. All BMI-associated SNPs were independent of previously-identified GWS associations with anthropometric and obesity-related traits within 500 kb. Six novel loci for WCadjBMI were near previous associations with related anthropometric traits. Of these six, association signals for rs6743226 near *HDLBP*, rs10269774 near *CDK6*, and rs6012558 near *ARFGEF2* were attenuated (P>1E-5 and β decreased by half) after conditioning on at least one nearby height and hip circumference adjusted for BMI (HIPadjBMI) SNP, but association signals remained independent of other related SNP-trait associations. For WHRadjBMI, conditioning on four other known SNPs associated with height, we found that our GWAS signal is attenuated by conditioning on two known height variants (rs6457374 and rs2247056), but remains significant for other conditional analyses.

Several additional loci were identified for Approaches 1 and 2 in at least one of the secondary meta-analyses that were not identified in any of the primary meta-analyses (**Table 2**, **Supplementary Tables 5-10**, **Supplementary Figure 6**). For BMI, 2 additional novel loci were identified by Approach 1, including 1 near *EPHA3* and 1 near *INADL*. For WCadjBMI, 2 additional novel loci were identified near *RAI14* and *PRNP*). For WHRadjBMI, five additional novel loci were identified in the secondary meta-analyses near *BBX, TRBI1*, *EHMT2*, *SMIM2* and *EYA4*. A comprehensive summary of the nearby genes for all novel loci and their potential biological relevance is provided in **Supplementary Note 2**.

Evidence for Modification of Genetic Predisposition to Obesity by Smoking Status

Using SMK-stratified meta-analysis results, Approach 3 directly evaluated the GxSMK interaction effects (**Table 3, Supplementary Tables 5-10**, **Figure 2**, **Supplementary Figures 7-8)**. For primary meta-analysis of BMI, two loci reached GWS for SNPint including a previously identified GxSMK interaction locus near *CHRNB4*[9], and a novel locus near *INPP4B*. Both loci exhibit GWS effects on BMI in smokers and no effects on BMI in nonsmokers. For the variant near *CHRNB4* (cholinergic nicotine receptor B4), the minor allele (G) exhibits a decreasing effect on BMI in current smokers (βsmk = - 0.047) but has no effect in nonsmokers (βnonsmk = 0.002). Previous studies have identified nearby SNPs in high LD associated with smoking (nonsynonymous, rs16969968 in *CHRNA5*)[9] and arterial calcification (rs3825807, a missense variant in *ADAMTS7)*[33]. Conditioning on these variants attenuated the interaction effect but did not eliminate it (**Supplementary Table 16**), suggesting a complex relationship between smoking, obesity, heart disease, and genetic variants in this region. Importantly, the *CHRNA5-CHRNA3-CHRNB4* gene cluster has been associated with lower BMI in current smokers[9,34], but with higher BMI in never smokers[9], which is supportive of the lack of association for BMI in nonsmokers in our study as well as a lack of previous GWAS findings on 15q25 for BMI (**Supplementary Table 17**)[6]. The *CHRNA5-CHRNA3-CHRNB4* genes encode the nicotinic acetylcholine receptor (nAChR) subunits α3, α5 and β4, which are expressed in the central nervous system[35,36]. Nicotine has differing effects on the body and brain, altering metabolism, caloric intake and feeding behaviors[37-39]. These findings suggest smoking exposure may modify genetic effects on 15q24-25 to influence smoking-related diseases through distinct pathways, such as obesity.

In our primary meta-analyses of WCadjBMI, one novel GWS locus (near *GRIN2A*) with opposite effect directions between smokers and nonsmokers was identified for Approach 3 (SNPint) (**Table 3, Supplementary Table 7, Figure 2, Supplementary Figures 7-8**). The T allele of rs4141488 exhibits a positive effect on WCadjBMI in current smokers as compared to a decreasing effect in nonsmokers (βsmk = 0.037, βnonsmk = -0.015). Additionally, in our secondary meta-analysis of European women-only, we identified an interaction between rs6076699, near *PRNP*, and SMK on WCadjBMI (**Table 4**, **Supplementary Table 9**, **Supplementary Figure 6**), a locus that was also identified in Approach 2 (SNPjoint) for European women. The major allele, A, has a positive effect on current smokers as compared to a weaker and negative effect on WC in nonsmokers (βsmk = 0.169, βnonsmk = -0.070). The opposing direction of effect in smokers compared to nonsmokers for these loci may explain why this variant remained undetected in previous GWAS of WCadjBMI (**Supplementary Table 17**).

We used an additional approach, Approach 4 (SNPscreen), (**Figure 1**, **Online Methods**) to evaluate GxSMK interactions whereby we screened SNPadjSMK results (from Approach 1) and then tested for interaction with a Bonferroni-correction in the subset of variants with GWS adjusted main effects (**Tables 3-4**, **Supplementary Tables 5-10**). While no loci were identified for WHRadjBMI using Approach 3, we identified significant interactions for WHRadjBMI with Approach 4. In primary meta-analyses, two SNPs, near *LYPLAL1* and *RSPO3*, were found to have significant interaction; both have previously published main effects on anthropometric traits[1]. These loci exhibit effects on WHRadjBMI in nonsmokers, but not in smokers (**Figure 2**). In our secondary meta-analyses, we identified three known loci with significant GxSMK interaction effects on WHRadjBMI using Approach 4 (SNPscreen) near *MAP3K1*, *HOXC4-HOXC6*, and *JUND* (**Table 2**, **Supplementary Tables 7** and **10**). Also, the variant, rs1809420, near the known *CHRNA5-CHRNA3-CHRNB4* gene cluster was identified using Approach 4 for BMI in the men-only, combined ancestries meta-analysis (**Supplementary Table 5**).

Power calculations demonstrate that Approach 4 has increased power to identify SNPs that show (i) an effect in one stratum (smokers or nonsmokers) and a less pronounced but concordant effect in the other stratum, or (ii) an effect in the larger nonsmoker stratum and no effect in smokers (**Figure 3**). In contrast, Approach 3 has increased power for SNPs that show (i) an effect in the smaller smoker stratum and no effect in nonsmokers, or (ii) an opposite effect between smokers and nonsmokers (**Figure 3**). Our findings for both approaches agree with the predictions from these power analyses, supporting the utility of employing both analytical approaches to identify GxSMK interactions.

Enrichment of Genetic Effects by Smoking Status on Obesity Traits

When examining the smoking specific effects for the BMI and WCadjBMI loci across our meta-analyses, no significant enrichment of genetic effects by smoking status were noted. (**Figure 2**, **Supplementary Figures 9-10**). Contrary to BMI and WCadjBMI, our results for WHRadjBMI were enriched for loci with a stronger effect in nonsmokers as compared to smokers, with 35 of 45 loci displaying numerically larger effects in nonsmokers ($P_{binomial}=1.2x10^{-4}$).

To further investigate whether the contribution of loci to adiposity outcomes differs by SMK, we calculated variance explained by subsets of SNPs selected on 15 increasing significance thresholds for Approach 1 from $P_{SNPadjSMK}=1x10^{-8}$ to $P_{SNPadjSMK}=0.1$ (**Supplementary Table 18**, **Figure 4**). Differences in variance explained between smokers and nonsmokers were significant (using a conservative P<0.003=0.05/15, Bonferroni-corrected for 15 thresholds) for BMI at each threshold, with more variance explained in smokers than in nonsmokers. So, while we did not find enrichment for loci with stronger effects in one stratum as compared to the other in our GWS loci, BMI loci explain more variation in smokers than in nonsmokers and this difference increases as the significance threshold increases. For WCadjBMI, the difference was significant for SNP sets beginning with $P_{SNPadjSMK}<3.16x10^{-4}$, and for WHRadjBMI at $P_{SNPadjSMK}<1x10^{-6}$, but in contrast to BMI, SNPs from Approach 1 explained a greater proportion of the variance in the nonsmokers than in current smokers. Differences in variance explained were greatest for BMI (differences ranged from 1.8% - 21% for smokers) and lowest for WHRadjBMI (ranging from 0.3% to 8.8% for nonsmokers).

These results suggest that tobacco smoking may increase the genetic susceptibility to overall adiposity, but attenuate the genetic effects on body fat distribution independent of overall adiposity. This contrast is not entirely surprising given phenotypic observations of higher overall adiposity and lower central adiposity in smokers[11,12,14-16]. Additionally, smoking increases oxidative stress and general inflammation in the body[40] and recent studies have found that oxidative stress may exacerbate weight gain, especially when compounded with increased calorie consumption[41,42]. Many genes implicated in BMI are involved in appetite regulation and feeding behavior[6]. However, for our waist traits, we have adjusted for overall adiposity and are likely highlighting other pathways through which smoking alters the genetic susceptibility to body fat distribution. Further investigations will be needed to determine whether or not these differences hold true for long-term and heavy smokers, as there is evidence that risk of obesity increases with the amount and duration of smoking[19-21,43]. Additionally, these results indicate that more loci remain to be discovered by adjusting for SMK as more variance in the trait can be explained as we drop the threshold for significant loci to be included.

Functional or Biological Role of Novel Obesity-related Loci

In order to gain insight into the biological relevance of our 23 newly identified loci, we conducted thorough searches of the literature and publicly available bioinformatics databases to understand the

functional role of all genes within 500 kb of our lead SNPs. Additionally, we systematically explored the potential role of our novel loci in affecting gene expression both with and without accounting for the influence of smoking behavior (**Online Methods**, **Supplementary Note 3, Supplementary Tables 19-21**).

Additional Support for Previously Implicated Biological Functions and Pathways

We found that the majority of the novel loci are near strong candidate genes with biological functions similar to previously identified adiposity-related loci, including regulation of body fat/weight, angiogenesis/adipogenesis, glucose and lipid homeostasis, general growth and development, etc. (**Supplementary Notes 2** and **3**).

For example, we identified rs17396340 associated with WCadjBMI for Approaches 1 and 2, an intronic variant in the *KIF1B* gene. This variant is associated with expression of *KIF1B* in whole blood with and without accounting for SMK[44] (GTeX and **Supplementary Tables 19** and **21**). Like other obesity-associated genes, *KIF1B* is highly expressed in the brain[45]. Knockout and mutant forms of *KIF1B* in mice have resulted in multiple brain abnormalities, including the morphology of the hippocampus[46], a region involved in memory and cognition, including food memory[47,48]. Additionally, the top SNP, rs17396340, is associated with the expression levels of *ARSA* in LCL tissue. Human adipocytes express functional *ARSA*, which turns dopamine sulfate into active dopamine. Dopamine is involved with long-term appetite regulation through the regulation of leptin and adiponectin levels, suggesting a regulatory role for *ARSA* in downstream regulation of appetite[49].

Additionally, a gene near our lead SNP for WHRadjBMI from Approach 1 for women-only, rs670752, may be of interest for follow-up studies. *CD47* (CD47 molecule) encodes a cell surface antigen involved in immune response to bacteria, cell adhesion, inflammatory response, and cell to cell signaling[50,51]. CD47 expression is significantly decreased in obese individuals and negatively correlated with BMI, WC, and Hip circumference[52]. Conversely, in mouse models, CD47 deficient mice show decreased weight gain on high fat diets, increased energy expenditure, improved glucose profile, and decreased inflammation[53].

Newly Implicated Biological Functions and Pathways

Several novel loci harbor genes involved in unique biological functions and pathways including those related to addictive behaviors and response to oxidative stress, highlighting the potential relevance for accounting for smoking in GWAS analyses. These potential candidate genes near our association signals are highly expressed in relevant tissues for regulation of adiposity and smoking behavior (e.g. brain, adipose tissue, liver, lung, muscle) (**Supplementary Note 2**, **Supplementary Table 19**).

For example, the *CHRNA5-CHRNA3-CHRNB4* cluster of genes are involved in the eNOS signaling pathway (Ingenuity KnowledgeBase, http://www.ingenuity.com) that is key for neutralizing reactive oxygen species introduced by tobacco smoke and obesity[54,55]. Disruption of this pathway has been associated with dysregulation of adiponectin in the adipocytes of obese mice, implicating this pathway in downstream effects on weight regulation[55,56]. This finding is especially important due to the compounded stress adiposity places on the body as it increases chronic oxidative stress itself[56]. While there is no direct relationship between *INPP4B* and obesity or oxidative stress, recent studies have implicated *INPP4B* in the regulation of the PI3K/Akt signaling pathway[57,58]. This pathway is important for a number of biological processes involved in cellular growth and proliferation, but also eNOS signaling, carbohydrate metabolism, and angiogenesis[59].

GRIN2A, near rs4141488 associated with WCadjBMI in Approach 3 (SNPint), assists in controlling long-term memory and learning through regulation and efficiency of synaptic transmission[60] and has been

associated with heroin addiction and obsessive-compulsive disorders[61-67]. Additionally, nicotine increases the expression of *GRIN2A* in the prefrontal cortex in murine models[68]. There are no established relationships between *GRIN2A* and obesity-related phenotypes in the literature, yet memantine and ketamine, pharmacological antagonists of GRIN2A activity[69-73], are implicated in the treatment for obesity-associated disorders, including binge-eating disorders and morbid obesity (ClinicalTrials.gov identifiers: NCT00330655, NCT02334059, NCT01997515, NCT01724983). Further, memantine is under clinical investigation for treatment of nicotine dependence (ClinicalTrials.gov identifiers: NCT01535040, NCT00136786, NCT00136747). While our lead SNP is not within a characterized gene, rs4141488 and variants in high LD (r^2>0.7) with our lead SNP are within active enhancer regions for several tissues, including liver, fetal leg muscle, smooth stomach and intestinal muscle, cortex, and several embryonic and pluripotent cell types (**Supplementary Note 2**), and therefore may represent an important regulatory region for nearby genes like *GRIN2A*.

In our secondary meta-analysis of European women-only, we identified a significant GxSMK interaction for rs6076699 on WCadjBMI (**Table 4**, **Supplementary Table 9**, **Supplementary Figure 6**). This SNP is 100kb upstream of the *PRNP* (prion protein) gene, which is a signaling transducer involved in multiple biological processes related to the nervous system, immune system, and many general cellular functions (**Supplementary Note 2**)[74]. Interestingly, alternate forms of the oligmers have been shown to form in response to oxidative stress caused by copper exposure[75]. Copper is present in cigarette smoke and elevated in serum of smokers, but is not considered outside of safe ranges according to the U.S. Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, and Office on Smoking and Health[76,77]. Another nearby gene to rs6076699, *SLC23A2* (Solute Carrier Family 23 [Ascorbic Acid Transporter], Member 2), is essential for the uptake and transport of Vitamin C, an important nutrient for DNA and cellular repair in response to oxidative stress both directly and through supporting the repair of Vitamin E after exposure to oxidative agents[78-80]. SLC23A2 is present in the adrenal glands and murine models indicate that it plays an important role in regulating the level of dopamine[81], a regulator of appetite[49]. Furthermore, this region is associated with success in smoking cessation and is implicated in addictive behaviors in general[82,83]. Our tag SNP is located within an active enhancer region (marked by open chromatin marks, DNAse hypersentivity, and transcription factor binding motifs); this regulatory activity appears tissue specific (sex-specific tissues and lungs) [HaploReg and UCSC Genome Browser].

Additionally, nicotinamide mononucleotide adenylyltransferease (*NMNAT1*), which is upstream of WCadjBMI variant rs17396340, is responsible for the synthesis of NAD from ATP and NMN[84-86]. NAD is necessary for cellular repair following oxidative stress. Upregulation of *NMNAT1* has been shown to protect against damage caused by reactive oxygen species in the brain, and specifically in the hippocampus[87,88]. Also for WCadjBMI variants, both *CDK6*, near our novel SNP rs10269774, and *FAM49B*, near our novel SNP rs6470765, are targets of the *BACH1* transcription factor, which is involved in cellular response to oxidative stress and management of the cell cycle[89].

Influence of Novel Loci on Related Traits

In order to determine if any of the 26 novel SNPs identified are associated with smoking behavior, we conducted a look-up in existing GWAS of smoking behavior[90]. **Supplementary Table 17** shows the meta-analysis p-values of our novel SNPs for the Ever/Never, Current/Non-Current and Smoking Quantity (SQ) phenotypes. Eight SNPs across the three adiposity traits show evidence of association at nominal significance for at least one of the smoking traits. However, after multiple test correction (P<0.05/26=0.0019), only one SNP remains significant: rs12902602, identified for Approaches 2

807 (SNPjoint) and 3 (SNPint) for BMI, showed association with SQ ($P=1.45x10^{-9}$), and is in LD with a known
808 region around the nicotine receptor cluster.
809
810 We conducted a search in the NHGRI-EBI GWAS Catalog[91,92] to determine if any of our newly identified
811 loci are in high LD with variants associated with related cardiometabolic and behavioral traits or
812 diseases. Of the seven novel BMI SNPs, only rs12902602 was in high LD ($r^2>0.7$) with SNPs previously
813 associated with smoking-related traits (e.g. nicotine dependence), lung cancer, and cardiovascular
814 diseases (e.g. coronary heart disease) (**Supplementary Table 22**). Of the 12 novel WCadjBMI SNPs, five
815 were in high LD with previously-reported GWAS variants for mean platelet volume, height, infant length,
816 and melanoma. Of the six novel WHRadjBMI SNPs, three were near several previously associated
817 variants, including cardiometabolic traits (e.g. LDL cholesterol, triglycerides, and measures of renal
818 function).
819
820 Lastly, as smoking has a negative (weight decreasing) effect on BMI, it is likely that smoking associated
821 genetic variants have an effect on BMI in current smokers. Therefore, it is expected that smoking
822 associated SNPs exhibit some interaction with smoking on BMI. To examine this further, we looked up
823 all GWAS-identified smoking behavior SNPs ($P<5x10^{-8}$) listed in the NHGRI-EBI GWAS Catalog[91,92],
824 including 10 variants in 6 loci. Two of these loci reached nominal significance ($P<0.05$) for GxSMK
825 interaction on BMI (**Supplementary Table 23**), but only one locus reached Bonferroni-corrected
826 significance ($P<0.005$). None of the smoking-associated SNPs exhibited GxSMK interaction. Therefore,
827 we did not see a strong enrichment for low interaction P-values among previously identified smoking
828 loci.
829
830 **Challenges with Accounting for Environmental Exposures in GWAS**
831
832 A possible limitation may be the definition and harmonization of the smoking status. We chose to
833 stratify based only on current smoking status without consideration of type of smoking (e.g. cigarette,
834 pipe) for two reasons. First, focusing on weight alone, former smokers tend to return to their expected
835 weight quickly following smoking cessation[16,23,25]. Second, this definition allowed us to maximize sample
836 size, as many of participating studies only had smoking status at time of data collection available. We
837 asked studies that had the information available to exclude participants that had only recently quit
838 smoking in an attempt to decrease trait heterogeneity. However, WC and WHR may not behave in the
839 same manner as weight and BMI with former smokers retaining excess fat around their waist[24]. Given
840 the difference between change in waist circumference and weight for smokers, it is possible that results
841 may differ slightly if we were able to harmonize the variable in a different manner.
842
843 Another limitation may be the potential for bias in our effect estimates when adjusting for a correlated
844 covariate in our association models (e.g. collider bias)[93,94]. This phenomenon is of particular concern
845 when the correlation between the outcome and the covariate is high and when significant genetic
846 associations occur with both traits in opposite directions. In other words, when the correlation between
847 the outcome and the covariate is positive, but the observed genetic effect directions are opposite. In
848 this paper, we have adjusted for BMI for both the WCadjBMI and WHRadjBMI traits. WHR has a
849 correlation of 0.49 with BMI, while WC has a correlation of 0.85[1,94]. Using previously published results
850 for BMI, WCadjBMI and WHRadjBMI, we find three of our novel loci for WCadjBMI (near *DOCK3,*
851 *ARFGEF2, TMEM38B*) and at two of our novel loci for WHRadjBMI (near *EHMT2, HLA-C*) (**Supplementary**
852 **Table 17**) with nominally significant associations with BMI and with opposite direction of effect. At these
853 loci, the genetic effect estimates should be interpreted with caution. Additionally, we have adjusted for
854 SMK in Approach 1 (SNPadjSMK). However binary smoking status that was used in the current analysis

has a very low correlation to BMI, WC, and WHR, as estimated in the ARIC study's European descent participants (-0.13, 0.08, and 0.12 respectively) and in the Framingham Heart Study (-0.05, 0.08, 0.16). Additionally, there are no loci identified in Approach 1 (SNPadjSMK) that are associated with any smoking behavior trait and that exhibit an opposite direction of effect from that identified in our adiposity traits (**Supplementary Table 17**). We therefore preclude potential bias and postulate true gain in power through SMK-adjustment at these loci.

In order to assess how much additional information is provided by accounting for SMK and GxSMK in GWAS for obesity traits, we calculated and compared genetic risk scores (GRSs) based on various subsets of lead SNP genotypes and based on various regression models (**Online Methods**). While any GRS was associated with its obesity trait ($P<1.6 \times 10^{-7}$, **Supplementary Table 24**), adding SMK and GxSMK terms to the regression model along with including novel variants in the calculation of the GRSs substantially increased variance explained. For example, variance explained increased by 38% for BMI (from 1.53% to 2.11%, $P=4.3\times10^{-5}$), by 27% for WCadjBMI (from 2.59% to 3.29%, $P=3.9\times10^{-6}$) and by 168% for WHRadjBMI (from 0.82% to 2.20%, $P=3.2\times10^{-11}$). Therefore, despite potential limitations, there is much to be gained by accounting for environmental exposures in GWAS studies.

DISCUSSION

To help us understand the effects of smoking on the genetic susceptibility to obesity, assessed by BMI, WCadjBMI, and WHRadjBMI, we conducted a series of meta-analyses to uncover genetic variants that may be masked when the environmental influence of smoking is not considered, and to discover genetic loci that interact with smoking to influence obesity-related traits. We identified 161 loci in total, including 23 novel loci influencing adiposity traits (6 for BMI, 11 for WCadjBMI, and 6 for WHRadjBMI). While many of our newly identified loci support the hypothesis that smoking may influence weight fluctuations through appetite regulation, these novel loci also have highlighted new biological processes and pathways implicated in the pathogenesis of obesity.

Importantly, we identified nine loci with convincing evidence of GxSMK interaction on obesity-related traits. We were able to replicate the previous GxSMK interaction with BMI finding within the *CHRNA5-CHRNA3-CHRNB4* gene cluster. Further, one novel BMI-associated locus near *INPP4B* and two novel WCadjBMI-associated loci near *GRIN2A* and *PRNP* displayed significant GxSMK interaction. By using robust analytical methods, we were able to identify significant GxSMK interaction for one known BMI-associated locus near *ADAMTS7* as well as for five known WHRadjBMI-associated loci near *LYPLAL1*, *RSPO3*, *MAP3K1*, *HOXC4-HOXC6* and *JUND*. Our results show that the majority of these loci harbor strong candidate genes with evidence for influence over adiposity, and also evidence that points to a possible role for the modulation of effects through tobacco use.

We identified 18 new loci in Approach 1 ($P_{SNPadjSMK}$) by adjusting for current smoking status. Our analyses did not allow us to determine whether these discoveries are due to different subsets of subjects included in the analyses compared to previous studies[6,7] or due to adjusting for current smoking. Adjustment for current smoking at these loci in our analyses, however, did reveal associations that were not previously identified. Specifically after accounting for smoking in our analyses, all novel BMI loci exhibit p values that are at least one order of magnitude lower than in previous GIANT investigations, despite a smaller sample size in the current analysis[7] (**Supplementary Table 14**). While the sample sizes for both WCadjBMI and WHRadjBMI are more comparable with previous GIANT investigations, our p-values for variants identified in Approach 1 are at least two orders of magnitude lower than previous findings. Thus, adjustment for smoking may have indeed revealed new loci. Further, loci identified in

Approach 2, including 9 novel loci, suggest that accounting for interaction improves our ability to detect these loci even in the presence of only modest evidence of GxSMK interaction.

We examined whether there were stronger effects in current smokers compared to nonsmokers. While we found that effects were not significantly stronger in current smokers as compared to nonsmokers for BMI, there is a greater proportion of variance in BMI explained by variants that are significant for Approach 1 (SNPadjSMK), which may be expected given that there are a greater number of variants with higher effect estimates in the current smokers. For WCadjBMI, there was no enrichment for stronger effects in one stratum compared to the others for our significant loci as well; however, there was a greater proportion of the variance in WCadjBMI for loci identified in Approach 1 (SNPadjSMK) in nonsmokers as compared to smokers. For WHRadjBMI, there were significantly more loci that exhibit greater effects in nonsmokers than in smokers, and this pattern was mirrored in the variance explained analysis. The large difference between the effect in smokers and nonsmokers likely explains the sub-GWS level of our loci in previous GIANT investigations[7]. For example, the T allele of rs7697556, 81kb from the *ADAMTS3* gene, was associated with increased WCadjBMI and exhibits a six-fold greater effect in nonsmokers compared to smokers, although the interaction effect was only nominal; in previous GWAS investigations by Locke et al. (2015) this variant was nearly GWS. These differences in effect estimates between smokers and nonsmokers across loci may help to explain inconsistent findings in previous analyses that show central adiposity increases with increased smoking, but is associated with decreased weight and BMI[13,20,21,95].

Our results support previous findings that implicate genes involved in transcription and gene expression, appetite regulation, macronutrient metabolism, and glucose homeostasis. Several of our novel loci have candidate genes within 500 kb of our tag variants that are highly expressed and/or active in brain tissue (*BBX*, *KIF1B*, *SOX11*, and *EPHA3*) and, like other obesity-associated genes, may be involved in previously-identified pathways linked to neuronal regulation of appetite (*KIF1B*, *GRIN2A*, and *SLC23A2*), adipo/angiogenesis (*ANGPTL3* and *TNF*) and glucose, lipid and energy homeostasis (*CD47*, *STK25*, *STK19*, *RAGE*, *AIF1*, *LYPLAL1*, *HDLBP*, *ANGPTL3*, *DOCK7*, *KIF1B*, *PREX1*, and *RPS12*).

However, many of our newly identified loci highlight novel biological functions and pathways where dysregulation may lead to increased susceptibility to obesity, including response to oxidative stress, addictive behavior, and newly identified regulatory functions. There is a growing body of evidence that supports the notion that exposure to oxidative stress leads to increased adiposity, risk of obesity, and poor cardiometabolic outcomes[55,96-99]. Our results for BMI and WCadjBMI, specifically associations identified near *CHRNA5-CHRNA3-CHRNB4, PRNP, SLC23A2, BACH1, and NMNAT1,* highlight new biological pathways and processes for future examination and may lead to a greater understanding of how oxidative stress leads to changes in obesity phenotypes and downstream cardiometabolic risk.

By incorporating current smoking into our analyses, we were able to identify 6 novel loci for BMI, 11 for WCadjBMI, and 6 for WHRadjBMI, and highlight novel biological processes and regulatory functions for genes implicated in increased obesity risk. Most previously identified loci were confirmed in our analyses with adjustment for current smoking status with smaller sample sizes than were needed in previous discovery analyses. A typical approach in large-scale GWAS meta-analyses is to not adjust for covariates such as current smoking; however, our findings highlight the importance of accounting for environmental exposures in genetic analyses.

ONLINE METHODS

Main Analyses

Study Design Overview

We applied four approaches to identify genetic loci that influence adiposity traits by accounting for current tobacco smoking status (**Figure 1**). We defined smokers as those who responded that they were currently smoking; not current smokers were those that responded "no" to currently smoking. We evaluated three traits: body mass index (BMI), waist circumference adjusted for BMI (WCadjBMI), and waist-to-hip ratio adjusted for BMI (WHRadjBMI). Our first two meta-analytical approaches were aimed at determining whether there are novel genetic variants that affect adiposity traits by adjusting for SMK (SNPadjSMK), or by jointly accounting for SMK and for interaction with SMK (SNPjoint); while Approaches 3 and 4 aimed to determine whether there are genetic variants that affect adiposity traits through interaction with SMK (SNPint and SNPscreen) (**Figure 1**). Our ***primary meta-analyses*** focused on results from all ancestries, sexes combined. ***Secondary meta-analyses*** were performed using the European-descent populations only, as well as stratified by sex (men-only and women-only) in all ancestries and in European-descent study populations.

Cohort Descriptions and Sample Sizes

The GIANT consortium was formed by an international group of researchers interested in understanding the genetic architecture of anthropometric traits (see **Supplemental Tables 1-4** for study sample sizes and descriptive statistics). In total, we included up to 79 studies comprising up to 241,258 individuals for BMI (51,080 smokers, 190,178 nonsmokers), 208,176 for WCadjBMI (43,226 smokers, 164,950 nonsmokers), and 189,180 for WHRadjBMI (40,543 smokers, 148,637 nonsmokers) with HapMap II imputed genome-wide chip data (up to 2.8M SNPs in association analyses), and/or with genotyped MetaboChip data (~195K SNPs in association analyses)[100]. In instances where studies submitted both Metabochip and GWAS data, these were for non-overlapping individuals. Each study's Institutional Review Board has approved this research and all study participants have provided written informed consent.

Phenotype descriptions

Our study highlights three traits of interest: BMI, WCadjBMI and WHRadjBMI. Height and weight, used to calculate BMI (kg/m^2), were measured in all studies; waist and hip circumferences were measured in the vast majority. For each sex, traits were adjusted using linear regression for age and age^2 (as well as for BMI for WCadjBMI and WHRadjBMI), and (when appropriate) for study site and principal components to account for ancestry. Family studies used linear mixed effects models to account for familial relationships and also conducted analyses for men and women combined including sex in the model. Phenotype residuals were obtained from the adjustment models and were inverse normally transformed subsequently to facilitate comparability across studies and with previously published analyses. The trait transformation was conducted separately for smokers and nonsmokers for the SMK-stratified model and using all individuals for the SMK-adjusted model.

Defining Smokers

The participating studies have varying levels of information on smoking, some with a simple binary variable and others with repeated, precise data. Since the effects of smoking cessation on adiposity appear to be immediate[16,17,23], a binary smoking trait (current smoker vs. not current smoker) is used for the analyses as most studies can readily derive this variable. We did not use a variable of 'ever smoker

vs. never' as it increases heterogeneity across studies, thus adding noise; also this definition would make harmonization across studies difficult.

Genotype Identification and Imputation

Studies with GWAS array data or Metabochip array data contributed to the results. Each study applied study-specific standard exclusions for sample call rate, gender checks, sample heterogeneity and ethnic group outliers (**Supplementary Table 2**). For most studies (except those that employed directly typed MetaboChip genotypes), genome-wide chip data was imputed to the HapMap II reference data set via MACH[101], IMPUTE[102], BimBam[103] or Beagle[104].

Study Level Analyses

To obtain study-specific summary statistics used in subsequent meta-analyses, the following linear models (or linear mixed effects models for studies with families/related individuals) were run separately for men and women and separately for cases and controls for case-control studies using phenotype residuals from the models described above. Studies with family data also conducted analyses with these models for men and women combined after accounting for dependency among family members as a function of their kinship correlations. We assumed an additive genetic model.

SMK-adjusted: $TRAIT = \beta_0 + \beta_1 SNP + \beta_2 SMK$

SMK-stratified: $TRAIT = \beta_0 + \beta_1 SNP$ (run in current smokers and nonsmokers separately)

The analyses were run using various GWAS software, including MACH2QTL [105], SNPTEST [106], ProbABEL[107], GenABEL[108], Merlin[109], PLINK[110] or QUICKTEST[111].

Quality control of study-specific summary statistics

The aggregated summary statistics were quality-controlled according to a standardized protocol[112]. These included checks for issues with trait transformations, allele frequencies and strand. Low quality SNPs in each study were excluded for the following criteria: (i) SNPs with low minor allele count (MAC <= 5, MAC = MAF * N) and monomorphic SNPs, (ii) genotyped SNPs with low SNP call-rate (< 95%) or low Hardy-Weinberg equilibrium test P-Value (< 10^{-6}), (iii) imputed SNPs with low imputation quality (MACH-Rsq or OEVAR<0.3, or information score <0.4 for SNPTEST/IMPUTE/IMPUTE2, or <0.8 for PLINK). To test for issues with relatedness or overlapping samples and to correct for potential population stratification, the study-specific standard errors and association P-Values were genomic control (GC) corrected using lambda factors[113] (**Supplementary Figure 1**). GC correction for GWAS data used all SNPs, but GC correction for MetaboChip data were restricted to chip QT interval SNPs only as the chip was enriched for associations with obesity-related traits. Any study-level GWAS file with a lambda > 1.5 was removed from further analyses. While we established this criterion, no study results were removed for this reason.

Meta-analyses

Meta-analyses used study-specific summary statistics for the phenotype associations for each of the above models. We used a fixed-effects inverse variance weighted method for the SNP main effect analyses. All meta-analyses were run in METAL[114]. As study results came in two separate batches (Stage 1 and Stage 2), meta-analyses from the two stages were further meta-analyzed (Stage 1 + Stage 2). A second GC correction was applied to all SNPs when combining Stage 1 and Stage 2 meta-analyses in the final meta-analysis. First, Hapmap-imputed GWAS data were meta-analyzed together, as were Metabochip studies. This step was followed by a combined GWAS + Metabochip meta-analysis. For primary analyses, we conducted meta-analyses across ancestries and sexes. For secondary meta-

analyses, we conducted meta-analyses in European-descent studies alone, and sex-specific meta-analyses. There were two reasons for conducting secondary meta-analyses. First, both WCadjBMI and WHRadjBMI have been shown to display sex-specific genetic effects[7,31,115]. Second, by including populations from multiple ancestries in our primary meta-analyses, we may be introducing heterogeneity due to differences in effect sizes, allele frequencies, and patterns of linkage disequilibrium across ancestries, potentially decreasing power to detect genetic effects. See **Supplementary Figure 1** for a summary of the primary meta-analysis study design. The obtained SMK-stratified summary statistics were later used to calculate summary SNPjoint and SNPint statistics using EasyStrata[116]. Briefly, this software implements a two-sample, large sample test of equal regression parameters between smokers and nonsmokers as described by Randall et al[31] for SNPint and the two degree of freedom test of main and interaction effects for SNPjoint as described by Aschard et al[30].

Lead SNP selection

Before selecting a lead SNP for each locus, SNPs with high heterogeneity $I^2 \geq 0.75$ or a minimum sample size below 50% of the maximum N for each strata (e.g. N> max[N Women Smokers]/2) were excluded. Lead SNPs that met significance criteria were selected based on distance (+/- 500 kb), and we defined the SNP with the lowest P-value as the top SNP for a locus. SNPs that reached genome-wide significance (GWS), but had no other SNPs within 500 kb with a P<1E-5 (lonely SNPs), were excluded from the SNP selection process. Two variants were excluded from Approach 2 based on this criterion, rs2149656 for WCadjBMI and rs2362267 for WHRadjBMI.

Approaches

Figure 1 outlines the four approaches that we used to identify novel SNPs. The left side of Figure 1 focuses on the first hypothesis that examines the effect of SNPs on adiposity traits. *Approach 1* considered a linear regression model that includes the SNP and SMK, thus adjusting for SMK (SNPadjSMK). Summary SNPadjSMK results were obtained from the SMK-adjusted meta-analysis. *Approach 2* used summary SMK-stratified meta-analysis results as described by Aschard et al.[30] to consider the joint hypothesis that a genetic variant has main and/or interaction effects on outcomes as a 2 degree of freedom test (SNPjoint). For this approach, the null hypothesis was that there is no main and no interaction effect on the outcome. Thus, rejection of this hypothesis could be due to either a main effect or an interaction effect or to both.

The right side of Figure 1 focuses on our second hypothesis, testing for interaction of a variant with SMK on adiposity traits as outcomes. *Approach 3* used the SMK-stratified results to directly contrast the regression coefficients for a test of interaction (SNPint)[31]. *Approach 4* used a screening strategy to evaluate interaction, whereby the SMK-adjusted main effect results (Approach 1) were screened for variants significant at the $P<5x10^{-8}$ level. These variants were then carried forward for a test of interaction, comparing the SMK-stratified specific regression coefficients in the second step (SNPscreen).

In *Approaches 1-3* variants significant at $P<5x10^{-8}$ were considered GWS. In *Approach 4* (SNPscreen) variants for which the p-value of the test of interaction is less than 0.05 divided by the number of variants carried forward were considered significant for interaction. We performed analytical power computations to demonstrate the usefulness and characteristic of the two interaction Approaches.

LocusZoom Plots

Regional association plots were generated for novel loci using the program Locuszoom[117]. For each plot, LD was calculated using a multiethnic sample of the 1000 Genomes Phase I reference panels[118,119], including EUR, AFR, EAS, and AMR. Previous SNP-trait associations highlighted within the plots include traits of interest (e.g. cardiometabolic, addiction, behavior, anthropometrics) found in the NHGRI-EMI GWAS Catalog and supplemented with recent GWAS studies from the literature[6,7,115,120].

Conditional Analyses

To determine if multiple association signals were present within a single locus, we used GCTA[32] to perform approximate joint conditional analyses on the SNPadjSMK and SMK- stratified data. The following criteria were used to select candidate loci for conditional analyses: nearby SNP (+/- 500kb) with an R^2>0.4 and an association P<1E-5 for any of our primary analyses. GCTA uses associations from our meta-analyses and LD estimates from reference data sets containing individual-level genotypic data to perform the conditional analyses. To calculate the LD structure, we used two U.S. cohorts, the Atherosclerosis Risk in Communities (ARIC) study consisting of 9,713 individuals of European descent and 580 individuals of African American descent, and the Framingham Heart Study (FramHS) consisting of 8,481 individuals of European ancestry, both studies imputed to HapMap r22. However, because our primary analyses were conducted in multiple ancestries, each study supplemented the genetic data using HapMap reference populations so that the final reference panel was composed of about 1-3% Asians (CHB + JPT) and 4-6% Africans (YRI for the FramHS) for the entire reference sample. We extracted each 1 MB region surrounding our candidate SNPs, performed joint approximate conditional analyses, and then repeated the steps for the appropriate Approach to identify additional association signals.

Many of the SNPs identified in the current analyses were nearby SNPs previously associated with related anthropometric and obesity traits (e.g. height, visceral adipose tissue). For all lead SNPs near a SNP previously associated with these traits, GCTA was also used to perform approximate conditional analyses on the SNPadjSMK and SMK-stratified data in order to determine if the loci identified here are independent of the previously identified SNP-trait associations.

Follow-up Analyses

Biological Summaries

To identify genes that may be implicated in the association between our lead SNPs (Tables 1-3) and BMI, WHRadjBMI, and WCadjBMI, and to shed light on the complex relationship between genetic variants, SMK and adiposity, we performed in-depth literature searches on nearby candidate genes. Snipper v1.2 (http://csg.sph.umich.edu/boehnke/snipper/) was used to identify any genes and cis- or trans-eQTLs within 500kb of our lead SNPs. All genes identified by Snipper were manually curated and examined for evidence of relationship with smoking and/or adiposity. To explore any potential regulatory or function role of the association regions, loci were also examined using several bioinformatic tools/databases, including HaploReg v4.1[121], UCSC Genome Browser[122] (available at http://genome.ucsc.edu/), GTeX Portal[123], and RegulomeDB[124].

eQTL Analyses

We used two approaches to systematically explore the role of novel loci in regulating gene expression. First, to gain a general overview of the regulatory role of newly identified GWAS regions, we conducted an eQTL lookup using >50 eQTL studies[125], with specific citations for >100 datasets included in the current query for blood cell related eQTL studies and relevant non-blood cell tissue eQTLs (e.g. adipose and brain tissues). Additional eQTL data was integrated from online sources including ScanDB, the Broad

Institute GTEx Portal, and the Pritchard Lab (eqtl.uchicago.edu). Additional details on the methods, including study references can be found in **Supplementary Note 3**. Only significant cis-eQTLS in high LD with our novel lead SNPs (r2>0.9, calculated in the CEU+YRI+CHB+JPT 1000 Genomes reference panel), or proxy SNPs, were retained for consideration.

Second, since public databases with eQTL data do not have information available on current smoking status, we also conducted a cis-eQTL association analysis using expression results derived from fasting peripheral whole blood using the Human Exon 1.0 ST Array (Affymetrix, Inc., Santa Clara, CA). The raw expression data were quantile-normalized, log2 transformed, followed by summarization using Robust Multi-array Average[126] and further adjusted for technical covariates, including the first principal component of the expression data, batch effect, the all-probeset-mean residual, blood cell counts, and cohort membership. We evaluated all transcripts +/- 1MB around each novel variant in the Framingham Heart Study while accounting for current smoking status, using the following four approaches similar to those used in our primary analyses of our traits: 1) eQTL adjusted for SMK, 2) eQTL stratified by SMK, 3) eQTL x SMK interaction, and 4) joint main + eQTLxSMK interaction). Significance level was evaluated by FDR < 5% per eQTL analysis and across all loci identified for that model in the primary meta-analysis. Additional details can be found in **Supplementary Note 3**.

Variance-explained estimates

We estimated the phenotypic variance in smokers and nonsmokers explained by the association signals using a method previously described by Kutalik et al.[127] For each associated region, we selected subsets of SNPs within 500 kb of our lead SNPs and based on varying P value thresholds (ranging from $1x10^{-8}$ to 0.1) from Approach 1 (SNPadjSMK model). First, each subset of SNPs was clumped into independent regions to identify the lead SNP for each region. The variance explained by each subset of SNPs in the SMK and nonSMK strata was estimated by summing the variance explained by the individual lead SNPs.

Smoking Behavior Lookups

In order to determine if any of the loci identified in the current study are associated with smoking behavior, we conducted a look-up of all lead SNPs from novel loci and Approach 3 in existing GWAS of smoking behavior[9]. The analysis consists of phasing study-specific GWAS samples contributing to the smoking behavior meta-analysis, imputation, association testing and meta-analysis. To ensure that all SNPs of interest were available in the smoking GWAS, the program SHAPEIT2[128] was used to phase a region 500Kb either side of each lead SNP, and imputation was carried out using IMPUTE2[129] with the 1000 Genomes Phase 3 dataset as a reference panel.

Each region was analyzed for 3 smoking related phenotypes: (i) Ever vs Never smokers, (ii) Current vs Non-current smokers, and (iii) a categorical measure of smoking quantity[90]. The smoking quantity levels were 0 (defined as 1-10 cigarettes per day [CPD]), 1 (11-20 CPD), 2 (21-30 CPD) and 3 (31 or more CPD). Each increment represents an increase in smoking quantity of 10 cigarettes per day. There were 10,058 Never smokers, 13,418 Ever smokers, 11,796 Non-current smokers, 6,966 Current smokers and 11,436 samples with the SQ phenotypes. SNPMETA[90] was used to perform an inverse-variance weighted fixed effects meta-analysis across cohorts at all SNPs in each region, and included a single GC correction. At each SNP, only those cohorts that had an imputation info score > 0.5 were included in the meta-analysis.

Main Effects Lookup in Previous GIANT Investigations

To better understand why our novel variants remained undiscovered in previous investigations that did not take SMK into account, we also conducted a lookup of our novel variants in published GWAS results examining genetic main effects on BMI, WCadjBMI, and WHRadjBMI.[6,7]

GWAS Catalog Lookups

To further investigate the identified genetic variants in this study and to gain additional insight into their functionality and possible effects on related cardiometabolic traits, we searched for previous SNP-trait associations nearby our lead SNPs. PLINK was used to find all SNPs within 500 kb of any of our lead SNPs and calculate r^2 values using a combined ancestry (AMR, AFR, EUR, ASN) 1000 Genomes Phase 1 reference panel[119] to allow for LD calculation for SNPs on the Illumina Metabochip and to best estimate LD in our multiethnic GWAS. All SNPs within the specified regions were compared with the NHGRI-EBI (National Human Genome Research Institute, European Bioinformatics Institute) GWAS Catalog, version 1.0 (www.ebi.ac.uk/gwas)[91,92] for overlap, and distances between the two SNPs were calculated using STATA v14, for the chromosome and base pair positions based on human genome reference build 19. All previous associations within 500 kb and with an R^2>0.5 with our lead SNP were retained for further interrogation.

Genetic risk score calculation

We calculated several unweighted genetic risk scores (GRSs) for each individual in the population-based KORA-S3 and KORA-S4 studies (total N = 3,457). We compared GRSs limited to previously known lead SNPs (see **Supplementary Tables 11-13** for lists of previously known lead SNPs**)** with GRSs based on previously known and novel lead SNPs from the current study (see **Tables 1-4** for lists of novel lead SNPs). Risk scores were tested for association with the obesity trait using the following linear regression models: The unadjusted GRS model (TRAIT = $\beta_0 + \beta_1$GRS), the adjusted GRS model (TRAIT = $\beta_0 + \beta_1$GRS + β_2SMK) and the GRSxSMK interaction model (TRAIT = $\beta_0 + \beta_1$GRS + β_2SMK + β_3GRSxSMK).

ACKNOWLEDGEMENTS

A full list of acknowledgments appears in the Supplementary Note 4. Funding for this study was provided by the Aase and Ejner Danielsens Foundation; Academy of Finland (41071, 77299, 102318, 110413, 117787, 121584, 123885, 124243, 124282, 126925, 129378, 134309, 286284); Accare Center for Child and Adolescent Psychiatry; Action on Hearing Loss (G51); Agence Nationale de la 359 Recherche; Agency for Health Care Policy Research (HS06516); ALF/LUA research grant in Gothenburg; ALFEDIAM; ALK-Abello´ A/S; Althingi; American Heart Association (13POST16500011); Amgen; Andrea and Charles Bronfman Philanthropies; Ardix Medical; Arthritis Research UK; Association Diabète Risque Vasculaire; Australian National Health and Medical Research Council (241944, 339462, 389875, 389891, 389892, 389927, 389938, 442915, 442981, 496739, 552485, 552498); Avera Institute; Bayer Diagnostics; Becton Dickinson; BHF (RG/14/5/30893); Boston Obesity Nutrition Research Center (DK46200), Bristol-Myers Squibb; British Heart Foundation (RG/10/12/28456, RG2008/08, RG2008/014, SP/04/002); Medical Research Council of Canada; Canadian Institutes for Health Research (FRCN-CCT-83028); Cancer Research UK; Cardionics; Cavadis B.V., Center for Medical Systems Biology; Center of Excellence in Genomics; CFI; CIHR; City of Kuopio; CNAMTS; Cohortes Santé TGIR; Contrat de Projets État-Région; Croatian Science Foundation (8875); Danish Agency for Science, Technology and Innovation; Danish Council for Independent Research (DFF-1333-00124, DFF-1331-00730B); County Council of Dalarna; Dalarna University; Danish Council for Strategic Research; Danish Diabetes Academy; Danish Medical Research Council; Department of Health, UK; Development Fund from the University of Tartu (SP1GVARENG); Diabetes Hilfs- und Forschungsfonds Deutschland; Diabetes UK; Donald W. Reynolds Foundation; Dr. Robert Pfleger-Stiftung; Dutch Brain Foundation; Dutch Diabetes Research Foundation; Dutch Inter University Cardiology Institute; Dutch Kidney Foundation (E033); Dutch Ministry of Justice; the DynaHEALTH action No 633595, Economic Structure Enhancing Fund of the Dutch Government; Else Kröner-Fresenius-Stiftung (2012_A147, P48/08//A11/08); Emil Aaltonen Foundation; Erasmus University Medical Center Rotterdam; Erasmus MC and Erasmus University Rotterdam; the Municipality of Rotterdam; Estonian Government (IUT20-60, IUT24-6); Estonian Research Roadmap through the Estonian Ministry of Education and Research (3.2.0304.11-0312); European Research Council (ERC Starting Grant); European Regional Development Fund; European Science Foundation (EU/QLRT-2001-01254); European Commission (018947, 018996, 201668, 223004, 230374, 279143, 284167, 305739, BBMRI-LPC-313010, HEALTH-2011.2.4.2-2-EU-MASCARA, HEALTH-2011-278913, HEALTH-2011-294713-EPLORE, HEALTH-F2-2008-201865-GEFOS, HEALTH-F2-2013-601456, HEALTH-F4-2007-201413, HEALTH-F4-2007-201550-HYPERGENES, HEALTH-F7-305507 HOMAGE, IMI/115006, LSHG-CT-2006-018947, LSHG-CT-2006-01947, LSHM-CT-2004-005272, LSHM-CT-2006-037697, LSHM-CT-2007-037273, QLG1-CT-2002-00896, QLG2-CT-2002-01254); Faculty of Biology and Medicine of Lausanne; Federal Ministry of Education and Research (01ZZ0103, 01ZZ0403, 01ZZ9603, 03IS2061A, 03ZIK012); Federal State of Mecklenburg-West Pomerania; Fédération Française de Cardiologie; Finnish Cultural Foundation; Finnish Diabetes Association; Finnish Foundation of Cardiovascular Research; Finnish Heart Association; Fondation Leducq; Food Standards Agency; Foundation for Strategic Research; French Ministry of Research; FRSQ; Genetic Association Information Network (GAIN) of the Foundation for the NIH; German Federal Ministry of Education and Research (BMBF, 01ER1206, 01ER1507); GlaxoSmithKline; Greek General Secretary of Research and Technology; Göteborg Medical Society; Health and Safety Executive; Healthcare NHS Trust; Healthway; Western Australia; Heart Foundation of Northern Sweden; Helmholtz Zentrum München – German Research Center for Environmental Health; Hjartavernd; Ingrid Thurings Foundation; INSERM; InterOmics (PB05 MIUR-CNR); INTERREG IV Oberrhein Program (A28); Interuniversity Cardiology Institute of the Netherlands (ICIN, 09.001); Italian Ministry of Health (ICS110.1/RF97.71); Italian Ministry of Economy and Finance (FaReBio di Qualità); Marianne and Marcus Wallenberg Foundation; the Ministry of Health, Welfare and Sports, the Netherlands; John D and Catherine T MacArthur Foundation Research Networks on Successful Midlife Development and

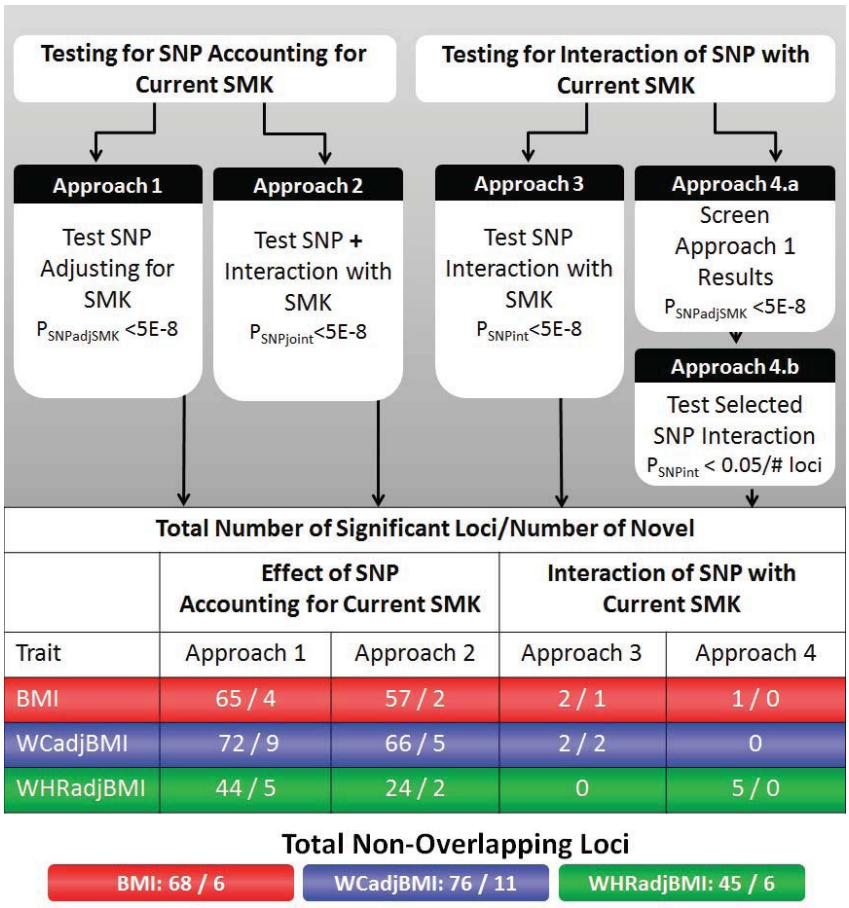
Socioeconomic Status and Health; Juho Vainio Foundation; Juvenile Diabetes Research Foundation International; KfH Stiftung Präventivmedizin e.V.; King's College London; Knut and Alice Wallenberg Foundation; Kuopio University Hospital; Kuopio, Tampere and Turku University Hospital Medical Funds (X51001); La Fondation de France; Leenaards Foundation; Lilly; LMUinnovativ; Lundberg Foundation; Magnus Bergvall Foundation; MDEIE; Medical Research Council UK (G0000934, G0601966, G0700931, MC_U106179471, MC_UU_12019/1); MEKOS Laboratories; Merck Santé; Ministry for Health, Welfare and Sports, The Netherlands; Ministry of Cultural Affairs of Mecklenburg-West Pomerania; Ministry of Economic Affairs, The Netherlands; Ministry of Education and Culture of Finland (627;2004-2011); Ministry of Education, Culture and Science, The Netherlands; Ministry of Science, Education and Sport in the Republic of Croatia (108-1080315-0302); MRC centre for Causal Analyses in Translational Epidemiology; MRC Human Genetics Unit; MRC-GlaxoSmithKline pilot programme (G0701863); MSD Stipend Diabetes; National Institute for Health Research; Netherlands Brain Foundation (F2013(1)-28); Netherlands CardioVascular Research Initiative (CVON2011-19); Netherlands Genomics Initiative (050-060-810); Netherlands Heart Foundation (2001 D 032, NHS2010B280); Netherlands Organization for Scientific Research (NWO) and Netherlands Organisation for Health Research and Development (ZonMW) (56-464-14192, 60-60600-97-118, 100-001-004, 261-98-710, 400-05-717, 480-04-004, 480-05-003, 481-08-013, 904-61-090, 904-61-193, 911-11-025, 985-10-002, Addiction-31160008, BBMRI–NL 184.021.007, GB-MaGW 452-04-314, GB-MaGW 452-06-004, GB-MaGW 480-01-006, GB-MaGW 480-07-001, GB-MW 940-38-011, Middelgroot-911-09-032, NBIC/BioAssist/RK 2008.024, Spinozapremie 175.010.2003.005, 175.010.2007.006); Neuroscience Campus Amsterdam; NHS Foundation Trust; National Institutes of Health (1RC2MH089951, 1Z01HG000024, 24152, 263MD9164, 263MD821336, 2R01LM010098, 32100-2, 32122, 32108, AA07535, AA10248, AA11998, AA13320, AA13321, AA13326, AA14041, AA17688, AG13196, CA047988, DA12854, DK56350, DK063491, DK078150, DK091718, DK100383, DK078616, ES10126, HG004790, HHSN268200625226C, HHSN268200800007C, HHSN268201200036C, HHSN268201500001I, HHSN268201100046C, HHSN268201100001C, HHSN268201100002C, HHSN268201100003C, HHSN268201100004C, HHSN271201100004C, HL043851, HL45670, HL080467, HL085144, HL087660, HL054457, HL119443, HL-118305, HL071981, HL034594, HL126024, MH66206, MH081802, N01AG12100, N01HC55015, N01HC55016, N01C55018, N01HC55019, N01HC55020, N01HC55021, N01HC55022, N01HC85079, N01HC85080, N01HC85081, N01HC85082, N01HC85083, N01HC85086, N01HC95159, N01HC95160, N01HC95161, N01HC95162, N01HC95163, N01HC95164, N01HC95165, N01HC95166, N01HC95167, N01HC95168, N01HC95169, N01HG65403, N01WH22110, N02HL6-4278, N01-HC-25195, P01CA33619, R01AG023629, R01CA63, R01D004215701A, R01DK075787, R01DK062370, R01DK072193, R01DK075787, R01DK089256, R01HL53353, R01HL59367, R01HL086694, R01HL087641, R01HL087652, R01HL103612, R01HL105756, R01HL117078, R01HL120393, R37CA54281, RC2AG036495, RC4AG039029, RPPG040710371, RR20649, TW008288, TW05596, U01AG009740, U01CA98758, U01CA136792, U01DK062418, U01HG004402, U01HG004802, U01HG007376, U01HL080295, UL1RR025005, UL1TR000040, UL1TR000124, UL1TR001079, 2T32HL007055-36, T32GM074905, HG002651, HL084729, N01-HC-25195, UM1CA182913); NIH, National Institute on Aging (Intramural funding, NO1-AG-1-2109); Northern Netherlands Collaboration of Provinces; Novartis Pharma; Novo Nordisk; Novo Nordisk Foundation; ONIVINS; Parnassia Bavo group; Pierre Fabre; Province of Groningen; Päivikki and Sakari Sohlberg Foundation; Påhlssons Foundation; Paavo Nurmi Foundation; Radboud Medical Center Nijmegen; Research Centre for Prevention and Health, the Capital Region of Denmark; the Research Institute for Diseases in the Elderly; Research into Ageing; Robert Dawson Evans Endowment of the Department of Medicine at Boston University School of Medicine and Boston Medical Center; Roche; Royal Society; Russian Foundation for Basic Research (NWO-RFBR 047.017.043); Rutgers University Cell and DNA Repository (NIMH U24 MH068457-06); Sanofi-Aventis; Scottish Government Health Directorates, Chief Scientist Office (CZD/16/6); Siemens Healthcare; Social Insurance Institution of Finland (4/26/2010); Social Ministry of

the Federal State of Mecklenburg-West Pomerania; Société Francophone du 358 Diabète; State of Bavaria; Stiftelsen för Gamla Tjänarinnor; Stockholm County Council (560183, 592229); Strategic Cardiovascular and Diabetes Programmes of Karolinska Institutet and Stockholm County Council; Stroke Association; Swedish Diabetes Association; Swedish Diabetes Foundation (2013-024); Swedish Foundation for Strategic Research; Swedish Heart-Lung Foundation (20120197); Swedish Research Council (0593, 8691, 2012-1397, 2012-1727, and 2012-2215); Swedish Society for Medical Research; Swiss Institute of Bioinformatics; Swiss National Science Foundation (3100AO-116323/1, 31003A-143914, 33CSCO-122661, 33CS30-139468, 33CS30-148401, 51RTP0_151019); Tampere Tuberculosis Foundation; Technology Foundation STW (11679); The Fonds voor Wetenschappelijk Onderzoek Vlaanderen, Ministry of the Flemish Community (G.0880.13, G.0881.13); The Great Wine Estates of the Margaret River Region of Western Australia; Timber Merchant Vilhelm Bangs Foundation; Topcon; Tore Nilsson Foundation; Torsten and Ragnar Söderberg's Foundation; United States – Israel Binational Science Foundation (Grant 2011036), Umeå University; University Hospital of Regensburg; University of Groningen; University Medical Center Groningen; University of Michigan; University of Utrecht; Uppsala Multidisciplinary Center for Advanced Computational Science (UPPMAX) (b2011036); Velux Foundation; VU University's Institute for Health and Care Research; Västra Götaland Foundation; Wellcome Trust (068545, 076113, 079895, 084723, 088869, WT064890, WT086596, WT098017, WT090532, WT098051, 098381); Wissenschaftsoffensive TMO; Yrjö Jahnsson Foundation; and Åke Wiberg Foundation. The views expressed in this manuscript are those of the authors and do not necessarily represent the views of the National Heart, Lung, and Blood Institute (NHLBI); the National Institutes of Health (NIH); or the U.S. Department of Health and Human Services.

COMPETING FINANCIAL INTERESTS

Bruce Psaty serves on the DSMB for a clinical trial funded by the device manufacturer (Zoll LifeCor) and on the Steering Committee of the Yale Open Data Access Project funded by Johnson & Johnson.

Figure 1. Summary of study design and results. Approach 1 uses both SNP and SMK in the association model. Approaches 2 and 3 use the SMK-stratified meta-analyses. Approach 4 screens loci based on Approach 1, then uses SMK-stratified results to identify loci with significant interaction effects (**Online Methods**).



Total Number of Significant Loci/Number of Novel

	Effect of SNP Accounting for Current SMK		Interaction of SNP with Current SMK	
Trait	Approach 1	Approach 2	Approach 3	Approach 4
BMI	65 / 4	57 / 2	2 / 1	1 / 0
WCadjBMI	72 / 9	66 / 5	2 / 2	0
WHRadjBMI	44 / 5	24 / 2	0	5 / 0

Total Non-Overlapping Loci

BMI: 68 / 6 | WCadjBMI: 76 / 11 | WHRadjBMI: 45 / 6

Figure 2. Estimated effect estimates (β ± 95% CI) per risk allele for A) BMI, B) WCadjBMI, and C) WHRadjBMI for novel loci from Approaches 1 and 2 (SNPadjSMK and SNPjoint, respectively) and all loci from Approaches 3 and 4 (SNPint and SNPscreen) identified in the primary meta-analyses. Loci are ordered by greater magnitude of effect in smokers compared to nonsmokers and labeled with the nearest gene. For the locus near *TMEM38B,* rs9409082 was used for effect estimates in this plot. (¥ loci identified for Approach 4, *loci identified for Approach 3).

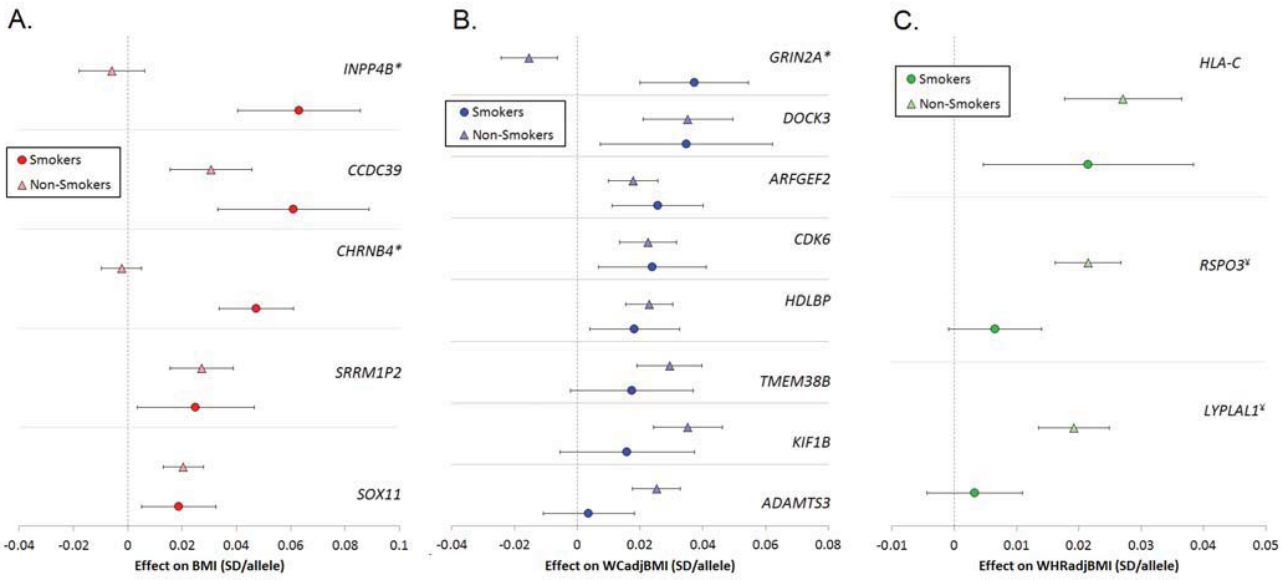


Figure 3. Power comparison. Shown is the power to identify GxSMK interaction for various types of GxSMK interaction effects given 50,000 smokers and 180,000 nonsmokers and using Approach 3 (SNPint, magenta) or Approach 4 (SNPscreen, orange). A: Assuming an effect in smokers (R^2_{SMK}=0.0007, MAF = 22%, similar to the realistic *BDNF* effect on BMI[2]) and for various effects in nonsmokers; B: Assuming an effect in nonsmokers (R^2_{NONSMK}=0.0007, MAF = 22%, similar to the realistic *BDNF* effect on BMI) and for various effects in smokers.

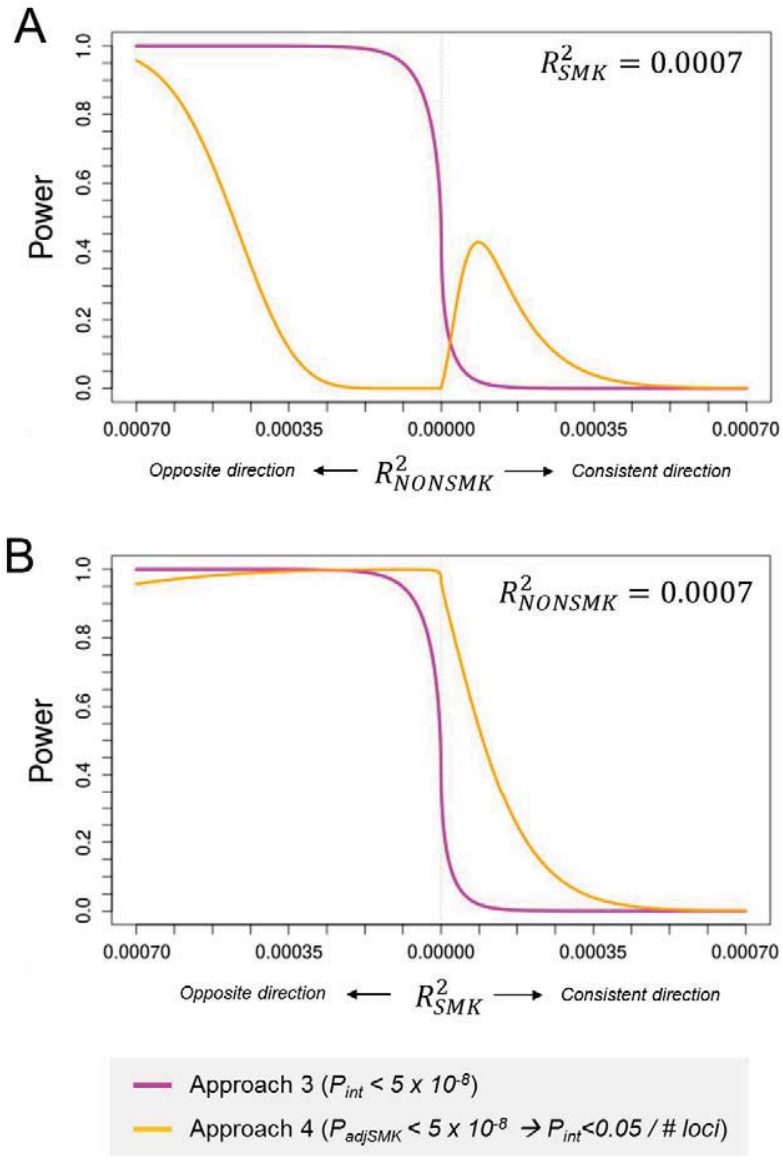


Figure 4. Total smoking status-specific explained variance by SNPs meeting varying thresholds of overall association in Approach 1 (SNPadjSMK) and the difference between the proportion of variance explained between smokers and nonsmokers for these same sets of SNPs in BMI (A,B), WCadjBMI (C,D), and for WHRadjBMI (E,F).

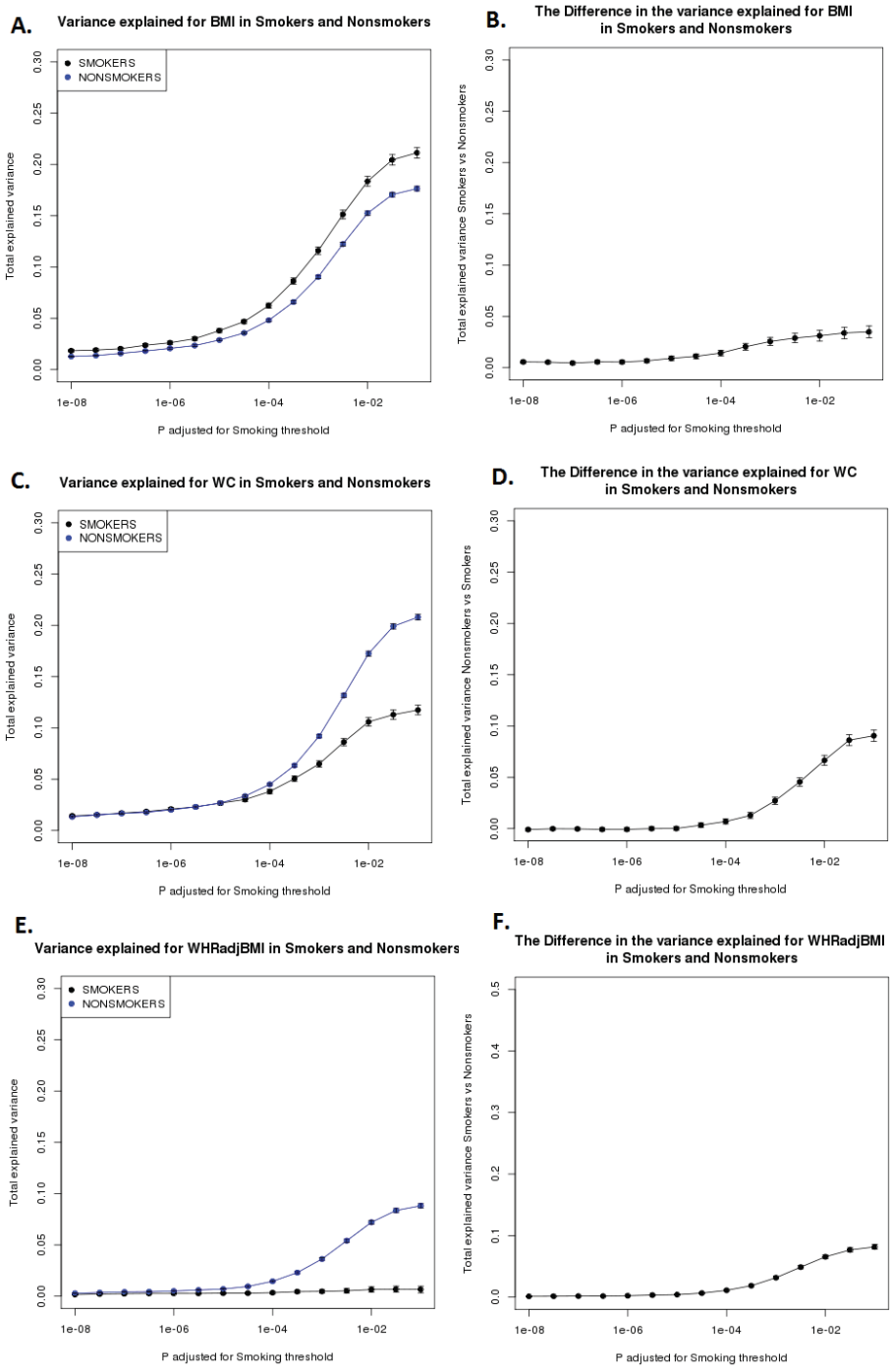


Table 1. Summary of association results for novel loci reaching genome-wide significance in Approach 1 ($P_{SNPadjSMK}$<5E-8) or Approach 2 ($P_{SNPjoint}$ <5E-8) for our primary meta-analysis in combined ancestries and combined sexes.

App	Marker	Chr:Pos (hg19)	Nearest Gene	N	EAF	Alleles E/O	SMOKERS		NON-SMOKERS		Main and Interaction Effects			
							β	P	β	P	β_{adj}	$P_{SNPadjSMK}$	P_{SNPint}	$P_{SNPjoint}$
BMI														
1,2	rs10929925	2:6155557	*SOX11*	225,067	0.55	C/A	0.019	7.8E-03	0.020	8.4E-08	0.020	**1.1E-09**	8.2E-01	**1.6E-08**
1	rs6794880	3:84451512	*SRRM1P2*	186,968	0.85	A/G	0.025	2.3E-02	0.027	3.9E-06	0.028	**4.3E-08**	8.5E-01	1.8E-06
2	rs13069244	3:180441172	*CCDC39*	233,776	0.08	A/G	0.061	1.8E-05	0.031	6.6E-05	0.035	1.2E-07	4.6E-02	**3.5E-08**
WCadjBMI														
1,2	rs17396340	1:10286176	*KIF1B*	206,485	0.14	A/G	0.016	1.4E-01	0.035	**4.7E-10**	0.028	**3.0E-08**	9.8E-02	**9.1E-10**
1,2	rs6743226	2:242236972	*HDLBP*	200,666	0.53	C/T	0.018	1.3E-02	0.023	**2.6E-09**	0.022	**1.2E-10**	5.5E-01	**5.8E-10**
1	rs4378999	3:51208646	*DOCK3*	156,566	0.13	T/A	0.035	1.3E-02	0.035	1.3E-06	0.036	**4.1E-08**	9.7E-01	4.1E-07
1,2	rs7697556	4:73515313	*ADAMTS3*	206,017	0.49	T/C	0.004	6.3E-01	0.025	**7.3E-11**	0.021	**5.2E-09**	6.7E-03	**7.6E-10**
1	rs10269774	7:92253972	*CDK6*	157,552	0.34	A/G	0.024	6.6E-03	0.023	1.1E-06	0.023	**2.9E-08**	8.8E-01	1.6E-07
1	rs6470765	8:130736697	*GSDMC*	157,450	0.76	A/C	0.032	1.9E-03	0.023	1.7E-05	0.026	**4.8E-08**	4.3E-01	9.5E-07
2	rs9408815	9:108890521	*TMEM38B*	156,427	0.75	C/G	0.012	2.3E-01	0.030	**4.2E-09**	0.026	**2.3E-08**	8.5E-02	**1.7E-08**
1	rs9409082	9:108901049		157,785	0.76	C/T	0.017	8.1E-02	0.029	**2.6E-08**	0.027	**1.5E-08**	2.7E-01	**4.6E-08**
1	rs6012558	20:47531286	*ARFGEF2*	208,004	0.41	A/G	0.026	5.4E-04	0.018	6.5E-06	0.020	**1.9E-08**	3.3E-01	1.3E-07
WHRadjBMI														
1,2	rs1049281	6:31236567	*HLA-C*	149,285	0.66	C/T	0.022	1.3E-02	0.027	**2.0E-08**	0.025	**2.2E-09**	5.6E-01	**5.3E-09**

Abbreviations: Chr- chromosome; Pos- position (bp); E/O- effect/other; EAF- effect allele frequency; adj- adjusted for smoking; int- interaction; App- Approach.

Table 2. Novel loci showing significant association in Approaches 1 (SNPadjSMK), 2 (SNPjoint), 3 (SNPint), and 4 (SNPscreen) for loci identified in secondary analysis samples, which were not identified in primary meta-analyses. All estimates are from the stratum specified in the Approach:Sample column (E-European-only, A- all ancestries, C- combined sexes, W-women only, M- men only). * This locus was filtered from approaches 2-4 due to low sample size in the SMK strata, and only p-values for Approach 1 are considered significant.

Approach: Strata	Marker	Chr:Pos (hg19)	Nearest Gene	N	EAF	Alleles E/O	SMOKERS β	SMOKERS P	NON-SMOKERS β	NON-SMOKERS P	Main and Interaction Effects β_{adj}	P_{adj}	P_{int}	P_{joint}
BMI														
1:EC	rs2481665	1:62594677	*INADL*	209,453	0.56	T/C	0.015	4.6E-02	0.021	8.9E-08	0.019	**3.5E-08**	4.0E-01	6.7E-08
1:AW	rs12629427	3:89145340	*EPHA3*	137,961	0.26	C/T	0.025	2.1E-02	0.028	3.6E-07	0.027	**4.8E-08**	8.0E-01	2.0E-07
1:EW	rs2173039	3:89142175		117,942	0.26	C/G	0.024	3.1E-02	0.032	8.9E-08	0.031	**7.3E-09**	5.7E-01	6.5E-08
WCadjBMI														
1:EM	rs1545348	5:34718343	*RAI14*	77,677	0.73	T/G	0.044	3.1E-04	0.030	1.9E-05	0.034	**1.8E-08**	3.2E-01	1.7E-07
2:EW	rs6076699	20:4566688	*PRNP*	76,930	0.97	A/G	0.169	1.4E-05	-0.070	1.2E-04	-0.034	3.5E-02	**1.4E-08**	**4.8E-08**
WHRadjBMI														
1:AW	rs670752	3:107312980	*BBX*	107,568	0.32	A/G	0.012	5.5E-02	0.009	1.5E-02	0.027	**4.9E-08**	6.8E-01	7.8E-03
1:EC	rs589428	6:31848220	*EHMT2*	162,918	0.66	G/T	0.006	1.2E-01	0.011	4.1E-04	0.022	**2.8E-08**	3.5E-01	7.0E-04
2:EC	rs1856293	6:133480940	*EYA4*	127,431	0.52	A/C	0.006	5.3E-01	-0.028	9.1E-09	-0.019	6.5E-06	5.4E-04	**4.7E-08**
1:AW	rs2001945	8:126477978	*TRIB1*	103,446	0.40	G/C	0.009	1.2E-01	0.013	1.0E-04	0.025	**4.7E-08**	5.9E-01	1.3E-04
1:EC	rs17065323	13:44627788	*SMIM2**	69,968	0.01	T/C	0.154	1.9E-01	-0.230	1.2E-10	-0.181	**9.2E-09**	1.4E-03	3.9E-10

Abbreviations: Chr- chromosome, Pos- position (bp), E/O- effect/other, EAF- effect allele frequency, Padj- adjusted for smoking, int- interaction.

Table 3. Summary of association results for loci showing significance for interaction with smoking in Approach 3 (SNPint) and/or Approach 4 (SNPscreen) in our primary meta-analyses of combined ancestries and combined sexes. Ł - known locus.

App	Marker	Chr:Pos (hg19)	Nearest Gene	N	EAF	Alleles E/O	SMOKERS		NON-SMOKERS		Main and Interaction Effects			
							β	P	β	P	β_{adj}	P_{adj}	P_{int}	P_{joint}
BMI														
3	rs336396	4:143062811	*INPP4B*	169,646	0.18	T/C	0.063	**4.8E-08**	-0.006	3.4E-01	0.007	2.3E-01	**2.1E-08**	1.9E-07
3	rs12902602 Ł	15:78967401	*CHRNB4*	240,135	0.62	A/G	0.047	**1.8E-11**	-0.002	5.5E-01	0.009	8.6E-03	**4.1E-11**	**1.1E-10**
WCadjBMI														
3	rs4141488	16:9629067	*GRIN2A*	153,892	0.50	T/C	0.037	2.2E-05	-0.015	9.6E-04	-0.003	4.4E-01	**2.7E-08**	5.0E-07
WHRadjBMI														
4	rs765751 Ł	1:219669226	*LYPLAL1*	189,028	0.64	C/T	0.003	3.9E-01	0.019	**3.1E-11**	0.029	**3.1E-16**	*7.3E-04*	**2.1E-10**
4	rs7766106 Ł	6:127455138	*RSPO3*	188,174	0.48	T/C	0.007	7.9E-02	0.022	**2.2E-15**	0.037	**3.7E-27**	*9.7E-04*	**3.8E-15**

Abbreviations: Chr- chromosome; Pos- position (bp); E/O- effect/other; EAF- effect allele frequency; adj- adjusted for smoking; int- interaction; App- Approach.

Table 4. Summary of association results for loci showing significance for interaction with smoking in Approach 3 (SNPint) and/or Approach 4 (SNPscreen) in our secondary meta-analyses not identified in primary meta-analyses. All estimates are from the stratum specified in the Approach:Sample column (E-European-only, A- all ancestries, C- combined sexes, W-women only, M- men only). Ł - known locus. The R^2 between the *ADAMTS7* (rs1809420) and *CHRNB4* variant (rs1290362) in **Table 3** is 0.72 (HapMap 2, CEU). Additionally, the *PRNP* variant (rs6076699) is the same as the variant that came up from Approach 2 (**Table 2**).

Approach: Strata	Marker	Chr:Pos (hg19)	Nearest Gene	N	EAF	Alleles E/O	SMOKERS β	SMOKERS P	NON-SMOKERS β	NON-SMOKERS P	Main and Interaction Effects β_{adj}	P_{adj}	P_{int}	P_{joint}
BMI														
4:AM	rs1809420 Ł	15:79056769	*ADAMTS7*	57,081	0.59	T/C	0.074	**9.8E-08**	0.023	2.0E-03	0.036	**4.9E-08**	*9.4E-04*	**5.6E-09**
WCadjBMI														
3:EW	rs6076699	20:4566688	*PRNP*	76,930	0.97	A/G	0.169	1.4E-05	-0.070	1.2E-04	-0.034	3.5E-02	**1.4E-08**	**4.8E-08**
WHRadjBMI														
4:EM	rs30000 Ł	5:55803533	*MAP3K1*	71,424	0.27	G/A	0.002	7.8E-01	0.031	**3.7E-08**	0.040	**1.7E-10**	*1.6E-04*	2.7E-07
4:AM	rs459193 Ł	5:55806751		80,852	0.27	A/G	0.004	5.0E-01	0.034	**4.1E-10**	0.043	**2.3E-13**	*6.8E-05*	**2.2E-09**
4:AM	rs2071449 Ł	12:54428011	*HOXC4-HOXC6*	70,868	0.37	A/C	0.003	6.0E-01	0.026	1.0E-06	0.034	**9.1E-09**	*1.1E-03*	5.7E-06
4:EM	rs754133 Ł	12:54418920		71,136	0.36	A/G	0.003	6.2E-01	0.026	8.2E-07	0.034	**3.0E-09**	*1.1E-03*	4.0E-06
4:AM	rs12608504 Ł	19:18389135	*JUND*	80,087	0.37	A/G	0.006	2.6E-01	0.025	5.0E-07	0.032	**4.7E-09**	*5.5E-03*	1.8E-06

Abbreviations: E/O- effect/other, EAF- effect allele frequency, SE- standard error; Chr- chromosome; Pos- position (bp); adj- adjusted for smoking; int- interaction; App- Approach.

REFERENCES

1. Heid, I.M. *et al.* Meta-analysis identifies 13 new loci associated with waist-hip ratio and reveals sexual dimorphism in the genetic basis of fat distribution. *Nat Genet* **42**, 949-60 (2010).
2. Speliotes, E.K. *et al.* Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nat Genet* **42**, 937-48 (2010).
3. Willer, C.J. *et al.* Six new loci associated with body mass index highlight a neuronal influence on body weight regulation. *Nat Genet* **41**, 25-34 (2009).
4. Loos, R.J. *et al.* Common variants near MC4R are associated with fat mass, weight and risk of obesity. *Nat Genet* **40**, 768-75 (2008).
5. Lindgren, C.M. *et al.* Genome-wide association scan meta-analysis identifies three Loci influencing adiposity and fat distribution. *PLoS Genet* **5**, e1000508 (2009).
6. Locke, A.E. *et al.* Genetic studies of body mass index yield new insights for obesity biology. *Nature* **518**, 197-206 (2015).
7. Shungin, D. *et al.* New genetic loci link adipose and insulin biology to body fat distribution. *Nature* **518**, 187-96 (2015).
8. Berndt, S.I. *et al.* Genome-wide meta-analysis identifies 11 new loci for anthropometric traits and provides insights into genetic architecture. *Nat Genet* **45**, 501-12 (2013).
9. Taylor, A.E. *et al.* Stratification by smoking status reveals an association of CHRNA5-A3-B4 genotype with body mass index in never smokers. *PLoS Genet* **10**, e1004799 (2014).
10. Koster-Rasmussen, R. *et al.* Back on track-Smoking cessation and weight changes over 9years in a community-based cohort study. *Prev Med* **81**, 320-5 (2015).
11. Pistelli, F., Aquilini, F. & Carrozzi, L. Weight gain after smoking cessation. *Monaldi Arch Chest Dis* **71**, 81-7 (2009).
12. Flegal, K.M., Troiano, R.P., Pamuk, E.R., Kuczmarski, R.J. & Campbell, S.M. The influence of smoking cessation on the prevalence of overweight in the United States. *N Engl J Med* **333**, 1165-70 (1995).
13. Morris, R.W. *et al.* Heavier smoking may lead to a relative increase in waist circumference: evidence for a causal relationship from a Mendelian randomisation meta-analysis. The CARTA consortium. *BMJ Open* **5**, e008808 (2015).
14. Barrett-Connor, E. & Khaw, K.T. Cigarette smoking and increased central adiposity. *Ann Intern Med* **111**, 783-7 (1989).
15. Kim, J.H. *et al.* Cigarette smoking increases abdominal and visceral obesity but not overall fatness: an observational study. *PLoS One* **7**, e45815 (2012).
16. Owen-Smith, V. & Hannaford, P.C. Stopping smoking and body weight in women living in the United Kingdom. *Br J Gen Pract* **49**, 989-90 (1999).
17. Lahmann, P.H., Lissner, L., Gullberg, B. & Berglund, G. Sociodemographic factors associated with long-term weight gain, current body fatness and central adiposity in Swedish women. *Int J Obes Relat Metab Disord* **24**, 685-94 (2000).
18. Martinez, J.A., Kearney, J.M., Kafatos, A., Paquet, S. & Martinez-Gonzalez, M.A. Variables independently associated with self-reported obesity in the European Union. *Public Health Nutr* **2**, 125-33 (1999).
19. Dare, S., Mackay, D.F. & Pell, J.P. Relationship between smoking and obesity: a cross-sectional study of 499,504 middle-aged adults in the UK general population. *PLoS One* **10**, e0123579 (2015).
20. Chiolero, A., Jacot-Sadowski, I., Faeh, D., Paccaud, F. & Cornuz, J. Association of cigarettes smoked daily with obesity in a general adult population. *Obesity (Silver Spring)* **15**, 1311-8 (2007).

21. Clair, C. *et al.* Dose-dependent positive association between cigarette smoking, abdominal obesity and body fat: cross-sectional data from a population-based survey. *BMC Public Health* **11**, 23 (2011).
22. Scherr, A. *et al.* Predictors of marked weight gain in a population of health care and industrial workers following smoking cessation. *BMC Public Health* **15**, 520 (2015).
23. Eliasson, B. & Smith, U. Leptin levels in smokers and long-term users of nicotine gum. *Eur J Clin Invest* **29**, 145-52 (1999).
24. Pisinger, C. & Jorgensen, T. Waist circumference and weight following smoking cessation in a general population: the Inter99 study. *Prev Med* **44**, 290-5 (2007).
25. Munafo, M.R., Johnstone, E.C., Murphy, M.F. & Aveyard, P. Lack of association of DRD2 rs1800497 (Taq1A) polymorphism with smoking cessation in a nicotine replacement therapy randomized trial. *Nicotine Tob Res* **11**, 404-7 (2009).
26. Weekley, C.K., 3rd, Klesges, R.C. & Reylea, G. Smoking as a weight-control strategy and its relationship to smoking status. *Addict Behav* **17**, 259-71 (1992).
27. Winter, A.L., de Guia, N.A., Ferrence, R. & Cohen, J.E. The relationship between body weight perceptions, weight control behaviours and smoking status among adolescents. *Can J Public Health* **93**, 362-5 (2002).
28. Pirie, P.L., Murray, D.M. & Luepker, R.V. Gender differences in cigarette smoking and quitting in a cohort of young adults. *Am J Public Health* **81**, 324-7 (1991).
29. Nicklas, B.J., Tomoyasu, N., Muir, J. & Goldberg, A.P. Effects of cigarette smoking and its cessation on body weight and plasma leptin levels. *Metabolism* **48**, 804-8 (1999).
30. Aschard, H., Hancock, D.B., London, S.J. & Kraft, P. Genome-wide meta-analysis of joint tests for genetic and gene-environment interaction effects. *Hum Hered* **70**, 292-300 (2010).
31. Randall, J.C. *et al.* Sex-stratified genome-wide association studies including 270,000 individuals show sexual dimorphism in genetic loci for anthropometric traits. *PLoS Genet* **9**, e1003500 (2013).
32. Yang, J., Lee, S.H., Goddard, M.E. & Visscher, P.M. GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet* **88**, 76-82 (2011).
33. van Setten, J. *et al.* Genome-wide association study of coronary and aortic calcification implicates risk loci for coronary artery disease and myocardial infarction. *Atherosclerosis* **228**, 400-5 (2013).
34. Winslow, U.C., Rode, L. & Nordestgaard, B.G. High tobacco consumption lowers body weight: a Mendelian randomization study of the Copenhagen General Population Study. *Int J Epidemiol* **44**, 540-50 (2015).
35. Morel, C. *et al.* Nicotine consumption is regulated by a human polymorphism in dopamine neurons. *Mol Psychiatry* **19**, 930-6 (2014).
36. Antolin-Fontes, B., Ables, J.L., Gorlich, A. & Ibanez-Tallon, I. The habenulo-interpeduncular pathway in nicotine aversion and withdrawal. *Neuropharmacology* **96**, 213-22 (2015).
37. Picciotto, M.R. & Mineur, Y.S. Molecules and circuits involved in nicotine addiction: The many faces of smoking. *Neuropharmacology* **76 Pt B**, 545-53 (2014).
38. Thorgeirsson, T.E. *et al.* A common biological basis of obesity and nicotine addiction. *Transl Psychiatry* **3**, e308 (2013).
39. Zoli, M. & Picciotto, M.R. Nicotinic regulation of energy homeostasis. *Nicotine Tob Res* **14**, 1270-90 (2012).
40. van der Vaart, H., Postma, D.S., Timens, W. & ten Hacken, N.H. Acute effects of cigarette smoke on inflammation and oxidative stress: a review. *Thorax* **59**, 713-21 (2004).
41. Aroor, A.R. & DeMarco, V.G. Oxidative stress and obesity: the chicken or the egg? *Diabetes* **63**, 2216-8 (2014).

42. Youn, J.Y. *et al.* Role of vascular oxidative stress in obesity and metabolic syndrome. *Diabetes* **63**, 2344-55 (2014).
43. Suidan, J.S. & Abboud, J.M. Treatment of the fetus in utero. *Middle East J Anaesthesiol* **11**, 163-79 (1991).
44. Westra, H.J. *et al.* Systematic identification of trans eQTLs as putative drivers of known disease associations. *Nat Genet* **45**, 1238-43 (2013).
45. Yang, H.W. *et al.* Genomic structure and mutational analysis of the human KIF1B gene which is homozygously deleted in neuroblastoma at chromosome 1p36.2. *Oncogene* **20**, 5075-83 (2001).
46. Zhao, C. *et al.* Charcot-Marie-Tooth disease type 2A caused by mutation in a microtubule motor KIF1Bbeta. *Cell* **105**, 587-97 (2001).
47. Higgs, S. Memory and its role in appetite regulation. *Physiol Behav* **85**, 67-72 (2005).
48. Davidson, T.L. *et al.* Contributions of the hippocampus and medial prefrontal cortex to energy and body weight regulation. *Hippocampus* **19**, 235-52 (2009).
49. Borcherding, D.C. *et al.* Dopamine receptors in human adipocytes: expression and functions. *PLoS One* **6**, e25537 (2011).
50. Lindberg, F.P. *et al.* Decreased resistance to bacterial infection and granulocyte defects in IAP-deficient mice. *Science* **274**, 795-8 (1996).
51. Finley, M.J., Clark, K.A., Alferiev, I.S., Levy, R.J. & Stachelek, S.J. Intracellular signaling mechanisms associated with CD47 modified surfaces. *Biomaterials* **34**, 8640-9 (2013).
52. Wiewiora, M., Piecuch, J., Sedek, L., Mazur, B. & Sosada, K. The effects of obesity on CD47 expression in erythrocytes. *Cytometry B Clin Cytom* (2015).
53. Maimaitiyiming, H., Norman, H., Zhou, Q. & Wang, S. CD47 deficiency protects mice from diet-induced obesity and improves whole body glucose tolerance and insulin sensitivity. *Sci Rep* **5**, 8846 (2015).
54. Forstermann, U. & Munzel, T. Endothelial nitric oxide synthase in vascular disease: from marvel to menace. *Circulation* **113**, 1708-14 (2006).
55. Holvoet, P. Stress in obesity and associated metabolic and cardiovascular disorders. *Scientifica (Cairo)* **2012**, 205027 (2012).
56. Furukawa, S. *et al.* Increased oxidative stress in obesity and its impact on metabolic syndrome. *J Clin Invest* **114**, 1752-61 (2004).
57. Gewinner, C. *et al.* Evidence that inositol polyphosphate 4-phosphatase type II is a tumor suppressor that inhibits PI3K signaling. *Cancer Cell* **16**, 115-25 (2009).
58. Gasser, J.A. *et al.* SGK3 mediates INPP4B-dependent PI3K signaling in breast cancer. *Mol Cell* **56**, 595-607 (2014).
59. Manning, B.D. & Cantley, L.C. AKT/PKB signaling: navigating downstream. *Cell* **129**, 1261-74 (2007).
60. Micu, I. *et al.* NMDA receptors mediate calcium accumulation in myelin during chemical ischaemia. *Nature* **439**, 988-92 (2006).
61. Zhong, H.J. *et al.* Functional polymorphisms of the glutamate receptor N-methyl D-aspartate 2A gene are associated with heroin addiction. *Genet Mol Res* **13**, 8714-21 (2014).
62. Turner, S.J. *et al.* GRIN2A: an aptly named gene for speech dysfunction. *Neurology* **84**, 586-93 (2015).
63. Liu, R. *et al.* Correlation of functional GRIN2A gene promoter polymorphisms with schizophrenia and serum D-serine levels. *Gene* **568**, 25-30 (2015).
64. Leuba, G. *et al.* Pathological reorganization of NMDA receptors subunits and postsynaptic protein PSD-95 distribution in Alzheimer's disease. *Curr Alzheimer Res* **11**, 86-96 (2014).
65. DeVries, S.P. & Patel, A.D. Two patients with a GRIN2A mutation and childhood-onset epilepsy. *Pediatr Neurol* **49**, 482-5 (2013).

66. Lemke, J.R. *et al.* Mutations in GRIN2A cause idiopathic focal epilepsy with rolandic spikes. *Nat Genet* **45**, 1067-72 (2013).
67. Carvill, G.L. *et al.* GRIN2A mutations cause epilepsy-aphasia spectrum disorders. *Nat Genet* **45**, 1073-6 (2013).
68. Rezvani, K., Teng, Y., Shim, D. & De Biasi, M. Nicotine regulates multiple synaptic proteins by inhibiting proteasomal activity. *J Neurosci* **27**, 10508-19 (2007).
69. Chen, H.S. & Lipton, S.A. Pharmacological implications of two distinct mechanisms of interaction of memantine with N-methyl-D-aspartate-gated channels. *J Pharmacol Exp Ther* **314**, 961-71 (2005).
70. Maler, J.M. *et al.* Memantine inhibits ethanol-induced NMDA receptor up-regulation in rat hippocampal neurons. *Brain Res* **1052**, 156-62 (2005).
71. Bresink, I. *et al.* Effects of memantine on recombinant rat NMDA receptors expressed in HEK 293 cells. *Br J Pharmacol* **119**, 195-204 (1996).
72. Rogawski, M.A. & Wenk, G.L. The neuropharmacological basis for the use of memantine in the treatment of Alzheimer's disease. *CNS Drug Rev* **9**, 275-308 (2003).
73. Xu, K. & Lipsky, R.H. Repeated ketamine administration alters N-methyl-D-aspartic acid receptor subunit gene expression: implication of genetic vulnerability for ketamine abuse and ketamine psychosis in humans. *Exp Biol Med (Maywood)* **240**, 145-55 (2015).
74. Linden, R. *et al.* Physiology of the prion protein. *Physiol Rev* **88**, 673-728 (2008).
75. Redecke, L. *et al.* Structural characterization of beta-sheeted oligomers formed on the pathway of oxidative prion protein aggregation in vitro. *J Struct Biol* **157**, 308-20 (2007).
76. United States. Public Health Service. Office of the Surgeon General. *How tobacco smoke causes disease : the biology and behavioral basis for smoking-attributable disease : a report of the Surgeon General*, xv, 704 p (U.S. Dept. of Health and Human Services, Public Health Service, Rockville, MD; Washington, DC, 2010).
77. Bernhard, D., Rossmann, A. & Wick, G. Metals in cigarette smoke. *IUBMB Life* **57**, 805-9 (2005).
78. Savini, I., Rossi, A., Pierro, C., Avigliano, L. & Catani, M.V. SVCT1 and SVCT2: key proteins for vitamin C uptake. *Amino Acids* **34**, 347-55 (2008).
79. Babaev, V.R. *et al.* Combined vitamin C and vitamin E deficiency worsens early atherosclerosis in apolipoprotein E-deficient mice. *Arterioscler Thromb Vasc Biol* **30**, 1751-7 (2010).
80. Catania, A.S., Barros, C.R. & Ferreira, S.R. [Vitamins and minerals with antioxidant properties and cardiometabolic risk: controversies and perspectives]. *Arq Bras Endocrinol Metabol* **53**, 550-9 (2009).
81. Bornstein, S.R. *et al.* Impaired adrenal catecholamine system function in mice with deficiency of the ascorbic acid transporter (SVCT2). *FASEB J* **17**, 1928-30 (2003).
82. Uhl, G.R., Drgon, T., Li, C.Y., Johnson, C. & Liu, Q.R. Smoking and smoking cessation in disadvantaged women: assessing genetic contributions. *Drug Alcohol Depend* **104 Suppl 1**, S58-63 (2009).
83. Rose, J.E., Behm, F.M., Drgon, T., Johnson, C. & Uhl, G.R. Personalized smoking cessation: interactions between nicotine dose, dependence and quit-success genotype score. *Mol Med* **16**, 247-53 (2010).
84. Sasaki, Y., Araki, T. & Milbrandt, J. Stimulation of nicotinamide adenine dinucleotide biosynthetic pathways delays axonal degeneration after axotomy. *J Neurosci* **26**, 8484-91 (2006).
85. Revollo, J.R., Grimm, A.A. & Imai, S. The NAD biosynthesis pathway mediated by nicotinamide phosphoribosyltransferase regulates Sir2 activity in mammalian cells. *J Biol Chem* **279**, 50754-63 (2004).
86. Jayaram, H.N., Kusumanchi, P. & Yalowitz, J.A. NMNAT expression and its relation to NAD metabolism. *Curr Med Chem* **18**, 1962-72 (2011).

87. Fang, C., Decker, H. & Banker, G. Axonal transport plays a crucial role in mediating the axon-protective effects of NmNAT. *Neurobiol Dis* **68**, 78-90 (2014).
88. Press, C. & Milbrandt, J. Nmnat delays axonal degeneration caused by mitochondrial and oxidative stress. *J Neurosci* **28**, 4861-71 (2008).
89. Warnatz, H.J. *et al.* The BTB and CNC homology 1 (BACH1) target genes are involved in the oxidative stress response and in control of the cell cycle. *J Biol Chem* **286**, 23521-32 (2011).
90. Liu, J.Z. *et al.* Meta-analysis and imputation refines the association of 15q25 with smoking quantity. *Nat Genet* **42**, 436-40 (2010).
91. Burdett, T. *et al.* The NHGRI-EBI Catalog of published genome-wide association studies. v1.0 edn Vol. 2015 (2015).
92. Hindorff, L.A. *et al.* Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc Natl Acad Sci U S A* **106**, 9362-7 (2009).
93. Day, F.R., Loh, P.R., Scott, R.A., Ong, K.K. & Perry, J.R. A Robust Example of Collider Bias in a Genetic Association Study. *Am J Hum Genet* **98**, 392-3 (2016).
94. Aschard, H., Vilhjalmsson, B.J., Joshi, A.D., Price, A.L. & Kraft, P. Adjusting for heritable covariates can bias effect estimates in genome-wide association studies. *Am J Hum Genet* **96**, 329-39 (2015).
95. Chiolero, A., Faeh, D., Paccaud, F. & Cornuz, J. Consequences of smoking for body weight, body fat distribution, and insulin resistance. *Am J Clin Nutr* **87**, 801-9 (2008).
96. Wei, Y. *et al.* Chronic exposure to air pollution particles increases the risk of obesity and metabolic syndrome: findings from a natural experiment in Beijing. *FASEB J* (2016).
97. Kondo, M. *et al.* Enhanced oxidative stress is associated with sleep-disordered breathing and obesity in patients with heart failure. *Int J Cardiol* **209**, 133-135 (2016).
98. Arany, I., Hall, S., Reed, D.K., Reed, C.T. & Dixit, M. Nicotine Enhances High-Fat Diet-Induced Oxidative Stress in the Kidney. *Nicotine Tob Res* (2016).
99. Weldy, C.S., Liu, Y., Liggitt, H.D. & Chin, M.T. In utero exposure to diesel exhaust air pollution promotes adverse intrauterine conditions, resulting in weight gain, altered blood pressure, and increased susceptibility to heart failure in adult mice. *PLoS One* **9**, e88582 (2014).
100. Voight, B.F. *et al.* The metabochip, a custom genotyping array for genetic studies of metabolic, cardiovascular, and anthropometric traits. *PLoS Genet* **8**, e1002793 (2012).
101. Li, Y., Willer, C., Sanna, S. & Abecasis, G. Genotype imputation. *Annu Rev Genomics Hum Genet* **10**, 387-406 (2009).
102. Howie, B.N., Donnelly, P. & Marchini, J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet* **5**, e1000529 (2009).
103. Servin, B. & Stephens, M. Imputation-based analysis of association studies: candidate regions and quantitative traits. *PLoS Genet* **3**, e114 (2007).
104. Browning, S.R. & Browning, B.L. Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *Am J Hum Genet* **81**, 1084-97 (2007).
105. Li, Y., Willer, C.J., Ding, J., Scheet, P. & Abecasis, G.R. MaCH: using sequence and genotype data to estimate haplotypes and unobserved genotypes. *Genet Epidemiol* **34**, 816-34 (2010).
106. Marchini, J., Howie, B., Myers, S., McVean, G. & Donnelly, P. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat Genet* **39**, 906-13 (2007).
107. Aulchenko, Y.S., Struchalin, M.V. & van Duijn, C.M. ProbABEL package for genome-wide association analysis of imputed data. *BMC Bioinformatics* **11**, 134 (2010).
108. Aulchenko, Y.S., Ripke, S., Isaacs, A. & van Duijn, C.M. GenABEL: an R library for genome-wide association analysis. *Bioinformatics* **23**, 1294-6 (2007).

109. Abecasis, G.R. & Wigginton, J.E. Handling marker-marker linkage disequilibrium: pedigree analysis with clustered markers. *Am J Hum Genet* **77**, 754-67 (2005).
110. Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* **81**, 559-75 (2007).
111. Kutalik, Z. *et al.* Methods for testing association between uncertain genotypes and quantitative traits. *Biostatistics* **12**, 1-17 (2011).
112. Winkler, T.W. *et al.* Quality control and conduct of genome-wide association meta-analyses. *Nat Protoc* **9**, 1192-212 (2014).
113. Devlin, B. & Roeder, K. Genomic control for association studies. *Biometrics* **55**, 997-1004 (1999).
114. Willer, C.J., Li, Y. & Abecasis, G.R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190-1 (2010).
115. Winkler, T.W. *et al.* The Influence of Age and Sex on Genetic Associations with Adult Body Size and Shape: A Large-Scale Genome-Wide Interaction Study. *PLoS Genet* **11**, e1005378 (2015).
116. Winkler, T.W. *et al.* EasyStrata: evaluation and visualization of stratified genome-wide association meta-analysis data. *Bioinformatics* **31**, 259-61 (2015).
117. Pruim, R.J. *et al.* LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics* **26**, 2336-7 (2010).
118. Abecasis, G.R. *et al.* An integrated map of genetic variation from 1,092 human genomes. *Nature* **491**, 56-65 (2012).
119. Abecasis, G.R. *et al.* A map of human genome variation from population-scale sequencing. *Nature* **467**, 1061-73 (2010).
120. Wood, A.R. *et al.* Defining the role of common variation in the genomic and biological architecture of adult human height. *Nat Genet* **46**, 1173-86 (2014).
121. Ward, L.D. & Kellis, M. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Res* **40**, D930-4 (2012).
122. Kuhn, R.M., Haussler, D. & Kent, W.J. The UCSC genome browser and associated tools. *Brief Bioinform* **14**, 144-61 (2013).
123. Consortium, G.T. The Genotype-Tissue Expression (GTEx) project. *Nat Genet* **45**, 580-5 (2013).
124. Boyle, A.P. *et al.* Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res* **22**, 1790-7 (2012).
125. Zhang, X. *et al.* Synthesis of 53 tissue and cell line expression QTL datasets reveals master eQTLs. *BMC Genomics* **15**, 532 (2014).
126. Irizarry, R.A. *et al.* Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* **4**, 249-64 (2003).
127. Kutalik, Z., Whittaker, J., Waterworth, D., Beckmann, J.S. & Bergmann, S. Novel method to estimate the phenotypic variation explained by genome-wide association studies reveals large fraction of the missing heritability. *Genet Epidemiol* **35**, 341-9 (2011).
128. Delaneau, O., Zagury, J.F. & Marchini, J. Improved whole-chromosome phasing for disease and population genetic studies. *Nat Methods* **10**, 5-6 (2013).
129. Howie, B., Fuchsberger, C., Stephens, M., Marchini, J. & Abecasis, G.R. Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. *Nat Genet* **44**, 955-9 (2012).

SUPPLEMENTARY NOTE 1. **Look-up of previously identified loci in our data set**

To fully explore the efficacy of accounting for smoking in GWAS of adiposity traits, we conducted a look-up in our data of recently published SNP associations with BMI, WHRadjBMI, and WCadjBMI identified in well-powered GWAS meta-analyses that did not account for SMK status[1,2]. Although our sample size was as little as one third of previously published GWAS[1,2], the majority of these loci (92% for BMI, 97% for WCadjBMI, and 92% for WHRadjBMI) reached Bonferroni corrected significant for at least one of the three Approaches in the current study.

All previously identified 97 BMI-associated SNPs were nominally significant (P<0.05) in Approach 1 (SNPadjSMK) for BMI including the sex-specific loci, 95 of the 97 for Approach 2 (SNPjoint), and seven for Approach 3 (SNPint). A total of 86 loci reached Bonferroni-corrected significance ($P<5.15x10^{-4}$) for Approach 1, 85 for Approach 2, and none for Approach 3. Finally, 41 loci from Approach 1 and 39 of the 97 from Approach 2 reached genome-wide significance (GWS, $P<5x10^{-8}$) (44 in total, 45%) (**Supplementary Table 11**). Of the 97 previously identified main effects loci for BMI, 3 of these were genome-wide significant GWS for women-only, 3 for men-only and the remaining in the sex-combined analysis in the previous publication. It is also worth noting that we report results for the All Ancestries meta-analysis, as this was our primary meta-analysis data-set; however, Locke et al. (2015) considered their European-descent only meta-analysis their primary data-set.

Of the 77 previously-identified WCadjBMI loci, 3 of these were GWS for women-only, 3 for men-only and the remaining in the sex-combined analysis as reported in Shungin et al[2]. Of these, 75 were nominally significant for Approachch 1 (SNPadjSMK) and Approach 2 (SNPjoint), and 5 for Approach 3 (SNPint). A total of 73 were Bonferroni-corrected significant ($P<6.49x10^{-4}$) for Approach 1 and 2; with 41 and 40 reaching GWS, respectively (43 non-overlapping, 56%) (**Supplementary Table 12**).

Eleven of the 68 previously published WHRadjBMI SNPs were associated in the women-only analyses in the previous investigation[2]. Of the 68 variants, 64 were nominally significant for Approach 1 (SNPadjSMK), 59 for Approach 2 (SNPjoint), and 10 for Approach 3 (SNPint). A total of 61 were Bonferroni-corrected significant ($P<6.49x10^{-4}$) for Approach 1 and 38 for Approach 2; with 36 and 8 reaching GWS, respectively (36 in total, 53%) (**Supplementary Table 13**).

In summary, we replicated all previously-identified BMI loci using one or more of our approaches (P<0.05 and concordant direction of effect), but did not replicate all previously-identified loci for WCadjBMI and WHRadjBMI in our current analyses. It is unclear if the lack of replication of previous findings is due to smaller sample size, patterns of linkage disequilibrium in our all ancestries sample, the adjustment of smoking status in the current discovery analysis, or even a combination of these factors.

SUPPLEMENTARY NOTE 2. Summary of literature search on genes nearest to the 21 novel loci and all GxSMK interaction loci.

We used SNIPPER (http://csg.sph.umich.edu/boehnke/snipper/) to identify potential biological functions of genes ±500kb of our novel association signals and those from Approach 3 (SNPint) for further investigation, and present a summary of those findings in this section (**Online Methods**).

Body Mass Index (BMI)

rs2481665 (*INADL*): There are seven genes within the 500kb region of the lead SNP rs2481665 on chromosome 1. These genes are *INADL, L1TD1, KANK4, USP1, DOCK7, TM2D1,* and *ANGPTL3*. The lead SNP is in intron (#15) of the *INADL* (InaD-Like) gene. *INADL* encodes the protein Palsi1-Associated Tight Junction (PATJ), which helps regulate the formation of tight junctions, and is involved in the processes of cell polarization and directional migration of epithelial cells[3,4]. A GWAS study (n= 815) designed to identify variants associated with childhood obesity in the Hispanic population, found near genome-wide significant associations between the exonic, non-synonymous SNP rs1056513 in *INADL* (204 kb downstream from our lead SNP) and the following fat distribution traits: weight [kg] (EAF[effect allele frequency]: 0.031, p-value: 1.18×10^{-07}); BMI [kg/m^2] (EAF: 0.021, p-value: 8.34×10^{-06}); fat mass [kg] (EAF: 0.035, p-value: 1.59×10^{-07}); trunk fat mass [kg] (EAF: 0.035, p-value: 2.36×10^{-07}); fat free mass [kg] (EAF: 0.034, p-value: 2.80×10^{-07}) and hip circumference (EAF: 0.022, p-value: 2.47×10^{-6}).[5] The SNP rs1056513 accounted for 3% of the variance in body weight and body composition[5]. However, this SNP is not in LD with the lead SNP rs2481665 in this study (R^2<0.2).

Farther away is the *DOCK7* gene, 326 kb downstream from the lead SNP. This gene encodes a guanine nucleotide exchange factor (GEF) protein that is involved in axon formation and neuronal polarization. GWAS studies have reported the association of variants located near the *DOCK7* gene with lipid levels. A GWAS study (n= up to 18,554) conducted with individuals of European ancestry identified the association of rs1213033 with triglycerides (eaf: -0.11, 2×10^{-8})[6]. Another GWAS meta-analysis found a genome-wide significant association between rs1168013 and triglycerides in individuals of European ancestry (n=17,723; eaf: 0.035 (0.007), p-value: 6.4×10^{-8})[7]. However, authors could not replicate this finding in other study samples consisting of 37,774 Europeans and 9,665 individuals of Indian Asian ethnicity. A GWAS replication study assessing the association between 15 SNPs and blood lipid and lipoprotein concentrations in individuals of Asian descent (n=4638), found a marginal association between the variant rs10889353, located in the intronic region of *DOCK7*, and triglycerides (eaf: -0.08, p-value: 6.5×10^{-04})[8]. None of the variants from the different GWAS studies discussed above are in LD with SNP rs2481665 (R^2<0.2).

TM2D1 is another gene in the 500kb area that is 404 kb upstream from rs2481665. This gene encodes a beta-amyloid peptide-binding protein (BBP), which is involved in neural death and in the decrease of cognitive skills that occurs in Alzheimer's disease**.** This protein may be targeted by the beta-amyloid peptide which has been linked to the formation of plaques resulting in neurotoxicity in Alzheimer's disease[9]. The APP, the precursor of beta-amyloid peptide, is expressed in adipose tissue and its expression is up-regulated in obesity[10,11].

ANGPTL3 (Angiopoietin-Like 3) is 469 kb upstream from the lead SNP, and upstream of the *DOCK7* gene. *ANGPTL3* encodes a protein that plays a role in angiogenesis. This protein is expressed mostly in the liver. Mutations in this gene lead to the disease familial hypobetalipoproteinemia type 2 (*FHBL2*), which causes low levels of apolipoprotein B (apoB), total cholesterol, low-density lipoprotein (LDL) cholesterol and high density lipoprotein cholesterol [12]. Several genetic association studies suggest that ANGPTL3 has a role in regulating plasma lipoprotein metabolism[6,8,13,14]. A few single-nucleotide polymorphisms, near the *ANGPTL3* gene, have been associated with lower triglyceride: rs1213033, rs213192, rs12042319[6]. One of these, rs1213033, is also near the *DOCK7* gene[6].

There are several nearby genes with no documented role in adiposity or related cardiometabolic traits. Including, *L1TD1* (Line-1 type transposase domain containing 1) located 66 kb upstream from the lead SNP. *L1TD1* encodes the protein ES Cell-Associated Protein 11, a RNA-binding protein that plays a role in maintaining the pluripotency of stem cells, and in the proliferation of cancer cells[15,16]. Also, *KANK4* (KN

motif and ankyrin repeat domains 4) is a gene located 107 kb downstream from our SNP of interest. It encodes the protein Ankyrin Repeat Domain 38, a member of the Kank family of proteins, which are involved in the control of cytoskeleton microfilaments by regulating the polymerization of actin. The Kank gene is a tumor suppressor in renal cell carcinoma[17]. *USP1*, 307 kb upstream from rs2481665, encodes a protein that cleaves ubiquitin, a peptide that is added to proteins to signal them for degradation, or modification of their cellular location or enzymatic activity.

The intronic rs2481665 variant does not seem to have a functional role (Score 4 in RegulomeDB[18]). Two eQTLs were found for rs2481665 (Gene: L1TD1, p-value: 2.1×10^{-7}, EAF: -0.73, tissue: brain-cerebellum) and (Gene: INALD, p-value: 4.0×10^{-6}, EAF: 0.29, tissue: heart-atrial appendage).

rs10929925 (*LOC400940*): *LOC400940* and *SOX11* are the two genes on Chr2 that are within 500 kb of the lead SNP rs10929925. SNP rs10929925 is downstream of *LOC400940*, the nearest gene, a non-coding RNA gene that remains uncharacterized. The variant is also 314 kb downstream from *SOX11*, a gene without introns that encodes a transcription factor that is part of the SOX (SRY-related HMG-box) family. This family of transcription factors is involved with processes that regulate embryonic development and cell fate[19]. One study has proposed that *SOX11* has a role in brain development after observing that mutations in the gene may lead to microcephaly, developmental delays and other features found in mild Coffin-Siris Syndrome, a genetic disorder that causes developmental delays[20]. A recent GWAS meta-analysis study of fat distribution, which included 224,459 individuals of European and non-European ancestry, identified a genome wide significant association ($p=4.5 \times 10^{-8}$) between rs10929925 and hip circumference unadjusted for BMI [2]. Based on a literature review, the study identified *SOX11* as the best candidate gene for rs10929925.[2]

There is no available information regarding the potential regulatory role of the lead SNP (RegulomeDB[18]). But there is evidence of an eQTL, although it does not reach 5% FDR (Gene: *SOX11*, P-value: 8.7×10^{-6}, Effect size: 0.39, Tissue: thyroid). In brain tissue, the SNP altered the TATA box motif of the DIx3 gene a homeodomain gene (HaploReg[21]).

rs6794880 (*SRRM1P2):* The 500kb region around the lead SNP, rs6794880, does not show the presence of any protein coding genes. The nearest genomic feature to rs6794880 is SRRM1P2, a pseudogene, named the serine/arginine repetitive matrix 1 pseudogene 2. Upstream rs6794880 is *LINC00971*, a long intergenic non-protein coding RNA gene that remains uncharacterized.

There is no evidence that the lead SNP rs6794880 has a functional/regulatory role (Score 6 in RegulomeDB[18]) in the genome. Additionally, there are no reports of eQTLs for this variant.

rs12629427 *(EPHA3)*: There is only one gene found within 500kb of the peak signal, rs12629427. *EPHA3* (EPH receptor A3) is 11kb downstream from rs12629427, and is a member of the ephrin receptor subfamily of the protein-tyrosine kinase family. EPH and EPH-related receptors have been implicated in mediating developmental events, particularly in the nervous system. This gene encodes a protein that binds ephrin-A ligands. *EPHA3* has been implicated in the pathogenesis of lung cancer[22-26]. The SNP rs12629427 has a score of 6 in RegulomeDB[18] (minimal binding evidence). No significant eQTLs were found for rs12629427 and no GWAS hits were identified within the 1MB region of the lead SNP.

rs2173039 (*EPHA3*): There is only one gene found within 500kb of rs2173039, which is 14.5kb upstream from EPHA3 (EPH receptor A3). See rs12629427 above.

rs13069244 (*CCDC39*): A total of 4 genes are found within 500kb of the lead marker, rs13069244. *CCDC39* (coiled-coil domain containing 39) is located 43.88kb downstream from the lead marker and encodes a protein involved in the motility of cilia and flagella. Defects in this gene cause primary ciliary dyskinesia type 14. Lung disease was worse in those with IDA/CA/MTD ultrastructural defects, most of whom had biallelic mutations in *CCDC39*[27]. *FXR1* (fragile X mental retardation, autosomal homolog 1) is located 189kb downstream from rs13069244, and codes for an RNA binding protein that shuttles between the nucleus and cytoplasm, and is associated with polyribosomes, predominantly with the 60S ribosomal subunit. Deregulation of FXR protein 1 by the lipodystrophic lamin A p.R482W mutation elicits a myogenic gene expression program in preadipocytes[28]. *DNAJC19* (DnaJ (Hsp40) homolog, subfamily C, member 19), located 260kb upstream from our lead marker, encodes a protein involved in the ATP-dependent transport of transit peptide-containing proteins from the inner cell membrane to the mitochondrial matrix. Defects in this gene are a cause of 3-methylglutaconic aciduria type 5 (MGA5), also known as dilated cardiomyopathy with ataxia (DCMA)[29-31]. The loss of DNAJC19/PHB complexes affects cardiolipin acylation and leads to the accumulation of cardiolipin species with altered acyl chains[32]. There is no evidence that rs13069244 has a functional/regulatory role (RegulomeDB[18] Score 6: minimal binding evidence) in the genome. No GWAS hits were identified within the 1Mb region of rs13069244 and no report of eQTL for the variant.

rs336396 (*INPP4B*): There are two genes found within 500kb of rs336396. The SNP lies within *INPP4B* (inositol polyphosphate-4-phosphatase, type II, 105kDa), which encodes inositol polyphosphate 4-phosphatase type II, one of the enzymes involved in phosphatidylinositol signaling pathways. INPP4B has been identified as a tumor suppressor by negatively regulating normal and malignant cell proliferation through regulation of the PI3K/Akt signaling pathway[33,34]. Different residues within the catalytic site of INPP4B are responsible for activity with lipid and protein substrates[35]. *IL15* (interleukin 15) is located 407kb upstream of rs336396. *IL15* encodes a cytokine that regulates T and natural killer (NK) cell activation and proliferation. This cytokine may act as an antagonist to IL2, which binds common hematopoietin receptor subunits, and may compete for the same receptor. This cytokine induces the activation of JAK kinases, as well as the phosphorylation and activation of transcription activators STAT3, STAT5, and STAT6. Murine models show that this cytokine may increase expression of apoptosis inhibitor BCL2L1/BCL-x(L), possibly through the transcription activation activity of STAT6, and thus prevent apoptosis. Cigarette smoke compromises IL-15 production – and as a result NK cell function – which could link to the higher incidence of cancers or viral infections observed among smokers[36]. A group of SNPs, upstream from *IL15*, were associated with both smoking status and quantity of cigarette consumption[37]. No data was provided for rs336396 by RegulomeDB[18]. No GWAS hits were identified within the 1Mb region of rs336396 and no report of an eQTL for the variant.

rs12902602 (*CHRNA5-CHRNA3-CHRNB4*): A total of 10 genes are found within 500kb of rs12902602. The SNP is located 33.81kb upstream of *CHRNB4* (cholinergic receptor, nicotinic beta 4). The *CHRNA5-CHRNA3-CHRNB4* gene cluster has consistently been associated with smoking quantity and nicotine dependence[38-40], COPD, lung cancer and peripheral artery disease[39,41,42], and increased risk of death[43]. Variants of *CHRNA5-CHRNA3-CHRNB4* have also been associated with lower birth weight from smoking mothers[44], and with lower BMI in current adult smokers[45,46], but with lower BMI in never smokers[46]. The *CHRNA5-CHRNA3-CHRNB4* genes encode the nicotinic acetylcholine receptor (nAChR) subunits α3, α5 and β4 that are expressed in mammalian brain[47,48]. GWASs have also identified loci at *ADAMTS7* (ADAM metallopeptidase with thrombospondin type 1 motif 7), at 84.14 kb downstream from the leader SNP rs12902602, associated with coronary artery disease and its risk factors[49-52].

Waist Circumference adjusted for BMI (WCADJBMI):

rs17396340 (*KIF1B*). A total of 10 genes are found within 500kb of the lead marker, rs17396340, which is intronic to *KIF1B*. We highlight four genes in the region here. *KIF1B* is involved in synaptic vesicle and mitochondrial transport, and may play a critical role in the development of hepatocellular carcinoma[53]. *6PGD* codes for an oxidative carboxylase responsible for reduction of 6-phosphogluconate. Cells lacking 6PGD appear to metabolize glucose as an inhibitor to induce senescence[54]. *RBP7* is involved in carotenoid metabolism. In avian model organisms, the *RBP7* promoter is important in regulating expression of several genes in adipose tissue at later developmental stages[55]. Nicotinamide mononucleotide adenylyltransferease (*NMNAT*) reversibly catalyzes the important step in the biosynthesis of NAD from ATP and NMN. NAD and NADP are used reversibly in anabolic and catabolic reactions. NAD is necessary for cell survival in oxidative stress and DNA damage. The top SNP, rs17396340, is associated with the expression levels of ARSA (p-value of 6.0e-05) at LCL tissue in *Homo sapiens*. Human adipocytes express functional DAR (Dopamine receptors) and ARSA, suggesting a regulatory role for peripheral dopamine in adipose functions[56]. It is speculated that the propensity of some DAR-activating antipsychotics to increase weight and alter metabolic homeostasis is due to their direct action on adipose tissue. Our lead SNP is also associated with mean platelet volume[57]. From HaploReg[21], the lead SNP, rs17396340, is annotated as KIF1B in GENCODE, and is functionally annotated as intronic. This lead SNP is associated with enhancer histone marks in 9 tissues; associated with regulatory motifs at GATA and Hoxa5; and with cis-eQTLs from various tissues (cells transformed fibroblasts, muscle skeletal, lymphoblastoid EUR exonlevel, lymphoblastoid EUR genelevel, and whole blood). The RegulomeDB[18] score for the lead SNP is 4.

rs6743226 (*HDLBP*). A total of 10 genes are found within 500kb of our lead marker, rs6743226. Three, of biological interest, are mentioned here. Our lead SNP, rs6743226, is intronic to *HDLBP,* which codes for a protein that binds high density lipoprotein (HDL) that functions to regulate excess cholesterol levels in cells.

STK25 codes for a serine/threonine kinase with important functions in the Golgi apparatus. This gene has been associated with severe hypoxia[58] and pseudohypoparathyroidism, symptoms of which include short stature and obesity[59]. Significantly higher serine/threonine kinase 25 (STK25) levels were observed in the skeletal muscle of type 2 diabetic patients, compared with individuals with normal glucose tolerance[60]. The overexpression of STK25 in conditions of excess dietary fuels associates with a shift in the metabolic balance in peripheral tissues from lipid oxidation to storage, leading to a systemic insulin resistance[61].

Expression of PAS domain containing serine/threonine kinase (*PASK*) is regulated by glucose and the encoded protein plays a role in the regulation of insulin gene expression. Down regulation of this gene may play a role in type 2 diabetes[62-64]. Far2 and Stk25 are candidate genes for the HDL cholesterol locus in mice[65]. The top SNP, rs6743226, is associated with the expression of B-cell CLL/lymphoma 10 (BCL10). The protein encoded by the gene *BCL10* contains a caspase recruitment domain (CARD), and induce apoptosis and to activate NF-kappaB MALT1 and this protein are thought to synergize in the activation of NF-kappaB, and the deregulation of either of them may contribute to the same pathogenetic process that leads to the malignancy[66].

There is no GWAS signal nearby the lead SNP rs6743226. This lead SNP is associated with enhancer histone marks in 4 tissues; associated with regulatory motifs changed at Goxa and TCF12; and with eQTL from various tissues including adipose subcutaneous, lung, and muscle tissues. The RegulomeDB[18] score for the lead SNP is 6.

rs4378999 *(DOCK3)*: A total of 4 genes are found near our lead marker, rs4378999, *DOCK3*, *MANF*, *VPRBP*, and *RBM15B*. Our lead variant is intronic to *DOCK3* (dedicator of cytokinesis 3), which is highly expressed in the central nervous system and like previously identified obesity related genes, is involved in neurite outgrowth downstream of BDNF-TrkB[67]. *MANF* (mesencephalic astrocyte-derived neurotrophic factor) is an endoplasmic reticulum protein that acts to protect ER in response to cellular/organismal stress [68], for example, expression is increased in skeletal muscle of the leg in rats in response to exercise [69]. Further, recent evidence shows that *MANF* may be an important factor in the protection of pancreatic beta cells and disruption of *MANF* expression can lead to diabetes [68]. There is very little known about *VPRBP*, and *RBM15B.*

Genome-wide association studies have reported the association within 1MB region of lead SNPs for height (R^2=0.35)[70,71] and melanoma (R^2=0.48)[72]. Our lead SNP is associated with regulatory motifs changed at Cdx2; and with eQTL from various tissues including adipose subcutaneous, and muscle skeletal. The lead SNP is associated eQTL in esophagus muscularis tissue based on GTEx[73] lookup. GWAS studies have report the association within 1Mb of lead SNP for height (R^2=0.38)[71], and fibrinogen (R^2=0.41)[74]. The RegulomeDB[18] does not have data for lead SNP rs4378999.

rs7697556 (*ADAMTS3*): One gene is found within 500kb of our lead marker, rs7697556. ADAM metallopeptidase with thrombospondin type 1 motif, 3 (*ADAMTS3*) is located 80 kb upstream of our variant, rs7697556. While there is no established role for ADAMTS3 in obesity-related traits, there are a number of variants within and near this gene associated with relate anthropometric and cardiometabolic traits, including height[70,71], lipid metabolism[75], and metabolites[76]. From There is no score assigned for our lead SNP in the RegulomeDB[18].

rs10269774 (*CDK6*): A total of 10 genes are found within 500 kb of the lead marker, rs10269774. The SNP is located within an intron in cyclin-dependent kinase 6 (*CDK6*). CDK family members are important regulators of cell cycle progression. GWAS have reported associations between *CDK6* variants with height[70,71,77-81]. The *CDK6*-rs2282978 associated with height is in complete LD with our lead marker (rs10269774: R^2=1, D'=1). Also, GWAS identified associations between *CDK6* variants with white blood cell counts[82] and rheumatoid arthritis[83,84]. *CDK6* rs42041 is associated with juvenile idiopathic arthritis (JIA)[85], and patients with JIA are significantly shorter and more often overweight or obese than controls[86]. Research suggests that the microRNA-103a-3p controls proliferation and osteogenic differentiation of human adipose tissue-derived stromal cells by binding to specific target sequences in the *CDK6* mRNA 3'-untranslated region[87]. Another study in the human placental transcriptome found that *CDK6* mRNA levels correlated with offspring birth weight and birth weight percentiles[88].

rs10269774 is located in enhancer regions (H3K4Me1 and H3K27ac) with histone modification enrichment in mammary epithelial tissue and lymphoblastoid cell lines. rs10269774 was suggested to have cis-acting associations with five gamma-glutamyltransferase (GGT) family gene expression in lymphoblastoid of Yoruba population (p=6E−05[89]). Elevated serum GGT is associated with waist circumference[90,91], BMI[91], visceral fat area[91], triglyceride levels[91], metabolic syndrome[90,92], coronary artery calcification[93] and biomarkers of atherosclerosis[94], arterial stiffness[95,96], incident CVD and death[92]. rs10269774 is located near to several transcription factor binding sites (*CTCF*, *EP300*, *JUN*, *POLR2A*, *FOS, NFIC*, and *RFX5*, among others).

rs9409082 and rs9408815 (*TMEM38B*): A total of 3 genes are found within 500 kb of the lead markers rs9409082 and rs9408815. At 364 kb downstream of rs9409082 is located *TMEM38B* (transmembrane

protein 38B, 9q31.2) gene, which encodes an intracellular monovalent cation channel that functions in maintenance of intracellular calcium release. Deletions in *TMEM38B* are associated with autosomal recessive osteogenesis imperfecta[97-99]. There is evidence of genome-wide association between rs9409082 with height[70]. Also, GWAS have reported several variants in this region associated with age at menarche[100-102], which is a risk factor to develop obesity, type 2 diabetes, cardiovascular disease, breast cancer and all-cause mortality[101]. However, the reported variants for age at menarche are in low-to-moderate LD ($0.005 < R^2 < 0.68$) with our lead marker from Approach 1, rs9409082. Variants on 9q31, in low LD with rs9409082, have shown suggestive association with visceral adipose to subcutaneous adipose ratio in men ($R^2=0.161$)[103] and with a protein quantitative trait locus modulating cellular response to chemotherapy ($R^2=0.002$)[104].

At 497.6 kb downstream of rs9409082 is the *FKTN* (fukutin, 9q31.2) gene that encodes a putative transmembrane protein of the cis-Golgi compartment. *FKTN* protein may be involved in the glycosylation of alpha-dystroglycan in skeletal muscle. Mutations in *FKTN* have shown association with congenital muscular dystrophy[105,106]. No significant eQTLs were found for SNP rs9409082 (GTEx[73], SNIPPER, RegulomeDB[18], and HaploReg[21]).

rs6012558 (*ARFGEF2*): A total of 11 genes are found within 500kb +/- of our lead SNP, rs6012558, which is 6,989 bp upstream of *ARFGEF2* (ADP-ribosylation factor guanine nucleotide-exchange factor 2). *ARFGEF2*'s primary function involves intracellular trafficking. Our lead variant is 86,866 bp upstream of *PREX1* (phosphatidylinositol-3,4,5-trisphosphate-dependent Rac exchange factor 1), a gene which encodes a protein involved in intraceullar signaling, lipid and protein binding, and regulation of GTPase activity [107-109]. *PREX1* is primarily expressed in the blood leukocytes and brain[107]. Recent mouse models indicate that PREX1 may be important for the regulation of thermogenic potential of brown adipose tissue and white preadipocytes, making this gene very important for energy expenditure[110]. Additionally, rs6012558 is a significant (<5% FDR) cis-acting expression quantitative trait locus (cis-eQTL) for *ARFGEF2* (subcutaneous adipose and sigmoid colon tissues), *CSE1L* (artery, thyroid, subcutaneous adipose, esophagus mucosa, and skeletal muscle tissues), and *STAU1* (transformed fibroblast cells) (GTEx[73]). Additional evidence that this variant lies in a potentially important regulatory region includes a RegulomeDB[18] score of 4 [18], it is nearby (<500kb +/- and $R^2>0.7$) other variants that rest in active enhancers for *ARFGEF2*, other cis-eQTLs for *ARFGEF2* (monocytes, whole blood, cerebellum, and temporal cortex), *DDX27* (monocytes), *C2orf199* (monocytes), *CSE1L* (whole blood), and *PREX1* (Cerebellum and Temporal Cortex) (HaploReg[21] and UCSC Browser[111]). Our lead SNP is within 500kb +/- of several previously identified GWAS SNPs for multiple traits, the nearest of which is rs6012564 associated with tendency toward anger (distance=10kb)[112]; however, all of these are in low LD with rs6012558 ($R^2<0.3$).

rs4141488 (*GRIN2A*): There are only two genes within 500 kb +/- of our lead SNP, rs4141488, which lies 218 kb downstream of *GRIN2A* (glutamate receptor, ionotropic, N-methyl D-aspartate 2A). The primary function of *GRIN2A* is to assist in controlling long-term memory and learning through regulation and efficiency of synaptic transmission. These receptors are essentially the gateway for calcium into post-synaptic cells[113]. Variants in this gene have been associated with various forms of epilepsy, sleep patterns, delayed psychomotor development, speech difficulties, seizures, mental retardation, and various mental disorders, including heroin addiction[114-120]. The only other gene within 500 kb of rs4141488 is C16orf72; little is known about the function of this gene. While GTEx[73] revealed no significant eQTLs nearby our lead variant, there is some evidence that this locus may lie within an important regulatory region. RegulomeDB[18] provided a score of 5 (minimal binding evidence) for rs4141488. Additionally, HaploReg[21] and UCSC browser show that our lead SNP and variants in high LD

(R^2>0.7) are within active enhancer regions for several tissues, including liver, fetal leg muscle, smooth stomach and intestinal muscle, cortex, and several embryonic and pluripotent cell types; and within altered binding motifs for EWSR1-FLI1, Elf3, STAT, CDP, HNF1, and SOX. Our lead SNP is within 500kb +/- of several previously identified GWAS SNPs for multiple traits, the nearest of which is rs17550532 associated with sudden cardiac arrest[121]. Other associations in this region include behavioral disinhibition[122], venous thromboembolism[123], and Transforming Growth Factor-β1 [5]; however, all of these are in low LD with rs4141488 (R^2<0.4).

rs1545348 (*RAI14*): Our lead SNP, rs1545348, lies within the intron of *RAI14* (Retinoic Acid Induced 14), although very little is known about the function of this gene in humans. There are four additional genes within 500 kb+/- of rs1545348, including *RAD1* (RAD1 checkpoint DNA exonuclease) 187 kb upstream. *RAD1* encodes a protein involved in stopping the cell cycle in response to DNA damage, as well as recruiting other proteins responsible for DNA repair[124,125], including in response to stress caused by cigarette smoke[126]. There is strong evidence of a regulatory role within the region surrounding our lead variant (RegulomeDB[18] score 4, minimal binding evidence). One significant (beta=-0.28, P=5.3E-6) eQTL between rs1545348 and *TTC23L* was found in sun exposed skin tissue (lower leg) (GTEx[73]). Additionally, HaploReg[21] and the UCSC browser reveal that the region surrounding our lead variant (+/- 500 kb, R^2>0.7) harbors marks of open and active chromatin and DNase hypersensitive regions across multiple tissues, including cancer, pluripotent, and normal tissue, brain and adipose tissue among others. Traits with nearby GWAS associations include several metabolite markers and left ventricular mass, although each of these associations are in low LD with rs1545348[127-131].

rs6470765 (*GSDMC*): There are three genes within 500 kb +/- of our lead SNP, rs6470765, which lies within an intron of *GSDMC* (gasdermin C). There is very little known about the function of *GSDMC*. Our lead SNP also lies 80 kb downstream of *FAM49B* (family with sequence similarity 49, member B). Similar to *CDK6*, a gene nearby another one of our novel variants, rs10269774, *FAM49B* is a target of *BACH1* transcription factor, which is involved in cellular response to oxidative stress and management of the cell cycle[132]. Also, *ASAP1* (ArfGAP With SH3 Domain, Ankyrin Repeat And PH Domain 1), a gene located 328 kb upstream of our association signal, may be involved in the differentiation of fibroblasts into adipocytes[133]. There is moderate evidence for the functional role of lead variant in regulation of gene expression (RegulomeDB[18] score of 6: minimal binding evidence). However, the GTEx[73] database indicates that rs6470765 is a significant eQTL for *GSDMC* in skeletal muscle, sun-exposed skin, and mucous in the esophagus. Furthermore, HaploReg[21] and the UCSC Browser highlight moderate evidence for regulatory elements in high LD >0.9, including DNAse hypersentive regions, and active enhancer and promotor regions in >20 tissue types (e.g. lung, adipose, skeletal muscle, epidermal and esophageal tissues, and many stem/pluripotent cell types). Our lead variant is within several altered binding sites for FOX1, FOX2 and SOX. Last, our lead SNP is in high LD with other potential cis-eQTLs for *GSDMC*. Nearby associations with other traits include height, hip circumference adjusted for BMI, and inflammatory bowel disorder[2,70,71,134].

rs6076699 (*PRNP*): There are seven genes within 500 kb+/- of our lead SNP, rs6076699. The lead SNP is 100kb upstream of *PRNP* (prion protein) is likely a signaling transducer involved in multiple biological processes related to nervous system, immune system, and general cellular functions[135-138]. Mutations in the repeat region as well as elsewhere in this gene have been associated with Creutzfeldt-Jakob disease, fatal familial insomnia, Gerstmann-Straussler disease, Huntington disease-like 1, and kuru [139-145].

Alternate forms of the oligmers have been shown to form in response to oxidative stress caused by copper exposure[146]. Copper is present in cigarette smoke and elevated in serum of smokers, but is not

outside of safe ranges according the U.S. Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, and Office on Smoking and Health[147,148]. Our lead SNP is 136 kb upstream from a related gene, *PRND* (prion protein 2), which is biochemically and structurally similar to *PRNP*[149]. Like PRNP, mutations in this gene may also be involved in neurocognitive disorders, although there are only weak associations[150,151]. A third prion protein (testes specific, *PRNT*) is found 145 kb away from our lead SNP; however no much is known about the function of this gene. Other nearby genes include *SLC23A2* (Solute Carrier Family 23 [Ascorbic Acid Transporter], Member 2), *ADRA1D* (Adrenoceptor Alpha 1D), *SMOX* (Spermine Oxidase), and *RASSF2* (Ras association [RalGDS/AF-6] domain family member 2). *SLC23A2* is essential for the uptake and transport of Vitamin C, which is an important nutrient for DNA and cellular repair in response to oxidative stress both directly and through supporting the repair of Vitamin E after exposure to oxidative agents[152-155]. Furthermore, this region is associated with success in smoking cessation and is implicated in addictive behaviors in general[156,157]. Nearby GWAS-identified associations include preeclampsia, and height[70,71,158]. There is little evidence that our association signal is involved in regulation of gene expression (RegulomeDB[18] score-5: minimal binding evidence) [18]. While our tag SNP is located within an active enhancer region (open chromatin marks, DNAse hypersentivity, and several transcription factor binding motifs), this activity appears tissue specific (sex-specific tissues and lungs)[21,111]. There are no other significant regulatory elements in high LD with rs6076699[21,73].

Waist-to-Hip Ratio adjusted for BMI (WHRadjBMI)

rs670752 (*BBX*): There are only three genes within 500 kb+/- of our lead SNP, rs670752, which lies within an intron of *BBX* (Bobby Sox Homolog [Drosophila]). While there is little known about the function of *BBX*, another nearby intronic variant, rs6437740, has been associated with smoking behavior in a previous GWAS[159]. Other nearby genes include *CCDC54* (coiled-coil domain containing 54) and *CD47* (CD47 molecule). Much is known about the function of *CD47* due to mouse models. *CD47* encodes a cell surface antigen involved in immune response to bacteria, cell adhesion, inflammatory response, and cell to cell signaling[160-162]. *CD47* expression is significantly decreased in obese individuals and negatively correlated with BMI, WC, and HIP in RBC [163].

Conversely, in mouse models, *CD47* deficient mice show decreased weight gain on high fat diets, increased energy expenditure, improved glucose profile, and decreased inflammation[164]. Our lead SNP, rs670752, has a score of 6 (very little binding evidence) in RegulomeDB[18] and no significant eQTLs were identified in GTEx[73]. However, our tag SNP was identified as a significant eQTL for *BBX* in brain tissue in HaploReg[21], Additionally, multiple SNPs in high LD with rs670752 provide several lines of evidence for nearby regulatory elements (e.g. active promoters, transcription factor binding motifs, strong and poised enhancers), mostly in pluripotent and embryonic cell lines, but also blood cell lines and brain tissue[21,111].

rs589428 (*EHMT2*). A total of seventy-seven genes are found near our lead SNP, rs589428, which is intronic within *EHMT2* (Euchromatic Histone-Lysine N-Methyltransferase 2). *EHMT2* encodes a histone methyltransferase, a group of genes involved in repression of transcription through the regulation of chromatin state [165]. The lead SNP is 302kb downstream of *TNF*. In patients with end-stage renal disease (ESRD) on long-term hemodialysis (HD), the SNP in the promoter region of the *IL-6* and *TNF-alpha*, and *IL-10*, show a strong association with indices of comorbidity and function, and biological and nutritional markers[166]. TNF-alpha promotes bone loss and inhibits bone formation and has an important role as a mediator of skeletal damage in inflammatory arthritis[167-170]. *TNF* is the master regulator of other inflammatory cytokines and the major cytokine in the pathogenesis of chronic inflammatory disease[171]. TNF-alpha exerts an important influence on adipose tissue metabolism and function. It inhibits the

expression of two major adipose tissue differentiation regulators: CCAAT and PPARγ-2 [172]. TNF-alpha promoter methylation levels could be involved in the susceptibility to stroke[173] and correlates with increased risk of coronary artery disease[174]. The risk of early childhood wheeze associated with early maternal smoking may be modified by TNF[175]. The lead SNP is also 287kb upstream of *NCR3*, which is associated with pulmonary function[176].

The top SNP is 17.5kb upstream of *NEU1* (Sialidase 1 (Lysosomal Sialidase)). The activity of *NEU1* is higher in epididymal fat and lower in the livers of two strains of obese and diabetic mice. Fluctuations in *NEU1* activity might be associated with the pathological status of these tissues in obesity[177]. The lead SNP is 50kb downstream of *HSPA1B*. Functional *HSPA1B* variants are associated with lung cancer risk and survival[178]. The top SNP is 65kb upstream of *CFB*. Increased concentrations of circulating binding factors fH and fB in subjects with altered glucose tolerance could reflect increased SVC-induced activation of the alternative pathway of the complement in omental adipose tissue linked to insulin resistance and metabolic disturbances[179]. The top SNP is 91kb upstream of *STK19*, which has been reported to be a pleiotropic gene for metabolic syndrome and inflammation and is associated with TG, BMI, WAIST, SBP and inflammatory markers including plasminogen activator inhibitor 1 (PAI-1) and white blood cell count (WBCC)[180]. Our top snp is 102kb upstream of *C4A*, which was identified as novel potential adipokine candidate regulator of obesity and adipose regions [181] between visceral and subcutaneous adipose tissue. The Top SNP is 102kb upstream of *C4B*. The carriers of C4B*Q0 (silent allele for the C4B gene) have a substantially increased risk to suffer from myocardial infarction or stroke. Compared to controls, C4B*Q0 carrier frequency was significantly higher at diagnosis in Icelandic smokers with angina pectoris (AP) or acute myocardial infarction (AMI) and Hungarian smokers with severe coronary artery disease, while no such difference was seen in nonsmokers. These findings indicate that C4B*Q0 genotype can be considered as a major covariate of smoking in precipitating the risk for AMI and associated mortality[182]. The top SNP is 150kb upstream of *DDAH2* in which SNP rs9267551 may confer increased risk for type 2 diabetes by affecting insulin sensitivity through increased asymmetric dimethylarginine (ADMA) levels [183,184].

Our top SNP is 222kb downstream of *APOM*. The PCSK9 pathway contributes to plasma apoM regulation in humans and the influence of PCSK9 on circulating apoM appears to be modified by adiposity [185]. In addition, APOM expression is related to FEV1/FVC (forced expiratory volume 1/ forced vital capacity) ratio and per cent emphysema [186]. The top SNP is 261kb downstream of *AGER/RAGE*. The lower level of soluble RAGE/AGER is associated with a number of components of metabolic syndrome (central obesity, hypertension, and hyperglycemia) [187]. Soluble RAGE is inversely associated with pancreatic cancer risk among Finnish male smokers [188]. The RAGE(2) haplotype is associated with diabetic nephropathy (DN) in type 2 diabetics and with earlier DN onset and, thus, can be regarded a marker for DN [189]. RAGE, via its interaction with ligands, serves as a cofactor exacerbating diabetic vascular disease [190]. Serum endogenous secretory RAGE (esRAGE) levels were inversely correlated with BMI and serum HDL-cholesterol[191]. In healthy subjects plasma levels of sRAGE were negatively correlated with BMI and waist/hip ratio supporting a possible protective role for these proteins before any evidence of diabetic or vascular complications[192].

The top SNP is 263 downstream of *AIF1*. The serum AIF-1 concentrations were positively correlated with levels of fasting plasma glucose, hemoglobin A1c, triglycerides, and uric acid, and with WC and BMI, and were inversely correlated with HDL cholesterol levels[193]. Also, the variants in AIF1 show evidence of association with adult obesity in the Greek population[194]. The top SNP is 306 downstream of *LTA*. SNPs in LTA are associated with chronic kidney disease in Type 2 diabetes[195]. The variability of LT-alpha

genotypes may have potential implications for individual susceptibility to asthma in atopic or in ever-smoking Chinese adults in Hong Kong[196].

The genome-wide association studies have reported the associations within 1Mb of region for age at menopause (R^2=0.32)[197], telomere length (R^2=0.22)[198], idiopathic membranous nephropathy[199] (R^2=0.45), chronic hepatitis B infection[200] (R^2= 0.45) and phospholipid levels (plasma) (R^2=0.23)[201]. This lead SNP is associated with regulatory motifs changed at Bcl6b, NF-kappaB, Pou5f1; associated with enhancer histone marks in stomach mucosa, HSMM cell derived skeletal muscle myotubes cell tissue; and in eQTL in various tissues including subcutaneous adipose, visceral omentum, lung and skeletal muscle tissues. The lead SNP is associated with eQTL in tibial artery and blood tissues from GTEx[73] analysis. The RegulomeDB[18] score for the lead SNP is 1f.

rs1856293 *(EYA4)*: A total of nine genes are found near our lead SNP, rs1856293. The lead SNP is 342kb downstream of *RPS12*. *RPS12* is a potential target gene of microRNA-377, which has been consistently upregulated in *in vitro* diabetic nephropathy (DN) models and in *in vivo* DN mouse models[202]. If *RPS12* is also upregulated in the diabetic milieu, it may contribute to the progression of DN. *RPS12* has been reported to be a strong candidate for diabetic nephropathy[203]. In addition, in the study of E3 rats, there were significant positive correlations between TG and the expression of *RPS12* gene[204]. The lead SNP is 83kb upstream of *EYA4*. Serum methylation levels of *EYA4* were significant discriminants between stage I colorectal cancer and healthy controls [205] and high methylation of the EYA4 gene is associated with ulcerative colitis with colorectal cancer[206]. The lead SNP is 446kb upstream of *VNN1*. Alternative splicing in *VNN1* is associated with colorectal cancer[207]. The combination of VNN1 and MMP9 may be used as a blood biomarker panel for the discrimination of pancreatic cancer-associated diabetes from type II diabetes[208]. There is no reported GWAS signal in high LD with the lead SNP. This lead SNP is associated with regulatory motifs changed at Esr2, LRH1, Myf_3, Sin3Ak-20_disc3 and T3R; and associated with enhancer histone marks in ESDR, SKIN and brain tissue. The RegulomeDB[18] score for the lead SNP is 6.

rs2001945 (*TRIB1*): There are five protein coding genes within 500 kb+/- of our lead SNP, rs2001945, which lies 27 kb downstream from *TRIB1*. *TRIB1* (tribbles pseudokinase 1) encodes a protein involved in ATP binding and the MAPK/ERK1/2 pathway[209]. Very little is known about the function of the other nearby genes, including *NSMCE2* (non-SMC element 2, *MMS21* homolog), *KIAA0196* (strumpellin), *SQLE* (qualene epoxidase), and *ZNF572* (Zinc Finger Protein 572). GTEx[73] indentified no significant eQTLs for our lead SNP; however, RegulomeDB[18] provided a score of 4 (minimal binding evidence [Transcription Factor binding + DNase peak]). Further, HaploReg[21]/UCSC Genome Browser reveal multiple lines of evidence across multiple tissues, including cis-eQTLs between rs2001945 for *TRIB1* and *NSMCE2* in brain tissue, strong DNAse hypersensitivity clusters both at the association peak and across SNPs in high LD with our lead SNP, transcription factor binding motifs, and open chromatin marks primarily in Human Umbilical Vein Endothelial Cells (HUVEC). There are several nearby previously-identified GWAS signals for related cardiometabolic and digestion-related traits, including lipids (e.g. triglycerides, LDL, HDL)[6,8,13,14,210-217], adiponectin[218], liver enzyme levels[219], gestational age[5], inflammatory bowel disease[134], Crohn's disease[220,221], and metabolite levels[222].

rs17065323 (*SMIM2*): A total of 6 genes are found within 500 kb of the lead marker, rs17065323. The SNP rs17065323, which is located 23.19 kb downstream of the long intergenic non-protein coding RNA 284 (*LINC00284*, 13q14.11), showed suggestive association with uric acid levels (p=8.7E-6, [223]). Variants of the *LACC1* (laccase (multicopper oxidoreductase) domain containing 1), at 159.72 downstream of rs17065323, were genome-wide associated with Crohn's disease [134,221], and a *LACC1* mutant showed evidence of association with systemic juvenile idiopathic arthritis [224]. In addition, GWASs have suggested

associations between variants on 13q14 with response to tocilizumab in rheumatoid arthritis (p=2E-7[225]), antineutrophil cytoplasmic antibody-associated vasculitis (p=3E-6[226]), and myotrophic lateral sclerosis (p=4E-6, [227]), as well as *SERP2* genotype-carbohydrate interaction influencing fasting insulin and homeostasis model assessment of insulin resistance (p=7E-6 and p=5E-6 , respectively [228]). The nearest protein-coding gene to our tag SNP is *SMIM2* (Small Integral Membrane Protein 2), located 89.5 kb upstream; however, very little is known about the function of *SMIM2*.

rs1049281 (*HLA-C*): Eighty-six genes are found within 500kb of rs1049281, which lies within the *HLA-C* gene at 6p21.3. *HLA-C*encodes an HLA class I heavy chain paralogue found in nearly all cells and important in the function of the immune system. There is strong evidence that our SNP is in a region likely to affect binding activity and gene expression in adipose tissue (RegulomeDB[18] score 1f). Over 100 alleles of the *HLA-C* gene have been described, and *HLA-C* has been associated with risk of various autoimmune diseases which can influence adiposity, including Type I diabetes, celiac disease, and psoriatic arthritis [229,230]. Our lead SNP is 314569 bp downstream of *DPCR1,* a gene associated with diffuse panbronchiolitis, a chronic inflammatory lung disease [231]. A variant near this gene (rs9368649), has been suggestively associated with smoking status (ever somker) and pack years (P~1.3E-07) [232], but not at GWS. This SNP is not in high LD with our lead SNP (R^2=0.152, D'=0.902). Our lead SNP is 190789 bp upstream of *HCP5*, a lncRNA. A variant (rs12175489) near this gene was suggestively associated (p=2.13E-06) with visceral adipose tissue (VAT) in men [103], but this variant is also not in high LD with our lead SNP (R^2=0.022, D'=0.478). Our lead SNP is 336394bp upstream of *AIF1*, 310030bp downstream of *NCR3*, and 341847 upstream of *BAT2*. Three variants in this region [rs2260000 (R^2=0.122, D'=0.526), rs1077393 (R^2=0.114, D'=0.434), and rs2844479 (R^2=0.100, D'=0.523) have been previously associated with variation in weight [233]. Another variant near *NCR3* (rs2070600) has been previously associated with ever-smoking and lung function, but is not in high LD with our lead SNP (R^2=0.137, D'=0.642) [176,232]. Our lead SNP is 340905bp downstream of *VARS2*, and a variant near this gene (rs7751505) has been suggestively associated with height change ($P<4.05 \times 10^{-6}$), though it is not in LD with our top SNP (R^2=0.054, D'=0.569). Two other variants in the region have been previously associated with extremes of height (p<5E-08), one of which is in strong LD with our lead SNP (rs2247056, 28923bp from rs1049281: R^2=0.814, D'=1.000; rs7741091: R^2=0.093, D'=0.652) [77].

SUPPLEMENTARY NOTE 3. **Detailed summary of eQTL methods and results.**

eQTL Methods

We used two approaches to systematically explore the role of novel loci in regulating gene expression. First, to gain a general overview of the regulatory role of newly identified GWAS regions, we conducted an eQTL lookup using >50 eQTL studies [234], with specific citations for >100 datasets included in the current query: 1) Blood cell related eQTL studies included fresh lymphocytes [235], fresh leukocytes [236], leukocyte samples in individuals with Celiac disease [237], whole blood samples [73,238-256], lymphoblastoid cell lines (LCL) derived from asthmatic children [257,258], HapMap LCL from 3 populations [259], a separate study on HapMap CEU LCL [260], additional LCL population samples [261-267],neutrophils [268,269],CD19+ B cells [270], primary PHA-stimulated T cells [261,264], CD4+ T cells [271], peripheral blood monocytes [267,270,272-275], long non-coding RNAs in monocytes [276] and CD14+ monocytes before and after stimulation with LPS or interferon-gamma [277], CD11+ dendritic cells before and after *Mycobacterium tuberculosis* infection [278] and a separate study of dendritic cells before or after stimulation with LPS, influenza or interferon-beta [279]. Micro-RNA QTLs [280,281], DNase-I QTLs [282], histone acetylation QTLs [283], and ribosomal occupancy QTLs [284] were also queried for LCL. Splicing QTLs [285] and micro-RNA QTLs [286] were queried in whole blood. 2) Non-blood cell tissue eQTLs searched included omental and subcutaneous adipose tissues[73,238,256,263,287], visceral adipose tissue [256], stomach [287], endometrial carcinomas [288], ER+ and ER- breast cancer tumor

cells [289], liver [256,287,290-293], osteoblasts [294], intestine [295] and normal and cancerous colon [296,297], skeletal muscle [256,298], breast tissue (normal and cancer)[299,300], lung [73,301-304], skin [73,263,267,305], primary fibroblasts [261,264,306], sputum [307], pancreatic islet cells [308], prostate [309], rectal mucosa [310], arterial wall [256] and heart tissue from left ventricles [73,311] and left and right atria [312]. Micro-RNA QTLs were also queried for gluteal and abdominal adipose [313] and liver [314]. Methylation QTLs were queried in pancreatic islet cells [315]. Further mRNA and micro-RNA QTLs were queried from ER+ invasive breast cancer samples, colon-, kidney renal clear-, lung- and prostate-adenocarcinoma samples [316]; 2 Brain eQTL studies included brain cortex [252,272,317-319], cerebellar cortex [320], cerebellum [289,318,321-323], frontal cortex [320,321,323], gliomas [324], hippocampus [320,323], inferior olivary nucleus (from medulla) [320], intralobular white matter [320], occiptal cortex [320], parietal lobe [322], pons [321], pre-frontal cortex [289,323,325,326], putamen (at the level of anterior commussure) [320], substantia nigra [320], temporal cortex [318,320,321,323], thalamus [323] and visual cortex [289].

Additional eQTL data was integrated from online sources including ScanDB (http://www.scandb.org/newinterface/about.html), the Broad Institute GTEx[73] Portal, and the Pritchard Lab (eqtl.uchicago.edu). Cerebellum, parietal lobe and liver eQTL data were downloaded from ScanDB. Cis-eQTLs were limited to those with P<1.0E-6 and trans-eQTLs with P<5.0E-8. Results for GTEx[73] Analysis V4 for 13 tissues were downloaded from the GTEx[73] Portal and then additionally filtered as described below [www. GTExportal.org: thyroid, leg skin (sun exposed), tibial nerve, aortic artery, tibial artery, skeletal muscle, esophagus mucosa, esophagus muscularis, lung, heart (left ventricle), stomach, whole blood, and subcutaneous adipose tissue [73]]. Splicing QTL (sQTL) results generated with sQTLseeker with false discovery rate P≤0.05 were retained. For all gene-level eQTLs, if at least 1 SNP passed the tissue-specific empirical threshold in GTEx[73], the best SNP for that eQTL was always retained. All gene-level eQTL SNPs with P<1.67E-11 were also retained, reflecting a global threshold correction of P=0.05/(30,000 genes X 1,000,000 tests).

Second, since public databases with eQTL data do not have information available on current smoking status, we also conducted an eQTL association analysis using expression results derived from fasting peripheral whole blood collected. Total RNA was isolated from frozen PAXgene blood tubes (PreAnalytiX, Hombrechtikon, Switzerland) and amplified using the WT-Ovation Pico RNA Amplification System (NuGEN, San Carlos, CA) according to the manufacturers' standard operating procedures. The obtained cDNA was hybridized to the Human Exon 1.0 ST Array (Affymetrix, Inc., Santa Clara, CA). The raw data were quantile-normalized, log2 transformed, followed by summarization using Robust Multi-array Average [327]and further adjusted for technical covariates, including the first principal component of the expression data, batch effect, and the all-probeset-mean residual. Study specific covariates in the association model included blood cell counts and cohort membership.
We evaluated all transcripts +/- 1MB around each novel variant in the Framingham Heart Study while accounting for current smoking status, using the following four approaches similar to those used in our primary analyses of our traits:

Model 1 (adjusted main effect of eQTL): Expression ~***SNP*** + SMK + age + age-squared + sex + study specific covariates

Model 2 (main effect of eQTL stratified by smoking status): Expression ~ ***SNP*** + age + age-squared + sex + study specific covariates

Model 3 (Interaction effect of eQTL): Expression ~ SNP + SMK + ***SNP*SMK*** + age + age-squared + sex + study specific covariates

Model 4 (Joint effect of eQTL): Expression ~ ***SNP*** + SMK + ***SNP*SMK*** + age + age-squared + sex + study specific covariates

Significance level was evaluated by FDR < 5% per eQTL analysis and across all loci identified for that model in the primary meta-analysis.

eQTL Results by Trait

Only significant cis-eQTLS in high LD with our novel lead SNPs (r^2>0.9, calculated in the CEU+YRI+CHB+JPT 1000 Genomes reference panel), or proxy SNPs, were retained for consideration.

For BMI, three of our seven novel SNPs across six loci that had at least one variant in high LD (r^2>0.9) with the tag SNP that is significantly (**Online Methods**) associated with expression of a gene transcript in the cerebellum and prefrontal cortex, or blood cell types, including *EPHA3*, *TTC14*, and *INADL*. Notably, our lead SNP, rs2481665, is a significant cis-eQTL for *INADL*, in prefrontal cortex tissue, and for *INADL* and *LITD1* in whole blood after adjusting for SMK (false discovery rate, FDR<5%). For the joint main + interaction effect eQTL analysis, we identified one significant eQTL for a BMI associated variant (rs12902602) for three gene transcripts (*PSMA4, CHRNA5*, and *CTSH*).

For WCadjBMI, five of our 12 novel SNPs were in high LD with a cis-eQTL for gene transcripts in the cerebellum, temporal cortex, prefrontal cortex, lymphoblastoid cells, liver, lung, lymph, omental adipose, subcutaneous adipose, Primary PHA-stimulated T cells, skin, and blood cell tissues in publicly available databases. In our cis-eQTL analyses adjusting for SMK, four of our nine novel lead SNPs were significant cis-eQTLs for 14 gene transcripts in 12 genes. Additionally, for the joint main + interaction effect eQTL analysis, we identified that two variants that were associated with the expression of *SEPT2, FARP2, PASK*, and *HDLBP* (rs6743226) and *KIF1B* (rs17396340).

For WHRadjBMI, three of out six novel SNPs were in high LD with a nearby cis-eQTL for gene transcripts in subcutaneous adipose tissue and blood cell types. We identified five novel WHRadjBMI variants near significant cis-eQTLs for 49 gene transcripts after adjusting for SMK, the most significant of which was between our tag SNP rs1049281 and *MSH5*. Additionally, for the joint main and interaction effect eQTL analysis, we identified two novel WHRadjBMI variants (rs1049281, rs1856293) were associated with 19 gene transcripts.

Across all of our three obesity-related traits, the majority of significant cis-eQTLs from public databases are found in blood cell lines (63% of unique SNP-transcript associations) (**Supplementary Table 16**). However, as in previous eQTL analyses of obesity-associated variants, we identify cis-eQTLs in brain and adipose tissue. Further analyses are needed to determine if these tissue-specific eQTLs remain significant after accounting for SMK, but our de-novo analysis in whole blood samples from the Framingham Heart Study using models to account for SMK indicate that gene expression may underlie our association signals in some instances and smoking exposure may play a role in influencing these associations (**Supplementary Tables 16-18**).

SUPPLEMENTARY NOTE 4. Full list of acknowledgments, including study-specific acknowledgements.

Writing and analysis of this study for AEJ were supported by the American Heart Association (13POST16500011) and NIH (2T32HL007055-36), for TOK by the Danish Council for Independent

Research (DFF – 1333-00124 and Sapere Aude program grant DFF – 1331-00730B), and analyses performed by JP were funded through NIH (T32GM074905).

Study Specific Acknowledgements

AE: Genotyping was funded by Cavadis B.V. Sander W. van der Laan is funded through grants from the Netherlands CardioVascular Research Initiative ("GENIUS", CVON2011-19), and the Interuniversity Cardiology Institute of the Netherlands (ICIN, 09.001). Sander W. van der Laan and Saskia Haitjema are funded through a grant from FP7 EU project CVgenes@target (HEALTH-F2-2013-601456). Marten A. Siemelink acknowledges funding by the European Union (BiomarCaRE, grant number: HEALTH-2011-278913) and Technology Foundation STW (Stichting voor de Technische Wetenschappen, Project 11679). Claudia Tersteeg, Krista den Ouden, Mirjam B. Smeets, and Loes B. Collé are graciously acknowledged for their work on the DNA extraction. Astrid E.M.W. Willems, Evelyn Velema, Kristy M. J. Vons, Sara Bregman, Timo R. ten Brinke, Sara van Laar, Sander M. van de Weg, Louise M. Catanzariti, Arjan H. Schoneveld, Petra H. Homoed-van der Kraak, Aryan Vink, and Joyce E.P. Vrijenhoek are graciously acknowledged for their past and continuing work on the Athero-Express Biobank Study. We would also like to thank all the (former) employees involved in the Athero-Express Biobank Study of the Departments of Surgery of the St. Antonius Hospital Nieuwegein and University Medical Center Utrecht for their continuing work. Jessica van Setten is graciously acknowledged for her help in the quality assurance and quality control of the genotype data. Lastly, we would like to thank all participants of the Athero-Express Biobank Study. MAS is funded by the European Union (BiomarCaRE, grant number: HEALTH-2011-278913), and the technology foundation "Stichting voor de Technische Wetenschappen" through the Danone partnership program (Project 11679). **SWvdL/SH**: SWvdL/SH: SWvdL is funded through grants from the Netherlands CardioVascular Research Initiative ("GENIUS", CVON2011-19) and the Interuniversity Cardiology Institute of the Netherlands (ICIN, 09.001). SWvdL and SH are both funded through the FP7 EU project CVgenes@target (HEALTH-F2-2013-601456).

AGES: The Reykjavik Study cohort originally comprised a random sample of 30,795 men and women born in 1907-1935 and living in Reykjavik in 1967. A total of 19,381 attended, resulting in 71% recruitment rate. Between 2002 and 2006, the AGES-Reykjavik study re-examined 5,764 survivors of the original cohort who had participated before in the Reykjavik Study. [Harris, T. B. et al. (2007). American Journal of Epidemiology, 165(9), 1076-1087. doi:10.1093/aje/kwk115]. This study has been funded by NIH contract N01-AG-1-2100, the NIA Intramural Research Program, Hjartavernd (the Icelandic Heart Association), and the Althingi (the Icelandic Parliament). The study is approved by the Icelandic National Bioethics Committee, VSN: 00-063. The researchers are indebted to the participants for their willingness to participate in the study.

ARIC: The Atherosclerosis Risk in Communities Study is carried out as a collaborative study supported by National Heart, Lung, and Blood Institute contracts N01-HC-55015, N01-HC-55016, N01-C-55018, N01-HC-55019, N01-HC-55020, N01-HC-55021, N01-HC-55022, R01HL087641, R01HL59367 and R01HL086694; National Human Genome Research Institute contract U01HG004402; and National Institutes of Health contract HHSN268200625226C. Infrastructure was partly supported by Grant Number UL1RR025005, a component of the National Institutes of Health and NIH Roadmap for Medical Research. The project described was supported by Grant Number UL1 RR 025005 from the National Center for Research Resources (NCRR), a component of the National Institutes of Health (NIH) and NIH Roadmap for Medical Research, and its contents are solely the responsibility of the authors and do not necessarily represent the official view of NCRR or NIH. The authors thank the staff and participants of the ARIC Study for their important contributions.

AUSTWIN: We acknowledge the contributions of many staff in the Genetic Epidemiology Unit, Queensland Institute of Medical Research, in interviewing study participants, sample processing and DNA extraction, and data management. Funding for aspects of this work was provided by the Australian National Health and Medical Research Council (241944, 339462, 389927, 389875, 389891, 389892, 389938, 442915, 442981, 496739, 552485, 552498), the EU 5th Framework Programme GenomeEUTwin Project (QLG2-CT-2002-01254), and the U.S. National Institutes of Health (AA07535, AA10248, AA11998, AA13320, AA13321, AA13326, AA14041, AA17688, DA12854, MH66206). A portion of the genotyping on which this study was based was carried out at the Center for Inherited Disease Research, Baltimore (CIDR) (Illumina 370K scans on 4300 individuals), through an access award to our late colleague Dr. Richard Todd. Parts of the statistical analyses were carried out on the Genetic Cluster Computer, which is financially supported by the Netherlands Scientific Organization (NWO 480-05-003). R.P.S.M. was, and G.W.M. is, supported by National Health and Medical Research Council (NHMRC) Fellowship Schemes.

BHS: The Busselton Health Study acknowledges the generous support for the 1994/5 follow-up study from Healthway, Western Australia and the numerous Busselton community volunteers who assisted with data collection and the study participants from the Shire of Busselton. The Busselton Health Study is supported by Department of Health and the Office of Science of the Government of Western Australia.

BioMe (MSSM): The Mount Sinai BioMe Biobank is supported by The Andrea and Charles Bronfman Philanthropies.

BLSA: The BLSA was supported by the Intramural Research Program of the NIH, National Institute on Aging.

B58C: We acknowledge use of phenotype and genotype data from the British 1958 Birth Cohort DNA collection, funded by the Medical Research Council grant G0000934 and the Wellcome Trust grant 068545/Z/02. Genotyping for the B58C-WTCCC subset was funded by the Wellcome Trust grant 076113/B/04/Z. The B58C-T1DGC genotyping utilized resources provided by the Type 1 Diabetes Genetics Consortium, a collaborative clinical study sponsored by the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), National Institute of Allergy and Infectious Diseases (NIAID), National Human Genome Research Institute (NHGRI), National Institute of Child Health and Human Development (NICHD), and Juvenile Diabetes Research Foundation International (JDRF) and supported by U01 DK062418. B58C-T1DGC GWAS data were deposited by the Diabetes and Inflammation Laboratory, Cambridge Institute for Medical Research (CIMR), University of Cambridge, which is funded by Juvenile Diabetes Research Foundation International, the Wellcome Trust and the National Institute for Health Research Cambridge Biomedical Research Centre; the CIMR is in receipt of a Wellcome Trust Strategic Award (079895). The B58C-GABRIEL genotyping was supported by a contract from the European Commission Framework Programme 6 (018996) and grants from the French Ministry of Research.

CHS: This CHS research was supported by NHLBI contracts HHSN268201200036C, HHSN268200800007C, N01HC55222, N01HC85079, N01HC85080, N01HC85081, N01HC85082, N01HC85083, N01HC85086; and NHLBI grants U01HL080295, R01HL087652, R01HL105756, R01HL103612, and R01HL120393 with additional contribution from the National Institute of Neurological Disorders and Stroke (NINDS). Additional support was provided through R01AG023629 from the National Institute on Aging (NIA). A full list of principal CHS investigators and institutions can be found at CHS-NHLBI.org. The provision of

genotyping data was supported in part by the National Center for Advancing Translational Sciences, CTSI grant UL1TR000124, and the National Institute of Diabetes and Digestive and Kidney Disease Diabetes Research Center (DRC) grant DK063491 to the Southern California Diabetes Endocrinology Research Center. Subjects for the present study were selected from CHS participants who had donated DNA samples for storage and provided informed consent for participation in DNA studies of cardiovascular-disease-related traits. The content of this work is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

CLHNS: The Cebu Longitudinal Health and Nutrition Survey (CLHNS) was supported by National Institutes of Health grants DK078150, TW05596, HL085144 and TW008288 and pilot funds from RR20649, ES10126, and DK56350. We thank the Office of Population Studies Foundation research and data collection teams and the study participants who generously provided their time for this study.

COLAUS: The CoLaus study was and is supported by research grants from GlaxoSmithKline, the Faculty of Biology and Medicine of Lausanne, and the Swiss National Science Foundation (grants 33CSCO-122661, 33CS30-139468 and 33CS30-148401). The authors thank Vincent Mooser and Dawn Waterworth, Co-PIs of the CoLaus study. Special thanks to Yolande Barreau, Mathieu Firmann, Vladimir Mayor, Anne-Lise Bastian, Binasa Ramic, Martine Moranville, Martine Baumer, Marcy Sagette, Jeanne Ecoffey and Sylvie Mermoud for data collection. SB is supported by the Swiss National Science Foundation (grant 3100AO-116323/1) and the Swiss Institute of Bioinformatics. ZK received financial support from the Leenaards Foundation, the Swiss Institute of Bioinformatics and the Swiss National Science Foundation (31003A-143914, 51RTP0_151019).

CROATIA-Korcula: We would like to acknowledge the contributions of the recruitment team in Korcula, the administrative teams in Croatia and Edinburgh and the people of Korcula. The SNP genotyping for the CROATIA-Korcula cohort was performed in Helmholtz Zentrum München, Neuherberg, Germany. The study was financed by the Medical Research Council UK, the Ministry of Science, Education and Sport in the Republic of Croatia (grant number 108-1080315-0302) and the Croatian Science Foundation (grant 8875).

CROATIA-Vis: We would like to acknowledge the staff of several institutions in Croatia that supported the field work, including but not limited to The University of Split and Zagreb Medical Schools, Institute for Anthropological Research in Zagreb and Croatian Institute for Public Health. The SNP genotyping for the CROATIA-Vis cohort was performed in the core genotyping laboratory of the Wellcome Trust Clinical Research Facility at the Western General Hospital, Edinburgh, Scotland. This study was supported through grants from the Medical Research Council UK, the Ministry of Science, Education and Sport of the Republic of Croatia (number 108-1080315-0302) and the European Union framework program 6 EUROSPAN project (contract no. LSHG-CT-2006-018947).

DESIR: This study was supported in part by grants from SFD ("Société Francophone du 358 Diabète"), CPER ("Contrat de Projets État-Région"), and ANR ("Agence Nationale de la 359 Recherche"). The D.E.S.I.R. study has been supported by INSERM contracts with CNAMTS, Lilly, Novartis Pharma and Sanofi-Aventis; by INSERM (Réseaux en Santé Publique, Interactions entre les déterminants de la santé), Cohortes Santé TGIR, the Association Diabète Risque Vasculaire, the Fédération Française de Cardiologie, La Fondation de France, ALFEDIAM, ONIVINS, Ardix Medical, Bayer Diagnostics, Becton Dickinson, Cardionics, Merck Santé, Novo Nordisk, Pierre Fabre, Roche, Topcon.

DR's EXTRA: The DR.s EXTRA Study was supported by grants to R. Rauramaa by the Ministry of Education and Culture of Finland (627;2004-2011), Academy of Finland (102318; 123885), Kuopio

University Hospital , Finnish Diabetes Association, Finnish Heart Association, Päivikki and Sakari Sohlberg Foundation and by grants from European Commission FP6 Integrated Project (EXGENESIS); LSHM-CT-2004-005272, City of Kuopio and Social Insurance Institution of Finland (4/26/2010).

EGCUT: EGCUT received support from EU FP7 grant Biobanking and Biomolecular Resources Research Infrastructure (BBMRI)-LPC 313010, targeted financing from Estonian Government IUT20-60, IUT24-6, Estonian Research Roadmap through the Estonian Ministry of Education and Research (3.2.0304.11-0312), Center of Excellence in Genomics (EXCEGEN), Development Fund from the University of Tartu (SP1GVARENG). This work was also supported by the US National Institute of Health [R01DK075787].

Ely: We are grateful to all the volunteers and to the staff of St. Mary's Street Surgery, Ely and the study team. The Ely Study was funded by the MRC (MC_U106179471) and Diabetes UK. Genotyping in the Ely and Fenland studies was supported in part by an MRC-GlaxoSmithKline pilot programme grant (G0701863).

EPIC: The EPIC Norfolk diabetes case cohort study is nested within the EPIC Norfolk Study, which is supported by programme grants from the Medical Research Council, and Cancer Research UK and with additional support from the European Union, Stroke Association, British Heart Foundation, Research into Ageing, Department of Health, The Wellcome Trust and the Food Standards Agency. Genotyping was in part supported by the MRC-GSK pilot programme grant. We acknowledge the contribution of the staff and participants of the EPIC-Norfolk Study.

EPIC-Norfolk: The EPIC Norfolk Study is funded by program grants from the Medical Research Council UK and Cancer Research UK, and by additional support from the European Union, Stroke Association, British Heart Foundation, Department of Health, Food Standards Agency, and the Wellcome Trust.

ERF: The ERF study as a part of EUROSPAN (European Special Populations Research Network) was supported by European Commission FP6 STRP grant number 018947 (LSHG-CT-2006-01947) and also received funding from the European Community's Seventh Framework Programme (FP7/2007-2013)/grant agreement HEALTH-F4-2007-201413 by the European Commission under the programme "Quality of Life and Management of the Living Resources" of 5th Framework Programme (no. QLG2-CT-2002-01254). High-throughput analysis of the ERF data was supported by joint grant from Netherlands Organization for Scientific Research and the Russian Foundation for Basic Research (NWO-RFBR 047.017.043). Exome sequencing analysis in ERF was supported by the ZonMw grant (project 91111025). We are grateful to all study participants and their relatives, general practitioners and neurologists for their contributions and to P. Veraart for her help in genealogy, J. Vergeer for the supervision of the laboratory work and P. Snijders for his help in data collection. Najaf Amin is supported by the Netherlands Brain Foundation (project number F2013(1)-28)

FamHS: The Family Heart Study was supported by grant R01-DK-089256 from NIDDK and grant R01HL117078 from NHLBI.

Fenland: The Fenland Study is funded by the Wellcome Trust and the Medical Research Council (MC_U106179471). We are grateful to all the volunteers for their time and help, and to the General Practitioners and practice staff for assistance with recruitment. We thank the Fenland Study Investigators, Fenland Study Co-ordination team and the Epidemiology Field, Data and Laboratory teams.

FramHS: This research was conducted in part using data and resources from the Framingham Heart Study of the National Heart Lung and Blood Institute of the National Institutes of Health and Boston University School of Medicine. The analyses reflect intellectual input and resource development from the Framingham Heart Study investigators participating in the SNP Health Association Resource (SHARe) project. This work was partially supported by the National Heart, Lung and Blood Institute's Framingham Heart Study (Contract No. N01-HC-25195 and Contract No. HHSN268201500001I) and its contract with Affymetrix, Inc for genotyping services (Contract No. N02-HL-6-4278). A portion of this research utilized the Linux Cluster for Genetic Analysis (LinGA-II) funded by the Robert Dawson Evans Endowment of the Department of Medicine at Boston University School of Medicine and Boston Medical Center. This research was partially supported by grant R01-DK089256 from the National Institute of Diabetes and Digestive and Kidney Diseases (MPIs: I.B. Borecki, L.A. Cupples, K. North).

FUSION: Support for FUSION was provided by NIH grants R01-DK062370 (to M.B.), R01-DK072193 (to K.L.M.), and intramural project number 1Z01-HG000024 (to F.S.C.). Genome-wide genotyping was conducted by the Johns Hopkins University Genetic Resources Core Facility SNP Center at the Center for Inherited Disease Research (CIDR), with support from CIDR NIH contract no. N01-HG-65403.

Gendian: The support of the physicians, the patients, and the staff of the Diabetes Zentrum Mergentheim (Head: Prof. Dr. Thomas Haak), the diabetes outpatient clinic Dr Nusser - Dr Kreisel, the dialysis centers KfH Amberg, KfH Bayreuth, KfH Deggendorf, KfH Donauwörth, KfH Freising, KfH Freyung, KfH Fürth, KfH Hof, KfH Ingolstadt, KfH Kelheim, KfH München Elsenheimerstraße, KfH München-Schwabing, KfH Neumarkt, KfH Neusäß, KfH Oberschleißheim, KfH Passau, KfH Plauen, KfH Regensburg Günzstraße, KfH Regensburg Caritas-Krankenhaus, KfH Straubing, KfH Sulzbach-Rosenberg, KfH Weiden, Dialysezentrum Augsburg Dr. Kirschner, Dialysezentrum Bad Alexandersbad, KfH Bamberg, Dialysezentrum Emmering, Dialysezentrum Klinikum Landshut, Dialysezentrum Landshut, Dialysezentrum Pfarrkirchen, Dialysezentrum Schwandorf, Dr. Angela Götz, the medical doctoral student Johanna Christ and the Study Nurse Ingrid Lugauer. The expert technical assistance of Claudia Strohmeier is gratefully acknowledged. Phenotyping was funded by the Dr. Robert Pfleger-Stiftung (Dr Carsten A. Böger), the MSD Stipend Diabetes (Dr Carsten A. Böger) and the University Hospital of Regensburg (intramural grant ReForM A to Dr. A. Götz, ReForM C to Dr. Carsten Böger). Genome-wide genotyping was funded by the KfH Stiftung Präventivmedizin e.V. (Dr. Carsten A. Böger, Dr. Jens Brüning), the Else Kröner-Fresenius-Stiftung (2012_A147 to Dr Carsten A. Böger and Dr Iris M. Heid) and the University Hospital Regensburg (Dr Carsten A. Böger). Data analysis was funded by the Else Kröner-Fresenius Stiftung (Dr. Iris M. Heid and Dr. Carsten A. Böger: 2012_A147; Dr. Carsten A. Böger and Dr. Bernhard K. Krämer: P48/08//A11/08).

Generation Scotland (GS): We would like to acknowledge the contributions of the families who took part in the Generation Scotland: Scottish Family Health Study, the general practitioners and Scottish School of Primary Care for their help in recruiting them, and the whole Generation Scotland team, which includes academic researchers, IT staff, laboratory technicians, statisticians and research managers. Genotyping was performed at the Wellcome Trust Clinical Research Facility Genetics Core at Western General Hospital, Edinburgh, UK. GS:SFHS is funded by the Scottish Executive Health Department, Chief Scientist Office, grant number CZD/16/6. Exome array genotyping for GS: SFHS was funded by the Medical Research Council UK. MIM is a Wellcome Trust Senior Investigator, and an NIHR Senior Investigator.

GENOA: The Genetic Epidemiology Network of Arteriopathy (GENOA) study is supported by the National Institutes of Health, grant numbers HL054457, HL087660, and HL119443 from National Heart, Lung,

Blood Institute. We thank Eric Boerwinkle, PhD from the Human Genetics Center and Institute of Molecular Medicine and Division of Epidemiology, University of Texas Health Science Center, Houston, Texas, USA and Julie Cunningham, PhD from the Department of Health Sciences Research, Mayo Clinic College of Medicine, Rochester, MN, USA for their help with genotyping.

GLACIER: The GLACIER study was funded by project grants to Paul W. Franks from Novo Nordisk, the Swedish Heart-Lung Foundation, the Swedish Diabetes Association, Påhlssons Foundation, the Swedish Research Council, Umeå University Career Development Award, and The Heart Foundation of Northern Sweden. Frida Renström was supported by a post-doctoral stipend from the Swedish Heart-Lung Foundation. The investigators thank the staff at the Wellcome Trust Sanger Institute for technical assistance with genotyping (WT098051 to the list of Wellcome Trust funded grants). The investigators are indebted to the study participants who dedicated their time and samples to these studies and the staff at the Umeå Medical Biobank and VIP for biomedical data collection and preparation.

GOOD: Financial support was received from the Swedish Research Council, the Swedish Foundation for Strategic Research, the ALF/LUA research grant in Gothenburg, the Lundberg Foundation, the Torsten and Ragnar Söderberg's Foundation, the Västra Götaland Foundation, the Göteborg Medical Society, the Novo Nordisk foundation, and the European Commission grant HEALTH-F2-2008-201865-GEFOS. We would like to acknowledge Maria Nethander at the genomics core facility at University of Gothenburg for statistical analyses.

GOYA: This study was conducted as part of the activities of the Gene-diet Interactions in Obesity project (GENDINOB, www.gendinob.dk) and the MRC centre for Causal Analyses in Translational Epidemiology (MRC CAiTE). We thank all the participants of the study. TSA was also funded by the GENDINOB project and acknowledges the same.

GxE: Our chief acknowledgement is to the participants in these studies for their willingness to contribute. We also thank Nurses Orgen Brown and Diedre Thomas for assistance with recruitment as well as past and present Laboratory technologists and drivers at TMRU for their invaluable technical assistance. This work was supported by NIH Grants R01HL53353 and R01DK075787.

Health2006: The Health2006 study was financially supported by grants from the Velux Foundation; the Danish Medical Research Council, Danish Agency for Science, Technology and Innovation; the Aase and Ejner Danielsens foundation; ALK-Abello´ A/S (Hørsholm, Denmark), Timber Merchant Vilhelm Bangs Foundation, MEKOS Laboratories (Denmark) and Research Centre for Prevention and Health, the Capital Region of Denmark. The Health2006 was approved by the Ethical Committee of Copenhagen (KA-20060011) and the Danish Data Protection Agency.

HERITAGE: The HERITAGE Family Study is supported by the National Heart, Lung, and Blood Institute Grants HL-45670 and HL118305.

HRS: The HRS is supported by the National Institute on Aging (NIA U01AG009740). The genotyping was funded separately by the National Institute on Aging (RC2 AG036495, RC4 AG039029). Our genotyping was conducted by the NIH Center for Inherited Disease Research (CIDR) at Johns Hopkins University. Genotyping quality control and final preparation of the data were performed by the Genetics Coordinating Center at the University of Washington.

HUNT2: The Nord-Trøndelag Health Study (The HUNT Study) is a collaboration between HUNT Research Centre (Faculty of Medicine, Norwegian University of Science and Technology NTNU), Nord-Trøndelag County Council, Central Norway Health Authority, and the Norwegian Institute of Public Health.

HYPERGENES: The study was supported by the European Union (FP7-HEALTH-F4-2007-201550-HYPERGENES, HEALTH-2011.2.4.2-2-EU-MASCARA, HEALTH-F7-305507 HOMAGE and the European Research Council Advanced Researcher Grant-2011-294713-EPLORE); InterOmics project (PB05 MIUR-CNR Italian Flagship Project); The Fonds voor Wetenschappelijk Onderzoek Vlaanderen, Ministry of the Flemish Community, Brussels, Belgium (G.0881.13 and G.088013).

IMPROVE: IMPROVE was supported by the European Commission (Contract number: QLG1-CT-2002-00896), the Swedish Heart-Lung Foundation, the Swedish Research Council (projects 8691 and 0593), the Knut and Alice Wallenberg Foundation, the Foundation for Strategic Research, the Stockholm County Council (project 592229), the Strategic Cardiovascular and Diabetes Programmes of Karolinska Institutet and Stockholm County Council, the European Union Framework Programme 7 (FP7/2007-2013) for the Innovative Medicine Initiative under grant agreement n° IMI/115006 (the SUMMIT consortium), the Academy of Finland (Grant #110413), the British Heart Foundation (RG2008/08, RG2008/014) and the Italian Ministry of Health (Ricerca Corrente).

InCHIANTI: The InCHIANTI study baseline (1998-2000) was supported as a "targeted project" (ICS110.1/RF97.71) by the Italian Ministry of Health and in part by the U.S. National Institute on Aging (Contracts: 263 MD 9164 and 263 MD 821336).

Inter99: The Inter99 study was funded by: Danish Research Councils; The Health Foundation; The Danish Centre for Evaluation and Health Technology Assessment; Copenhagen County; Danish Heart Foundation; Ministry of Health and Prevention; Danish Pharmaceutical Association; Augustinus Foundation; Novo Nordisk; Velux Foundation; Becket Foundation and Ib Henriksens Foundation.

KORA: The KORA study was initiated and financed by the Helmholtz Zentrum München – German Research Center for Environmental Health, which is funded by the German Federal Ministry of Education and Research (BMBF, 01ER1206 and 01ER1507 for IMH) and by the State of Bavaria. Furthermore, KORA research was supported within the Munich Center of Health Sciences (MC-Health), Ludwig-Maximilians-Universität, as part of LMUinnovativ. [PMID: 16032513] [PMID: 16032514]

Lifelines: The Lifelines Cohort Study, and generation and management of GWAS genotype data for the Lifelines Cohort Study is supported by the Netherlands Organization of Scientific Research NWO (grant 175.010.2007.006), the Economic Structure Enhancing Fund (FES) of the Dutch government, the Ministry of Economic Affairs, the Ministry of Education, Culture and Science, the Ministry for Health, Welfare and Sports, the Northern Netherlands Collaboration of Provinces (SNN), the Province of Groningen, University Medical Center Groningen, the University of Groningen, Dutch Kidney Foundation and Dutch Diabetes Research Foundation.

LOLIPOP: The LOLIPOP study is supported by the National Institute for Health Research (NIHR) Comprehensive Biomedical Research Centre Imperial College Healthcare NHS Trust, the British Heart Foundation (SP/04/002), the Medical Research Council (G0601966,G0700931), the Wellcome Trust (084723/Z/08/Z) the NIHR (RP-PG-0407-10371), European Union FP7 (EpiMigrant, 279143) and Action on Hearing Loss (G51). The work was carried out in part at the NIHR/Wellcome Trust Imperial Clinical Research Facility. We thank the participants and research staff who made the study possible.

LURIC: We thank the LURIC study team who were either temporarily or permanently involved in patient recruitment as well as sample and data handling, in addition to the laboratory staff at the Ludwigshafen General Hospital and the Universities of Freiburg and Ulm, Germany. This work was supported by the 7th Framework Program (integrated project AtheroRemo, grant agreement number 201668 and RiskyCAD, grant agreement number 305739) of the European Union and by the INTERREG IV Oberrhein Program (Project A28, Genetic mechanisms of cardiovascular diseases) with support from the European Regional Development Fund (ERDF) and the Wissenschaftsoffensive TMO.

MESA: MESA and the MESA SHARe project are conducted and supported by the National Heart, Lung, and Blood Institute (NHLBI) in collaboration with MESA investigators. Support for MESA is provided by contracts N01-HC-95159, N01-HC-95160, N01-HC-95161, N01-HC-95162, N01-HC-95163, N01-HC-95164, N01-HC-95165, N01-HC-95166, N01-HC-95167, N01-HC-95168, N01-HC-95169, UL1-TR-001079, and UL1-TR-000040. Funding for MESA SHARe genotyping was provided by NHLBI Contract N02-HL-6-4278. The provision of genotyping data was supported in part by the National Center for Advancing Translational Sciences, CTSI grant UL1TR000124, and the National Institute of Diabetes and Digestive and Kidney Disease Diabetes Research Center (DRC) grant DK063491 to the Southern California Diabetes Endocrinology Research Center. The authors thank the MESA participants, as well as the Coordinating Centers, investigators, and study staff for their valuable contributions. A full list of participating MESA investigators and institutions can be found at http://www.mesa-nhlbi.org.

METSIM: The METSIM study was funded by the Academy of Finland (grants no. 77299 and 124243).

MRC NSHD: This work was funded by the Medical Research Council (MC_UU_12019/1), the British Heart Foundation (RG/10/12/28456) and the Wellcome Trust (088869/B/09/Z). We are very grateful to the members of this birth cohort for their continuing interest and participation in the study. We would like to acknowledge the Swallow group, UCL, who performed the DNA extractions (Rousseau, et al 2006). DOI: 10.1111/j.1469-1809.2006.00250.x.

MrOS Sweden: This work was supported by the Swedish Research Council, the Swedish Foundation for Strategic Research, The ALF/LUA research grant in Gothenburg, the Lundberg Foundation, the Torsten and Ragnar Söderberg's Foundation, Magnus Bergvall Foundation, Åke Wiberg Foundation, Tore Nilson Foundation and The Swedish Society for Medical Research.

NHS: The study was supported by grants from the National Heart, Lung, and Blood Institute (HL071981, HL034594, HL126024), the National Institute of Diabetes and Digestive and Kidney Diseases (DK091718, DK100383, DK078616), the Boston Obesity Nutrition Research Center (DK46200), and United States – Israel Binational Science Foundation Grant2011036.

NTR: This study was funded by the Netherlands Organization for Scientific Research (NWO) and The Netherlands Organisation for Health Research and Development (ZonMW) grants 904-61-090, 985-10-002, 904-61-193,480-04-004, 400-05-717, Addiction-31160008, Middelgroot-911-09-032, Spinozapremie 56-464-14192, Center for Medical Systems Biology (CSMB, NWO Genomics), NBIC/BioAssist/RK(2008.024), Biobanking and Biomolecular Resources Research Infrastructure (BBMRI – NL, 184.021.007). VU University's Institute for Health and Care Research (EMGO+) and Neuroscience Campus Amsterdam (NCA); the European Science Foundation (ESF, EU/QLRT-2001-01254), the European Community's Seventh Framework Program (FP7/2007-2013), ENGAGE (HEALTH-F4-2007-201413); the European Research Council (ERC Advanced, 230374, Starting grant 284 167), Rutgers University Cell and

DNA Repository (NIMH U24 MH068457-06), the Avera Institute, Sioux Falls, South Dakota (USA) and the National Institutes of Health (NIH, R01D0042157-01A, MH081802, Grand Opportunity grant 1RC2 MH089951). Part of the genotyping and analyses were funded by the Genetic Association Information Network (GAIN) of the Foundation for the National Institutes of Health.

NFBC1966/Oxford Univ: The Northern Finland Birth Cohort (NFBC) Research program, received financial support from Academy of Finland (1114194, 24300796), NHLBI grant 5R01HL087679 through the STAMPEED program (1RL1MH083268-01), ENGAGE project and grant agreement HEALTH-F4-2007-201413, the Medical Research Council (grant G0500539, centre grant G0600705, PrevMetSyn), and the Wellcome Trust (project grant GR069224), UK. The program is currently being funded by the H2020-633595 DynaHEALTH action and Academy of Finland EGEA-project. University of Oxford, UK, was funded by the British Heart Foundation (grant code SP/13/2/30111), the European Commission (ENGAGE: HEALTH-F4-2007-201413), Medical Research Council (G0601261), and the Wellcome Trust (090532, 098381). M-R.J. received funding from the European Union's Horizon 2020 research and innovation programme [under grant agreement No 633595].

ORCADES: ORCADES was supported by the Chief Scientist Office of the Scottish Government, the Royal Society, the MRC Human Genetics Unit, Arthritis Research UK and the European Union framework program 6 EUROSPAN project (contract no. LSHG-CT-2006-018947). DNA extractions were performed at the Wellcome Trust Clinical Research Facility in Edinburgh. We would like to acknowledge the invaluable contributions of Lorraine Anderson and the research nurses in Orkney, the administrative team in Edinburgh and the people of Orkney.

PIVUS: PIVUS was supported by Knut and Alice Wallenberg Foundation (Wallenberg Academy Fellow), European Research Council (ERC Starting Grant), Swedish Diabetes Foundation (2013-024), Swedish Research Council (2012-1397, 2012-1727, and 2012-2215), Marianne and Marcus Wallenberg Foundation, County Council of Dalarna, Dalarna University, and Swedish Heart-Lung Foundation (20120197). The computations were performed on resources provided by SNIC through Uppsala Multidisciplinary Center for Advanced Computational Science (UPPMAX) under Project b2011036. Genotyping was funded by the Wellcome Trust under awards WT064890 and WT086596. Analysis of genetic data was funded by the Wellcome Trust under awards WT098017 and WT090532. We thank the SNP&SEQ Technology Platform in Uppsala (www.genotyping.se) for excellent genotyping. Andrew P Morris is a Wellcome Trust Senior Fellow in Basic Biomedical Science under award WT098017.

PREVEND: The PREVEND genetics is supported by the Dutch Kidney Foundation (Grant E033), the EU project grant GENECURE (FP-6 LSHM CT 2006 037697), the National Institutes of Health (grant 2R01LM010098), The Netherlands organization for health research and development (NWO-Groot grant 175.010.2007.006, NWO VENI grant 916.761.70, ZonMw grant 90.700.441), and the Dutch Inter University Cardiology Institute Netherlands (ICIN). N. Verweij is supported by the Netherlands Heart Foundation (grant NHS2010B280).

PROSPER: The PROSPER study was supported by an investigator initiated grant obtained from Bristol-Myers Squibb. Prof. Dr. J. W. Jukema is an Established Clinical Investigator of the Netherlands Heart Foundation (grant 2001 D 032). Support for genotyping was provided by the seventh framework program of the European commission (grant 223004) and by the Netherlands Genomics Initiative (Netherlands Consortium for Healthy Aging grant 050-060-810).

QFS: The Quebec Family Study (QFS) was funded by multiple grants from the Medical Research Council of Canada and the Canadian Institutes for Health Research. This work was supported by a team grant from the Canadian Institutes for Health Research (FRCN-CCT-83028).

RS1: We thank the Genetic Laboratory of the Department of Internal Medicine of the Erasmus MC and specifically Pascal Arp, Mila Jhamai, Marijn Verkerk, and Carolina Medina-Gomez for their help in creating the GWAS database and the creation and analysis of imputed data. The dedication, commitment, and contribution of inhabitants, general practitioners, and pharmacists of the Ommoord district to the Rotterdam Study are gratefully acknowledged. We also thank the patients participating in the Erasmus Stroke Study.

RS2: The Rotterdam study is supported by the Erasmus MC and Erasmus University Rotterdam; the Netherlands Organisation for Scientific Research; the Netherlands Organisation for Health Research and Development (Zorg onderzoek Nederland Medische Wetenschappen); the Research Institute for Diseases in the Elderly; the Netherlands Genomics Initiative; the Ministry of Education, Culture and Science; the Ministry of Health, Welfare and Sports; the European Commission (Directorate-General XII); and the Municipality of Rotterdam.

RS3: None of the funders had any role in design and conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, review, or approval of this article.

SardiNIA: We thank the many individuals who generously participated in this study. This work was supported by Contract NO1-AG-1-2109 from the National Institute of Aging, and in part by a grant from the Italian Ministry of Economy and Finance to the CNR for the Project “FaReBio di Qualità” to F Cucca. The efforts of GR Abecasis were supported in part by contract 263-MA-410953 from the NIA to the University of Michigan and by research grant HG002651 and HL084729 from the NIH.

SCARFSHEEP: The work was supported by the European Commission (LSHM-CT- 2007- 037273), the Swedish Heart-Lung Foundation, the Swedish Research Council (8691, 09533), the Knut and Alice Wallenberg Foundation, the Foundation for Strategic Research, the Torsten and Ragnar Söderberg Foundation, the Strategic Cardiovascular Programme of Karolinska Institutet and the Stockholm County Council and the Stockholm County Council (560183); the Magnus Bergvall Foundation, Stiftelsen för Gamla Tjänarinnor, and the Tore Nilsson and Fredrik och Ingrid Thurings Foundations.

SHIP: SHIP is part of the Community Medicine Research net of the University of Greifswald, Germany, which is funded by the Federal Ministry of Education and Research (grants no. 01ZZ9603, 01ZZ0103, and 01ZZ0403), the Ministry of Cultural Affairs as well as the Social Ministry of the Federal State of Mecklenburg-West Pomerania, and the network ‘Greifswald Approach to Individualized Medicine (GANI_MED)’ funded by the Federal Ministry of Education and Research (grant 03IS2061A). Genome-wide data have been supported by the Federal Ministry of Education and Research (grant no. 03ZIK012) and a joint grant from Siemens Healthcare, Erlangen, Germany and the Federal State of Mecklenburg-West Pomerania. The University of Greifswald is a member of the Caché Campus program of the InterSystems GmbH.

SORBS: This work was supported by grants from the German Research Council (SFB- 1052 “Obesity mechanisms”, SPP 1629 TO 718/2-1), from the German Diabetes Association and from the DHFD (Diabetes Hilfs- und Forschungsfonds Deutschland). Inga Prokopenko was funded in part through the

European Community's Seventh Framework Programme (FP7/2007-2013), ENGAGE project, grant agreement HEALTH-F4-2007-201413.

SPT: Our chief acknowledgement is to the participants in these studies for their willingness to contribute. We also thank Nurses Orgen Brown and Diedre Thomas for assistance with recruitment as well as past and present Laboratory technologists and drivers at TMRU for their invaluable technical assistance. This work was supported by NIH Grants R01HL53353 and R01DK075787.

THISEAS: Recruitment for The Hellenic study of Interactions between SNPs and Eating in Atherosclerosis Susceptibility (THISEAS) study was partially funded by a research grant (PENED 2003) from the Greek General Secretary of Research and Technology. We thank all the dieticians and clinicians for their contribution to the project and the Genotyping Facility at the Wellcome Trust Sanger Institute for SNP typing. Analysis was partly supported by BHF grant (Deloukas) RG/14/5/30893 and the Barts Cardiovascular Biomedical Research Unit which is supported and funded by the National Institute for Health Research.

TRAILS: TRAILS (TRacking Adolescents' Individual Lives Survey) is a collaborative project involving various departments of the University Medical Center and University of Groningen, the Erasmus University Medical Center Rotterdam, the University of Utrecht, the Radboud Medical Center Nijmegen, and the Parnassia Bavo group, all in the Netherlands. TRAILS has been financially supported by grants from the Netherlands Organization for Scientific Research NWO (Medical Research Council program grant GB-MW 940-38-011; ZonMW Brainpower grant 100-001-004; ZonMw Risk Behavior and Dependence grant 60-60600-97-118; ZonMw Culture and Health grant 261-98-710; Social Sciences Council medium-sized investment grants GB-MaGW 480-01-006 and GB-MaGW 480-07-001; Social Sciences Council project grants GB-MaGW 452-04-314 and GB-MaGW 452-06-004; NWO large-sized investment grant 175.010.2003.005; NWO Longitudinal Survey and Panel Funding 481-08-013); the Dutch Ministry of Justice (WODC), the European Science Foundation (EuroSTRESS project FP-006), Biobanking and Biomolecular Resources Research Infrastructure BBMRI-NL (CP 32), the participating universities, and Accare Center for Child and Adolescent Psychiatry. We are grateful to all adolescents, their parents and teachers who participated in this research and to everyone who worked on this project and made it possible. Statistical analyses were carried out on the Genetic Cluster Computer (http://www.geneticcluster.org), which is financially supported by the Netherlands Scientific Organization (NWO 480-05-003) along with a supplement from the Dutch Brain Foundation.

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

WGHS: The WGHS is supported by HL043851 and HL080467 from the National Heart, Lung, and Blood Institute, and CA047988 and UM1CA182913 from the National Cancer Institute (NCI), the Donald W. Reynolds Foundation and the Fondation Leducq, with collaborative scientific support and funding for genotyping provided by Amgen.

WHI: Funding support for the "Epidemiology of putative genetic variants: The Women's Health Initiative" study is provided through the NHGRI PAGE program (U01HG007376, HG004790, and its NHGRI ARRA supplement). The WHI acknowledgment statement that you have is out of date. The

statement should be The WHI program is funded by the National Heart, Lung, and Blood Institute, National Institutes of Health, U.S. Department of Health and Human Services through contracts HHSN268201100046C, HHSN268201100001C, HHSN268201100002C, HHSN268201100003C, HHSN268201100004C, and HHSN271201100004C.

Whitehall: Dr. Kumari's and Professor Kivimaki's time on this manuscript was partially supported by the National Heart Lung and Blood Institute (NHLBI: HL36310). The Whitehall-II study has been supported by grants from the Medical Research Council (MRC); British Heart Foundation; Health and Safety Executive; Department of Health; National Institute on Aging (AG13196), US, NIH; Agency for Health Care Policy Research (HS06516); and the John D and Catherine T MacArthur Foundation Research Networks on Successful Midlife Development and Socioeconomic Status and Health.

YFS: The Young Finns Study has been financially supported by the Academy of Finland: grants 286284 (T.L.), 134309 (Eye), 126925, 121584, 124282, 129378 (Salve), 117787 (Gendi), and 41071 (Skidi); the Social Insurance Institution of Finland; Kuopio, Tampere and Turku University Hospital Medical Funds (grant X51001 for T.L.); Juho Vainio Foundation; Paavo Nurmi Foundation; Finnish Foundation of Cardiovascular Research (T.L.); Finnish Cultural Foundation; Tampere Tuberculosis Foundation (T.L.); Emil Aaltonen Foundation (T.L.); and Yrjö Jahnsson Foundation (T.L.). The expert technical assistance in the statistical analyses by Ville Aalto and Irina Lisinen is gratefully acknowledged.

BIBLIOGRAPHY

1. Locke, A.E. *et al.* Genetic studies of body mass index yield new insights for obesity biology. *Nature* **518**, 197-206 (2015).
2. Shungin, D. *et al.* New genetic loci link adipose and insulin biology to body fat distribution. *Nature* **518**, 187-96 (2015).
3. Pena, J.C., Duhalt, R., Navarrette, R. & Garcia Zozaya, J.L. [Periodic hemodialysis in the treatment of chronic renal insufficiency]. *Gac Med Mex* **98**, 150-67 (1968).
4. Shin, K., Wang, Q. & Margolis, B. PATJ regulates directional migration of mammalian epithelial cells. *EMBO Rep* **8**, 158-64 (2007).
5. Comuzzie, A.G. *et al.* Novel genetic loci identified for the pathophysiology of childhood obesity in the Hispanic population. *PLoS One* **7**, e51954 (2012).
6. Kathiresan, S. *et al.* Six new loci associated with blood low-density lipoprotein cholesterol, high-density lipoprotein cholesterol or triglycerides in humans. *Nat Genet* **40**, 189-97 (2008).
7. Zhang, Z. *et al.* Association of genetic loci with blood lipids in the Chinese population. *PLoS One* **6**, e27305 (2011).
8. Waterworth, D.M. *et al.* Genetic variants influencing circulating lipid levels and risk of coronary artery disease. *Arterioscler Thromb Vasc Biol* **30**, 2264-76 (2010).
9. Murphy, M.P. & LeVine, H., 3rd. Alzheimer's disease and the amyloid-beta peptide. *J Alzheimers Dis* **19**, 311-23 (2010).
10. Lee, Y.H. *et al.* Amyloid precursor protein expression is upregulated in adipocytes in obesity. *Obesity (Silver Spring)* **16**, 1493-500 (2008).
11. Puig, K.L., Floden, A.M., Adhikari, R., Golovko, M.Y. & Combs, C.K. Amyloid precursor protein and proinflammatory changes are regulated in brain and adipose tissue in a murine model of high fat diet-induced obesity. *PLoS One* **7**, e30378 (2012).

12. Martin-Campos, J.M. *et al.* Identification of a novel mutation in the ANGPTL3 gene in two families diagnosed of familial hypobetalipoproteinemia without APOB mutation. *Clin Chim Acta* **413**, 552-5 (2012).
13. Kathiresan, S. *et al.* Common variants at 30 loci contribute to polygenic dyslipidemia. *Nat Genet* **41**, 56-65 (2009).
14. Willer, C.J. *et al.* Newly identified loci that influence lipid concentrations and risk of coronary artery disease. *Nat Genet* **40**, 161-9 (2008).
15. Wong, R.C. *et al.* L1TD1 is a marker for undifferentiated human embryonic stem cells. *PLoS One* **6**, e19355 (2011).
16. Santos, M.C. *et al.* Embryonic Stem Cell-Related Protein L1TD1 Is Required for Cell Viability, Neurosphere Formation, and Chemoresistance in Medulloblastoma. *Stem Cells Dev* **24**, 2700-8 (2015).
17. Zhu, Y., Kakinuma, N., Wang, Y. & Kiyama, R. Kank proteins: a new family of ankyrin-repeat domain-containing proteins. *Biochim Biophys Acta* **1780**, 128-33 (2008).
18. Boyle, A.P. *et al.* Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res* **22**, 1790-7 (2012).
19. de la Rocha, A.M., Sampron, N., Alonso, M.M. & Matheu, A. Role of SOX family of transcription factors in central nervous system tumors. *Am J Cancer Res* **4**, 312-24 (2014).
20. Hempel, A. *et al.* Deletions and de novo mutations of SOX11 are associated with a neurodevelopmental disorder with features of Coffin-Siris syndrome. *J Med Genet* **53**, 152-62 (2016).
21. Ward, L.D. & Kellis, M. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Res* **40**, D930-4 (2012).
22. Zhuang, G. *et al.* Effects of cancer-associated EPHA3 mutations on lung cancer. *J Natl Cancer Inst* **104**, 1182-97 (2012).
23. Wood, L.D. *et al.* Somatic mutations of GUCY2F, EPHA3, and NTRK3 in human cancers. *Hum Mutat* **27**, 1060-1 (2006).
24. Lisabeth, E.M., Fernandez, C. & Pasquale, E.B. Cancer somatic mutations disrupt functions of the EphA3 receptor tyrosine kinase through multiple mechanisms. *Biochemistry* **51**, 1464-75 (2012).
25. Lee, D.J. *et al.* Multiple tumor-suppressor genes on chromosome 3p contribute to head and neck squamous cell carcinoma tumorigenesis. *Cancer Biol Ther* **10**, 689-93 (2010).
26. Lahtela, J. *et al.* A high-content cellular senescence screen identifies candidate tumor suppressors, including EPHA3. *Cell Cycle* **12**, 625-34 (2013).
27. Davis, S.D. *et al.* Clinical features of childhood primary ciliary dyskinesia by genotype and ultrastructural phenotype. *Am J Respir Crit Care Med* **191**, 316-24 (2015).
28. Oldenburg, A.R., Delbarre, E., Thiede, B., Vigouroux, C. & Collas, P. Deregulation of Fragile X-related protein 1 by the lipodystrophic lamin A p.R482W mutation elicits a myogenic gene expression program in preadipocytes. *Hum Mol Genet* **23**, 1151-62 (2014).
29. Sparkes, R., Patton, D. & Bernier, F. Cardiac features of a novel autosomal recessive dilated cardiomyopathic syndrome due to defective importation of mitochondrial protein. *Cardiol Young* **17**, 215-7 (2007).
30. Ojala, T. *et al.* New mutation of mitochondrial DNAJC19 causing dilated and noncompaction cardiomyopathy, anemia, ataxia, and male genital anomalies. *Pediatr Res* **72**, 432-7 (2012).
31. Davey, K.M. *et al.* Mutation of DNAJC19, a human homologue of yeast inner mitochondrial membrane co-chaperones, causes DCMA syndrome, a novel autosomal recessive Barth syndrome-like condition. *J Med Genet* **43**, 385-93 (2006).

32. Richter-Dennerlein, R. *et al.* DNAJC19, a mitochondrial cochaperone associated with cardiomyopathy, forms a complex with prohibitins to regulate cardiolipin remodeling. *Cell Metab* **20**, 158-71 (2014).
33. Perez-Lorenzo, R. *et al.* A tumor suppressor function for the lipid phosphatase INPP4B in melanocytic neoplasms. *J Invest Dermatol* **134**, 1359-68 (2014).
34. Fedele, C.G. *et al.* Inositol polyphosphate 4-phosphatase II regulates PI3K/Akt signaling and is lost in human basal-like breast cancers. *Proc Natl Acad Sci U S A* **107**, 22231-6 (2010).
35. Lopez, S.M. *et al.* Determinants of the tumor suppressor INPP4B protein and lipid phosphatase activities. *Biochem Biophys Res Commun* **440**, 277-82 (2013).
36. Mian, M.F., Pek, E.A., Mossman, K.L., Stampfli, M.R. & Ashkar, A.A. Exposure to cigarette smoke suppresses IL-15 generation and its regulatory NK cell functions in poly I:C-augmented human PBMCs. *Mol Immunol* **46**, 3108-16 (2009).
37. Liu, Y.Z. *et al.* Genome-wide association analyses suggested a novel mechanism for smoking behavior regulated by IL15. *Mol Psychiatry* **14**, 668-80 (2009).
38. Thorgeirsson, T.E. *et al.* Sequence variants at CHRNB3-CHRNA6 and CYP2A6 affect smoking behavior. *Nat Genet* **42**, 448-53 (2010).
39. Saccone, N.L. *et al.* Multiple independent loci at chromosome 15q25.1 affect smoking quantity: a meta-analysis and comparison with lung cancer and COPD. *PLoS Genet* **6**(2010).
40. Hancock, D.B. *et al.* Genome-wide meta-analysis reveals common splice site acceptor variant in CHRNA4 associated with nicotine dependence. *Transl Psychiatry* **5**, e651 (2015).
41. Thorgeirsson, T.E. *et al.* A variant associated with nicotine dependence, lung cancer and peripheral arterial disease. *Nature* **452**, 638-42 (2008).
42. Chen, L.S. *et al.* CHRNA5 risk variant predicts delayed smoking cessation and earlier lung cancer diagnosis--a meta-analysis. *J Natl Cancer Inst* **107**(2015).
43. Hallden, S. *et al.* Gene variance in the nicotinic receptor cluster (CHRNA5-CHRNA3-CHRNB4) predicts death from cardiopulmonary disease and cancer in smokers. *J Intern Med* (2015).
44. Tyrrell, J. *et al.* Genetic variation in the 15q25 nicotinic acetylcholine receptor gene cluster (CHRNA5-CHRNA3-CHRNB4) interacts with maternal self-reported smoking status during pregnancy to influence birth weight. *Hum Mol Genet* **21**, 5344-58 (2012).
45. Winslow, U.C., Rode, L. & Nordestgaard, B.G. High tobacco consumption lowers body weight: a Mendelian randomization study of the Copenhagen General Population Study. *Int J Epidemiol* **44**, 540-50 (2015).
46. Taylor, A.E. *et al.* Stratification by smoking status reveals an association of CHRNA5-A3-B4 genotype with body mass index in never smokers. *PLoS Genet* **10**, e1004799 (2014).
47. Morel, C. *et al.* Nicotine consumption is regulated by a human polymorphism in dopamine neurons. *Mol Psychiatry* **19**, 930-6 (2014).
48. Antolin-Fontes, B., Ables, J.L., Gorlich, A. & Ibanez-Tallon, I. The habenulo-interpeduncular pathway in nicotine aversion and withdrawal. *Neuropharmacology* **96**, 213-22 (2015).
49. van Setten, J. *et al.* Genome-wide association study of coronary and aortic calcification implicates risk loci for coronary artery disease and myocardial infarction. *Atherosclerosis* **228**, 400-5 (2013).
50. Schunkert, H. *et al.* Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. *Nat Genet* **43**, 333-8 (2011).
51. Reilly, M.P. *et al.* Identification of ADAMTS7 as a novel locus for coronary atherosclerosis and association of ABO with myocardial infarction in the presence of coronary atherosclerosis: two genome-wide association studies. *Lancet* **377**, 383-92 (2011).
52. Dichgans, M. *et al.* Shared genetic susceptibility to ischemic stroke and coronary artery disease: a genome-wide analysis of common variants. *Stroke* **45**, 24-36 (2014).

53. Wang, Z.C. *et al.* Genetic polymorphism of the kinesin-like protein KIF1B gene and the risk of hepatocellular carcinoma. *PLoS One* **8**, e62571 (2013).
54. Sukhatme, V.P. & Chan, B. Glycolytic cancer cells lacking 6-phosphogluconate dehydrogenase metabolize glucose to induce senescence. *FEBS Lett* **586**, 2389-95 (2012).
55. Ahn, J. *et al.* Identification of the avian RBP7 gene as a new adipose-specific gene and RBP7 promoter-driven GFP expression in adipose tissue of transgenic quail. *PLoS One* **10**, e0124768 (2015).
56. Borcherding, D.C. *et al.* Dopamine receptors in human adipocytes: expression and functions. *PLoS One* **6**, e25537 (2011).
57. Gieger, C. *et al.* New gene functions in megakaryopoiesis and platelet formation. *Nature* **480**, 201-8 (2011).
58. Nogueira, E. *et al.* SOK1 translocates from the Golgi to the nucleus upon chemical anoxia and induces apoptotic cell death. *J Biol Chem* **283**, 16248-58 (2008).
59. Davids, M.S. *et al.* STK25 is a candidate gene for pseudopseudohypoparathyroidism. *Genomics* **77**, 2-4 (2001).
60. Nerstedt, A. *et al.* Serine/threonine protein kinase 25 (STK25): a novel negative regulator of lipid and glucose metabolism in rodent and human skeletal muscle. *Diabetologia* **55**, 1797-807 (2012).
61. Cansby, E. *et al.* Increased expression of STK25 leads to impaired glucose utilization and insulin sensitivity in mice challenged with a high-fat diet. *FASEB J* **27**, 3660-71 (2013).
62. Semplici, F. *et al.* Human mutation within Per-Arnt-Sim (PAS) domain-containing protein kinase (PASK) causes basal insulin hypersecretion. *J Biol Chem* **286**, 44005-14 (2011).
63. Grose, J.H. & Rutter, J. The role of PAS kinase in PASsing the glucose signal. *Sensors (Basel)* **10**, 5668-82 (2010).
64. da Silva Xavier, G., Rutter, J. & Rutter, G.A. Involvement of Per-Arnt-Sim (PAS) kinase in the stimulation of preproinsulin and pancreatic duodenum homeobox 1 gene expression by glucose. *Proc Natl Acad Sci U S A* **101**, 8319-24 (2004).
65. Su, Z., Cox, A., Shen, Y., Stylianou, I.M. & Paigen, B. Farp2 and Stk25 are candidate genes for the HDL cholesterol locus on mouse chromosome 1. *Arterioscler Thromb Vasc Biol* **29**, 107-13 (2009).
66. Li, B.Z. *et al.* [Abnormal expression of bcl-10 protein in extranodal marginal zone B cell lymphoma of mucosa-associated lymphoid tissue lymphoma type]. *Zhonghua Bing Li Xue Za Zhi* **36**, 819-24 (2007).
67. Namekata, K. *et al.* Dock3 regulates BDNF-TrkB signaling for neurite outgrowth by forming a ternary complex with Elmo and RhoG. *Genes Cells* **17**, 688-97 (2012).
68. Liu, H., Tang, X. & Gong, L. Mesencephalic astrocyte-derived neurotrophic factor and cerebral dopamine neurotrophic factor: New endoplasmic reticulum stress response proteins. *Eur J Pharmacol* **750**, 118-22 (2015).
69. Padilla, J. *et al.* Transcriptome-wide RNA sequencing analysis of rat skeletal muscle feed arteries. II. Impact of exercise training in obesity. *J Appl Physiol (1985)* **116**, 1033-47 (2014).
70. Wood, A.R. *et al.* Defining the role of common variation in the genomic and biological architecture of adult human height. *Nat Genet* **46**, 1173-86 (2014).
71. Lango Allen, H. *et al.* Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature* **467**, 832-8 (2010).
72. Song, F. *et al.* Identification of a melanoma susceptibility locus and somatic mutation in TET2. *Carcinogenesis* **35**, 2097-101 (2014).
73. Consortium, G.T. The Genotype-Tissue Expression (GTEx) project. *Nat Genet* **45**, 580-5 (2013).

74. Sabater-Lleal, M. *et al.* Multiethnic meta-analysis of genome-wide association studies in >100 000 subjects identifies 23 fibrinogen-associated Loci but no strong evidence of a causal association between circulating fibrinogen and cardiovascular disease. *Circulation* **128**, 1310-24 (2013).
75. Kettunen, J. *et al.* Genome-wide association study identifies multiple loci influencing human serum metabolite levels. *Nat Genet* **44**, 269-76 (2012).
76. Inouye, M. *et al.* Novel Loci for metabolic networks and multi-tissue expression studies reveal genes for atherosclerosis. *PLoS Genet* **8**, e1002907 (2012).
77. Berndt, S.I. *et al.* Genome-wide meta-analysis identifies 11 new loci for anthropometric traits and provides insights into genetic architecture. *Nat Genet* **45**, 501-12 (2013).
78. Weedon, M.N. *et al.* Genome-wide association analysis identifies 20 loci that influence adult height. *Nat Genet* **40**, 575-83 (2008).
79. Soranzo, N. *et al.* Meta-analysis of genome-wide scans for human adult stature identifies novel Loci and associations with measures of skeletal frame size. *PLoS Genet* **5**, e1000445 (2009).
80. Lettre, G. *et al.* Identification of ten loci associated with height highlights new biological pathways in human growth. *Nat Genet* **40**, 584-91 (2008).
81. Gudbjartsson, D.F. *et al.* Many sequence variants affecting diversity of adult human height. *Nat Genet* **40**, 609-15 (2008).
82. Reiner, A.P. *et al.* Genome-wide association study of white blood cell count in 16,388 African Americans: the continental origins and genetic epidemiology network (COGENT). *PLoS Genet* **7**, e1002108 (2011).
83. Okada, Y. *et al.* Genetics of rheumatoid arthritis contributes to biology and drug discovery. *Nature* **506**, 376-81 (2014).
84. Raychaudhuri, S. *et al.* Common variants at CD40 and other loci confer risk of rheumatoid arthritis. *Nat Genet* **40**, 1216-23 (2008).
85. Albers, H.M. *et al.* Genetic variation in VTCN1 (B7-H4) is associated with course of disease in juvenile idiopathic arthritis. *Ann Rheum Dis* **73**, 1198-201 (2014).
86. Markula-Patjas, K.P. *et al.* High adiposity and serum leptin accompanied by altered bone turnover markers in severe juvenile idiopathic arthritis. *J Rheumatol* **41**, 2474-81 (2014).
87. Kim da, S., Lee, S.Y., Lee, J.H., Bae, Y.C. & Jung, J.S. MicroRNA-103a-3p controls proliferation and osteogenic differentiation of human adipose tissue-derived stromal cells. *Exp Mol Med* **47**, e172 (2015).
88. Sedlmeier, E.M. *et al.* Human placental transcriptome shows sexually dimorphic gene expression and responsiveness to maternal dietary n-3 long-chain polyunsaturated fatty acid intervention during pregnancy. *BMC Genomics* **15**, 941 (2014).
89. Gamazon, E.R. *et al.* SCAN: SNP and copy number annotation. *Bioinformatics* **26**, 259-62 (2010).
90. Li, M., Campbell, S. & McDermott, R. gamma-Glutamyltransferase, obesity, physical activity, and the metabolic syndrome in indigenous Australian adults. *Obesity (Silver Spring)* **17**, 809-13 (2009).
91. Iwasaki, T. *et al.* Hepatic fat content-independent association of the serum level of gamma-glutamyltransferase with visceral adiposity, but not subcutaneous adiposity. *Diabetes Res Clin Pract* **79**, e13-4 (2008).
92. Lee, D.S. *et al.* Gamma glutamyl transferase and metabolic syndrome, cardiovascular disease, and mortality risk: the Framingham Heart Study. *Arterioscler Thromb Vasc Biol* **27**, 127-33 (2007).
93. Atar, A.I. *et al.* Association between gamma-glutamyltransferase and coronary artery calcification. *Int J Cardiol* **167**, 1264-7 (2013).

94. Bradley, R.D. *et al.* Associations between gamma-glutamyltransferase (GGT) and biomarkers of atherosclerosis: the Multi-ethnic Study of Atherosclerosis (MESA). *Atherosclerosis* **233**, 387-93 (2014).
95. Zhu, C. *et al.* Association of serum gamma-glutamyltransferase with arterial stiffness in established coronary artery disease. *Angiology* **64**, 15-20 (2013).
96. Park, J.S. *et al.* Association between gamma-glutamyltransferase, adiponectin and arterial stiffness. *J Atheroscler Thromb* **19**, 90-7 (2012).
97. Volodarsky, M. *et al.* A deletion mutation in TMEM38B associated with autosomal recessive osteogenesis imperfecta. *Hum Mutat* **34**, 582-6 (2013).
98. Shaheen, R. *et al.* Study of autosomal recessive osteogenesis imperfecta in Arabia reveals a novel locus defined by TMEM38B mutation. *J Med Genet* **49**, 630-5 (2012).
99. Rubinato, E. *et al.* A novel deletion mutation involving TMEM38B in a patient with autosomal recessive osteogenesis imperfecta. *Gene* **545**, 290-2 (2014).
100. Perry, J.R. *et al.* Meta-analysis of genome-wide association data identifies two loci influencing age at menarche. *Nat Genet* **41**, 648-50 (2009).
101. Perry, J.R. *et al.* Parent-of-origin-specific allelic associations among 106 genomic loci for age at menarche. *Nature* **514**, 92-7 (2014).
102. Elks, C.E. *et al.* Thirty new loci for age at menarche identified by a meta-analysis of genome-wide association studies. *Nat Genet* **42**, 1077-85 (2010).
103. Fox, C.S. *et al.* Genome-wide association for abdominal subcutaneous and visceral adipose reveals a novel locus for visceral fat in women. *PLoS Genet* **8**, e1002695 (2012).
104. Melzer, D. *et al.* A genome-wide association study identifies protein quantitative trait loci (pQTLs). *PLoS Genet* **4**, e1000072 (2008).
105. Ismail, S., Schaffer, A.E., Rosti, R.O., Gleeson, J.G. & Zaki, M.S. Novel mutation in the fukutin gene in an Egyptian family with Fukuyama congenital muscular dystrophy and microcephaly. *Gene* **539**, 279-82 (2014).
106. Costa, C. *et al.* A Portuguese case of Fukuyama congenital muscular dystrophy caused by a multi-exonic duplication in the fukutin gene. *Neuromuscul Disord* **23**, 557-61 (2013).
107. Welch, H.C. *et al.* P-Rex1, a PtdIns(3,4,5)P3- and Gbetagamma-regulated guanine-nucleotide exchange factor for Rac. *Cell* **108**, 809-21 (2002).
108. Kimura, S., Sato, K., Banno, Y., Nagase, T. & Ueda, H. The importance of interaction with membrane lipids through the pleckstrin homology domain of the guanine nucleotide exchange factor for rho family small guanosine triphosphatase, FLJ00018. *Biol Pharm Bull* **36**, 1204-7 (2013).
109. Damoulakis, G. *et al.* P-Rex1 directly activates RhoG to regulate GPCR-driven Rac signalling and actin polarity in neutrophils. *J Cell Sci* **127**, 2589-600 (2014).
110. Xue, R. *et al.* Clonal analyses and gene profiling identify genetic biomarkers of the thermogenic potential of human brown and white preadipocytes. *Nat Med* **21**, 760-8 (2015).
111. Kuhn, R.M., Haussler, D. & Kent, W.J. The UCSC genome browser and associated tools. *Brief Bioinform* **14**, 144-61 (2013).
112. Mick, E. *et al.* Genome-wide association study of proneness to anger. *PLoS One* **9**, e87257 (2014).
113. Micu, I. *et al.* NMDA receptors mediate calcium accumulation in myelin during chemical ischaemia. *Nature* **439**, 988-92 (2006).
114. Zhong, H.J. *et al.* Functional polymorphisms of the glutamate receptor N-methyl D-aspartate 2A gene are associated with heroin addiction. *Genet Mol Res* **13**, 8714-21 (2014).
115. Turner, S.J. *et al.* GRIN2A: an aptly named gene for speech dysfunction. *Neurology* **84**, 586-93 (2015).

116. Liu, R. *et al.* Correlation of functional GRIN2A gene promoter polymorphisms with schizophrenia and serum D-serine levels. *Gene* **568**, 25-30 (2015).
117. Leuba, G. *et al.* Pathological reorganization of NMDA receptors subunits and postsynaptic protein PSD-95 distribution in Alzheimer's disease. *Curr Alzheimer Res* **11**, 86-96 (2014).
118. Lemke, J.R. *et al.* Mutations in GRIN2A cause idiopathic focal epilepsy with rolandic spikes. *Nat Genet* **45**, 1067-72 (2013).
119. DeVries, S.P. & Patel, A.D. Two patients with a GRIN2A mutation and childhood-onset epilepsy. *Pediatr Neurol* **49**, 482-5 (2013).
120. Carvill, G.L. *et al.* GRIN2A mutations cause epilepsy-aphasia spectrum disorders. *Nat Genet* **45**, 1073-6 (2013).
121. Aouizerat, B.E. *et al.* GWAS for discovery and replication of genetic loci associated with sudden cardiac arrest in patients with coronary artery disease. *BMC Cardiovasc Disord* **11**, 29 (2011).
122. McGue, M. *et al.* A genome-wide association study of behavioral disinhibition. *Behav Genet* **43**, 363-73 (2013).
123. Greliche, N. *et al.* A genome-wide search for common SNP x SNP interactions on the risk of venous thrombosis. *BMC Med Genet* **14**, 36 (2013).
124. Volkmer, E. & Karnitz, L.M. Human homologs of Schizosaccharomyces pombe rad1, hus1, and rad9 form a DNA damage-responsive protein complex. *J Biol Chem* **274**, 567-70 (1999).
125. Marathi, U.K. *et al.* RAD1, a human structural homolog of the Schizosaccharomyces pombe RAD1 cell cycle checkpoint gene. *Genomics* **54**, 344-7 (1998).
126. Chaudhuri, S.P. *et al.* Activation of S phase checkpoint by cigarette smoke extract in Schizosaccharomyces pombe. *Yeast* **22**, 1223-38 (2005).
127. Suhre, K. *et al.* A genome-wide association study of metabolic traits in human urine. *Nat Genet* **43**, 565-9 (2011).
128. Seppala, I. *et al.* Genome-wide association study on dimethylarginines reveals novel AGXT2 variants associated with heart rate variability but not with overall mortality. *Eur Heart J* **35**, 524-31 (2014).
129. Rueedi, R. *et al.* Genome-wide association study of metabolic traits reveals novel gene-metabolite-disease links. *PLoS Genet* **10**, e1004132 (2014).
130. Nicholson, G. *et al.* A genome-wide metabolic QTL analysis in Europeans implicates two loci shaped by recent positive selection. *PLoS Genet* **7**, e1002270 (2011).
131. Arnett, D.K. *et al.* Genome-wide association study identifies single-nucleotide polymorphism in KCNB1 associated with left ventricular mass in humans: the HyperGEN Study. *BMC Med Genet* **10**, 43 (2009).
132. Warnatz, H.J. *et al.* The BTB and CNC homology 1 (BACH1) target genes are involved in the oxidative stress response and in control of the cell cycle. *J Biol Chem* **286**, 23521-32 (2011).
133. Kim, J. *et al.* Functional genomic screen for modulators of ciliogenesis and cilium length. *Nature* **464**, 1048-51 (2010).
134. Jostins, L. *et al.* Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature* **491**, 119-24 (2012).
135. Linden, R. *et al.* Physiology of the prion protein. *Physiol Rev* **88**, 673-728 (2008).
136. Vanik, D.L. & Surewicz, W.K. Disease-associated F198S mutation increases the propensity of the recombinant prion protein for conformational conversion to scrapie-like form. *J Biol Chem* **277**, 49065-70 (2002).
137. Steele, A.D., Emsley, J.G., Ozdinler, P.H., Lindquist, S. & Macklis, J.D. Prion protein (PrPc) positively regulates neural precursor proliferation during developmental and adult mammalian neurogenesis. *Proc Natl Acad Sci U S A* **103**, 3416-21 (2006).
138. Mouillet-Richard, S. *et al.* Signal transduction through prion protein. *Science* **289**, 1925-8 (2000).

139. Telling, G.C. *et al.* Evidence for the conformation of the pathologic isoform of the prion protein enciphering and propagating prion diversity. *Science* **274**, 2079-82 (1996).
140. Poulter, M. *et al.* Inherited prion disease with 144 base pair gene insertion. 1. Genealogical and molecular studies. *Brain* **115 (Pt 3)**, 675-85 (1992).
141. Moore, R.C. *et al.* Huntington disease phenocopy is a familial prion disease. *Am J Hum Genet* **69**, 1385-8 (2001).
142. Mead, S. *et al.* A novel protective prion protein variant that colocalizes with kuru exposure. *N Engl J Med* **361**, 2056-65 (2009).
143. Haik, S. *et al.* Striking PrPsc heterogeneity in inherited prion diseases with the D178N mutation. *Ann Neurol* **56**, 909-10; author reply 910-1 (2004).
144. Collinge, J. *et al.* Diagnosis of Gerstmann-Straussler syndrome in familial dementia with prion protein gene analysis. *Lancet* **2**, 15-7 (1989).
145. Basset-Leobon, C. *et al.* Familial Creutzfeldt-Jakob disease with an R208H-129V haplotype and Kuru plaques. *Arch Neurol* **63**, 449-52 (2006).
146. Redecke, L. *et al.* Structural characterization of beta-sheeted oligomers formed on the pathway of oxidative prion protein aggregation in vitro. *J Struct Biol* **157**, 308-20 (2007).
147. United States. Public Health Service. Office of the Surgeon General. *How tobacco smoke causes disease : the biology and behavioral basis for smoking-attributable disease : a report of the Surgeon General*, xv, 704 p (U.S. Dept. of Health and Human Services, Public Health Service, Rockville, MD; Washington, DC, 2010).
148. Bernhard, D., Rossmann, A. & Wick, G. Metals in cigarette smoke. *IUBMB Life* **57**, 805-9 (2005).
149. Moore, R.C. *et al.* Ataxia in prion protein (PrP)-deficient mice is associated with upregulation of the novel PrP-like protein doppel. *J Mol Biol* **292**, 797-817 (1999).
150. Croes, E.A. *et al.* Polymorphisms in the prion protein gene and in the doppel gene increase susceptibility for Creutzfeldt-Jakob disease. *Eur J Hum Genet* **12**, 389-94 (2004).
151. Schroder, B. *et al.* Polymorphisms within the prion-like protein gene (Prnd) and their implications in human prion diseases, Alzheimer's disease and other neurological disorders. *Hum Genet* **109**, 319-25 (2001).
152. Savini, I., Rossi, A., Pierro, C., Avigliano, L. & Catani, M.V. SVCT1 and SVCT2: key proteins for vitamin C uptake. *Amino Acids* **34**, 347-55 (2008).
153. Catania, A.S., Barros, C.R. & Ferreira, S.R. [Vitamins and minerals with antioxidant properties and cardiometabolic risk: controversies and perspectives]. *Arq Bras Endocrinol Metabol* **53**, 550-9 (2009).
154. Babaev, V.R. *et al.* Combined vitamin C and vitamin E deficiency worsens early atherosclerosis in apolipoprotein E-deficient mice. *Arterioscler Thromb Vasc Biol* **30**, 1751-7 (2010).
155. Hediger, M.A. New view at C. *Nat Med* **8**, 445-6 (2002).
156. Uhl, G.R., Drgon, T., Li, C.Y., Johnson, C. & Liu, Q.R. Smoking and smoking cessation in disadvantaged women: assessing genetic contributions. *Drug Alcohol Depend* **104 Suppl 1**, S58-63 (2009).
157. Rose, J.E., Behm, F.M., Drgon, T., Johnson, C. & Uhl, G.R. Personalized smoking cessation: interactions between nicotine dose, dependence and quit-success genotype score. *Mol Med* **16**, 247-53 (2010).
158. Zhao, L., Bracken, M.B. & DeWan, A.T. Genome-wide association study of pre-eclampsia detects novel maternal single nucleotide polymorphisms and copy-number variants in subsets of the Hyperglycemia and Adverse Pregnancy Outcome (HAPO) study cohort. *Ann Hum Genet* **77**, 277-87 (2013).
159. Caporaso, N. *et al.* Genome-wide and candidate gene association study of cigarette smoking behaviors. *PLoS One* **4**, e4653 (2009).

160. Oldenborg, P.A. *et al.* Role of CD47 as a marker of self on red blood cells. *Science* **288**, 2051-4 (2000).
161. Lindberg, F.P. *et al.* Decreased resistance to bacterial infection and granulocyte defects in IAP-deficient mice. *Science* **274**, 795-8 (1996).
162. Finley, M.J., Clark, K.A., Alferiev, I.S., Levy, R.J. & Stachelek, S.J. Intracellular signaling mechanisms associated with CD47 modified surfaces. *Biomaterials* **34**, 8640-9 (2013).
163. Wiewiora, M., Piecuch, J., Sedek, L., Mazur, B. & Sosada, K. The effects of obesity on CD47 expression in erythrocytes. *Cytometry B Clin Cytom* (2015).
164. Maimaitiyiming, H., Norman, H., Zhou, Q. & Wang, S. CD47 deficiency protects mice from diet-induced obesity and improves whole body glucose tolerance and insulin sensitivity. *Sci Rep* **5**, 8846 (2015).
165. Mozzetta, C., Pontis, J. & Ait-Si-Ali, S. Functional Crosstalk Between Lysine Methyltransferases on Histone Substrates: The Case of G9A/GLP and Polycomb Repressive Complex 2. *Antioxid Redox Signal* **22**, 1365-81 (2015).
166. Balakrishnan, V.S. *et al.* Cytokine gene polymorphisms in hemodialysis patients: association with comorbidity, functionality, and serum albumin. *Kidney Int* **65**, 1449-60 (2004).
167. Simmonds, R.E. & Foxwell, B.M. Signalling, inflammation and arthritis: NF-kappaB and its relevance to arthritis and inflammation. *Rheumatology (Oxford)* **47**, 584-90 (2008).
168. Nanes, M.S. Tumor necrosis factor-alpha: molecular and cellular mechanisms in skeletal pathology. *Gene* **321**, 1-15 (2003).
169. Boyce, B.F., Schwarz, E.M. & Xing, L. Osteoclast precursors: cytokine-stimulated immunomodulators of inflammatory bone disease. *Curr Opin Rheumatol* **18**, 427-32 (2006).
170. Lu, X. *et al.* Identification of the homeobox protein Prx1 (MHox, Prrx-1) as a regulator of osterix expression and mediator of tumor necrosis factor alpha action in osteoblast differentiation. *J Bone Miner Res* **26**, 209-19 (2011).
171. Clark, I.A. How TNF was recognized as a key mechanism of disease. *Cytokine Growth Factor Rev* **18**, 335-43 (2007).
172. Popko, K. *et al.* Proinflammatory cytokines Il-6 and TNF-alpha and the development of inflammation in obese subjects. *Eur J Med Res* **15 Suppl 2**, 120-2 (2010).
173. Gomez-Uriz, A.M. *et al.* Epigenetic patterns of two gene promoters (TNF-alpha and PON) in stroke considering obesity condition and dietary intake. *J Physiol Biochem* **70**, 603-14 (2014).
174. Elahi, M.M., Gilmour, A., Matata, B.M. & Mastana, S.S. A variant of position -308 of the Tumour necrosis factor alpha gene promoter and the risk of coronary heart disease. *Heart Lung Circ* **17**, 14-8 (2008).
175. Panasevich, S. *et al.* Interaction between early maternal smoking and variants in TNF and GSTP1 in childhood wheezing. *Clin Exp Allergy* **40**, 458-67 (2010).
176. Soler Artigas, M. *et al.* Genome-wide association and large-scale follow up identifies 16 new loci influencing lung function. *Nat Genet* **43**, 1082-90 (2011).
177. Natori, Y., Ohkura, N., Nasui, M., Atsumi, G. & Kihara-Negishi, F. Acidic sialidase activity is aberrant in obese and diabetic mice. *Biol Pharm Bull* **36**, 1027-31 (2013).
178. Guo, H. *et al.* Variations in HSPA1B at 6p21.3 are associated with lung cancer risk and prognosis in Chinese populations. *Cancer Res* **71**, 7576-86 (2011).
179. Moreno-Navarrete, J.M. *et al.* Complement factor H is expressed in adipose tissue in association with insulin resistance. *Diabetes* **59**, 200-9 (2010).
180. Kraja, A.T. *et al.* Pleiotropic genes for metabolic syndrome and inflammation. *Mol Genet Metab* **112**, 317-38 (2014).
181. Dahlman, I. *et al.* Functional annotation of the human fat cell secretome. *Arch Physiol Biochem* **118**, 84-91 (2012).

182. Arason, G.J. *et al.* Smoking and a complement gene polymorphism interact in promoting cardiovascular disease morbidity and mortality. *Clin Exp Immunol* **149**, 132-8 (2007).
183. Sesti, G. *et al.* A functional variant of the dimethylarginine dimethylaminohydrolase-2 gene is associated with chronic kidney disease. *Atherosclerosis* **231**, 141-4 (2013).
184. Andreozzi, F. *et al.* A functional variant of the dimethylarginine dimethylaminohydrolase-2 gene is associated with insulin sensitivity. *PLoS One* **7**, e36224 (2012).
185. Kappelle, P.J., Lambert, G., Dahlback, B., Nielsen, L.B. & Dullaart, R.P. Relationship of plasma apolipoprotein M with proprotein convertase subtilisin-kexin type 9 levels in non-diabetic subjects. *Atherosclerosis* **214**, 492-4 (2011).
186. Burkart, K.M. *et al.* APOM and high-density lipoprotein cholesterol are associated with lung function and per cent emphysema. *Eur Respir J* **43**, 1003-17 (2014).
187. Hudson, B.I. *et al.* Serum levels of soluble receptor for advanced glycation end-products and metabolic syndrome: the Northern Manhattan Study. *Metabolism* **63**, 1125-30 (2014).
188. Jiao, L. *et al.* Evidence that serum levels of the soluble receptor for advanced glycation end products are inversely associated with pancreatic cancer risk: a prospective study. *Cancer Res* **71**, 3582-9 (2011).
189. Kankova, K., Stejskalova, A., Hertlova, M. & Znojil, V. Haplotype analysis of the RAGE gene: identification of a haplotype marker for diabetic nephropathy in type 2 diabetes mellitus. *Nephrol Dial Transplant* **20**, 1093-102 (2005).
190. Nawroth, P., Bierhaus, A., Marrero, M., Yamamoto, H. & Stern, D.M. Atherosclerosis and restenosis: is there a role for RAGE? *Curr Diab Rep* **5**, 11-6 (2005).
191. Gohda, T. *et al.* Increased serum endogenous secretory receptor for advanced glycation end-product (esRAGE) levels in type 2 diabetic patients with decreased renal function. *Diabetes Res Clin Pract* **81**, 196-201 (2008).
192. Norata, G.D. *et al.* Circulating soluble receptor for advanced glycation end products is inversely associated with body mass index and waist/hip ratio in the general population. *Nutr Metab Cardiovasc Dis* **19**, 129-34 (2009).
193. Fukui, M. *et al.* The serum concentration of allograft inflammatory factor-1 is correlated with metabolic parameters in healthy subjects. *Metabolism* **61**, 1021-5 (2012).
194. Rouskas, K. *et al.* Common variants in FTO, MC4R, TMEM18, PRL, AIF1, and PCSK1 show evidence of association with adult obesity in the Greek population. *Obesity (Silver Spring)* **20**, 389-95 (2012).
195. Wang, Y. *et al.* Predictive role of multilocus genetic polymorphisms in cardiovascular disease and inflammation-related genes on chronic kidney disease in Type 2 diabetes--an 8-year prospective cohort analysis of 1163 patients. *Nephrol Dial Transplant* **27**, 190-6 (2012).
196. Mak, J.C. *et al.* Polymorphisms in the IL-4, IL-4 receptor alpha chain, TNF-alpha, and lymphotoxin-alpha genes and risk of asthma in Hong Kong Chinese adults. *Int Arch Allergy Immunol* **144**, 114-22 (2007).
197. Stolk, L. *et al.* Loci at chromosomes 13, 19 and 20 influence age at natural menopause. *Nat Genet* **41**, 645-7 (2009).
198. Levy, D. *et al.* Genome-wide association identifies OBFC1 as a locus involved in human leukocyte telomere biology. *Proc Natl Acad Sci U S A* **107**, 9293-8 (2010).
199. Stanescu, H.C. *et al.* Risk HLA-DQA1 and PLA(2)R1 alleles in idiopathic membranous nephropathy. *N Engl J Med* **364**, 616-26 (2011).
200. Kim, Y.J. *et al.* A genome-wide association study identified new variants associated with the risk of chronic hepatitis B. *Hum Mol Genet* **22**, 4233-8 (2013).

201. Walt, A.J., Bouwman, D.L., Weaver, D.W. & Sachs, R.J. The impact of technology on the management of pancreatic pseudocyst. Fifth annual Samuel Jason Mixter Lecture. *Arch Surg* **125**, 759-63 (1990).
202. Wang, Q. *et al.* MicroRNA-377 is up-regulated and can lead to increased fibronectin production in diabetic nephropathy. *FASEB J* **22**, 4126-35 (2008).
203. McDonough, C.W. *et al.* A genome-wide association study for diabetic nephropathy genes in African Americans. *Kidney Int* **79**, 563-72 (2011).
204. Lan, X. *et al.* Identification of differentially expressed genes related to metabolic syndrome induced with high-fat diet in E3 rats. *Exp Biol Med (Maywood)* **240**, 235-41 (2015).
205. Liu, Y. *et al.* Serum methylation levels of TAC1. SEPT9 and EYA4 as diagnostic markers for early colorectal cancers: a pilot study. *Biomarkers* **18**, 399-405 (2013).
206. Kisiel, J.B., Garrity-Park, M.M., Taylor, W.R., Smyrk, T.C. & Ahlquist, D.A. Methylated eyes absent 4 (EYA4) gene promotor in non-neoplastic mucosa of ulcerative colitis patients with colorectal cancer: evidence for a field effect. *Inflamm Bowel Dis* **19**, 2079-83 (2013).
207. Lovf, M. *et al.* A novel transcript, VNN1-AB, as a biomarker for colorectal cancer. *Int J Cancer* **135**, 2077-84 (2014).
208. Huang, H. *et al.* Novel blood biomarkers of pancreatic cancer-associated diabetes mellitus identified by peripheral blood-based gene expression profiles. *Am J Gastroenterol* **105**, 1661-9 (2010).
209. Soubeyrand, S., Naing, T., Martinuk, A. & McPherson, R. ERK1/2 regulates hepatocyte Trib1 in response to mitochondrial dysfunction. *Biochim Biophys Acta* **1833**, 3405-14 (2013).
210. Zhou, L. *et al.* A genome wide association study identifies common variants associated with lipid levels in the Chinese population. *PLoS One* **8**, e82420 (2013).
211. Kim, Y.J. *et al.* Large-scale genome-wide association studies in East Asians identify new genetic loci influencing metabolic traits. *Nat Genet* **43**, 990-5 (2011).
212. Kraja, A.T. *et al.* A bivariate genome-wide approach to metabolic syndrome: STAMPEED consortium. *Diabetes* **60**, 1329-39 (2011).
213. Ko, A. *et al.* Amerindian-specific regions under positive selection harbour new lipid variants in Latinos. *Nat Commun* **5**, 3983 (2014).
214. Kamatani, Y. *et al.* Genome-wide association study of hematological and biochemical traits in a Japanese population. *Nat Genet* **42**, 210-5 (2010).
215. Global Lipids Genetics, C. *et al.* Discovery and refinement of loci associated with lipid levels. *Nat Genet* **45**, 1274-83 (2013).
216. Aulchenko, Y.S. *et al.* Loci influencing lipid levels and coronary heart disease risk in 16 European population cohorts. *Nat Genet* **41**, 47-55 (2009).
217. Teslovich, T.M. *et al.* Biological, clinical and population relevance of 95 loci for blood lipids. *Nature* **466**, 707-13 (2010).
218. Dastani, Z. *et al.* Novel loci for adiponectin levels and their influence on type 2 diabetes and metabolic traits: a multi-ethnic meta-analysis of 45,891 individuals. *PLoS Genet* **8**, e1002607 (2012).
219. Chambers, J.C. *et al.* Genome-wide association study identifies loci influencing concentrations of liver enzymes in plasma. *Nat Genet* **43**, 1131-8 (2011).
220. Barrett, J.C. *et al.* Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. *Nat Genet* **40**, 955-62 (2008).
221. Franke, A. *et al.* Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. *Nat Genet* **42**, 1118-25 (2010).

222. Yu, B. *et al.* Genome-wide association study of a heart failure related metabolomic profile among African Americans in the Atherosclerosis Risk in Communities (ARIC) study. *Genet Epidemiol* **37**, 840-5 (2013).
223. McArdle, P.F. *et al.* Association of a common nonsynonymous variant in GLUT9 with serum uric acid levels in old order amish. *Arthritis Rheum* **58**, 2874-81 (2008).
224. Wakil, S.M. *et al.* Association of a mutation in LACC1 with a monogenic form of systemic juvenile idiopathic arthritis. *Arthritis Rheumatol* **67**, 288-95 (2015).
225. Wang, J. *et al.* Genome-wide association analysis implicates the involvement of eight loci with response to tocilizumab for the treatment of rheumatoid arthritis. *Pharmacogenomics J* **13**, 235-41 (2013).
226. Lyons, P.A. *et al.* Genetically distinct subsets within ANCA-associated vasculitis. *N Engl J Med* **367**, 214-23 (2012).
227. Ahmeti, K.B. *et al.* Age of onset of amyotrophic lateral sclerosis is modulated by a locus on 1p34.1. *Neurobiol Aging* **34**, 357 e7-19 (2013).
228. Zheng, J.S. *et al.* Genome-wide contribution of genotype by environment interaction to variation of diabetes-related traits. *PLoS One* **8**, e77442 (2013).
229. Sokolik, R. *et al.* Significance of association of HLA-C and HLA-E with psoriatic arthritis. *Hum Immunol* **75**, 1188-91 (2014).
230. Smigoc Schweiger, D. *et al.* Genetic risk for co-occurrence of type 1 diabetes and celiac disease is modified by HLA-C and killer immunoglobulin-like receptors. *Tissue Antigens* **84**, 471-8 (2014).
231. Matsuzaka, Y. *et al.* Identification of novel candidate genes in the diffuse panbronchiolitis critical region of the class I human MHC. *Immunogenetics* **54**, 301-9 (2002).
232. Hancock, D.B. *et al.* Genome-wide joint meta-analysis of SNP and SNP-by-smoking interaction identifies novel loci for pulmonary function. *PLoS Genet* **8**, e1003098 (2012).
233. Thorleifsson, G. *et al.* Genome-wide association yields new sequence variants at seven loci that associate with measures of obesity. *Nat Genet* **41**, 18-24 (2009).
234. Zhang, X. *et al.* Synthesis of 53 tissue and cell line expression QTL datasets reveals master eQTLs. *BMC Genomics* **15**, 532 (2014).
235. Goring, H.H. *et al.* Discovery of expression QTLs using large-scale transcriptional profiling in human lymphocytes. *Nat Genet* **39**, 1208-16 (2007).
236. Idaghdour, Y. *et al.* Geographical genomics of human leukocyte gene expression variation in southern Morocco. *Nat Genet* **42**, 62-7 (2010).
237. Heap, G.A. *et al.* Complex nature of SNP genotype effects on gene expression in primary human leucocytes. *BMC Med Genomics* **2**, 1 (2009).
238. Emilsson, V. *et al.* Genetics of gene expression and its effect on disease. *Nature* **452**, 423-8 (2008).
239. Fehrmann, R.S. *et al.* Trans-eQTLs reveal that independent genetic variants associated with a complex phenotype converge on intermediate genes, with a major role for the HLA. *PLoS Genet* **7**, e1002197 (2011).
240. Mehta, D. *et al.* Impact of common regulatory single-nucleotide variants on gene expression profiles in whole blood. *Eur J Hum Genet* **21**, 48-54 (2013).
241. Zhernakova, D.V. *et al.* DeepSAGE reveals genetic variants associated with alternative polyadenylation and expression of coding and non-coding transcripts. *PLoS Genet* **9**, e1003594 (2013).
242. Sasayama, D. *et al.* Identification of single nucleotide polymorphisms regulating peripheral blood mRNA expression with genome-wide significance: an eQTL study in the Japanese population. *PLoS One* **8**, e54967 (2013).

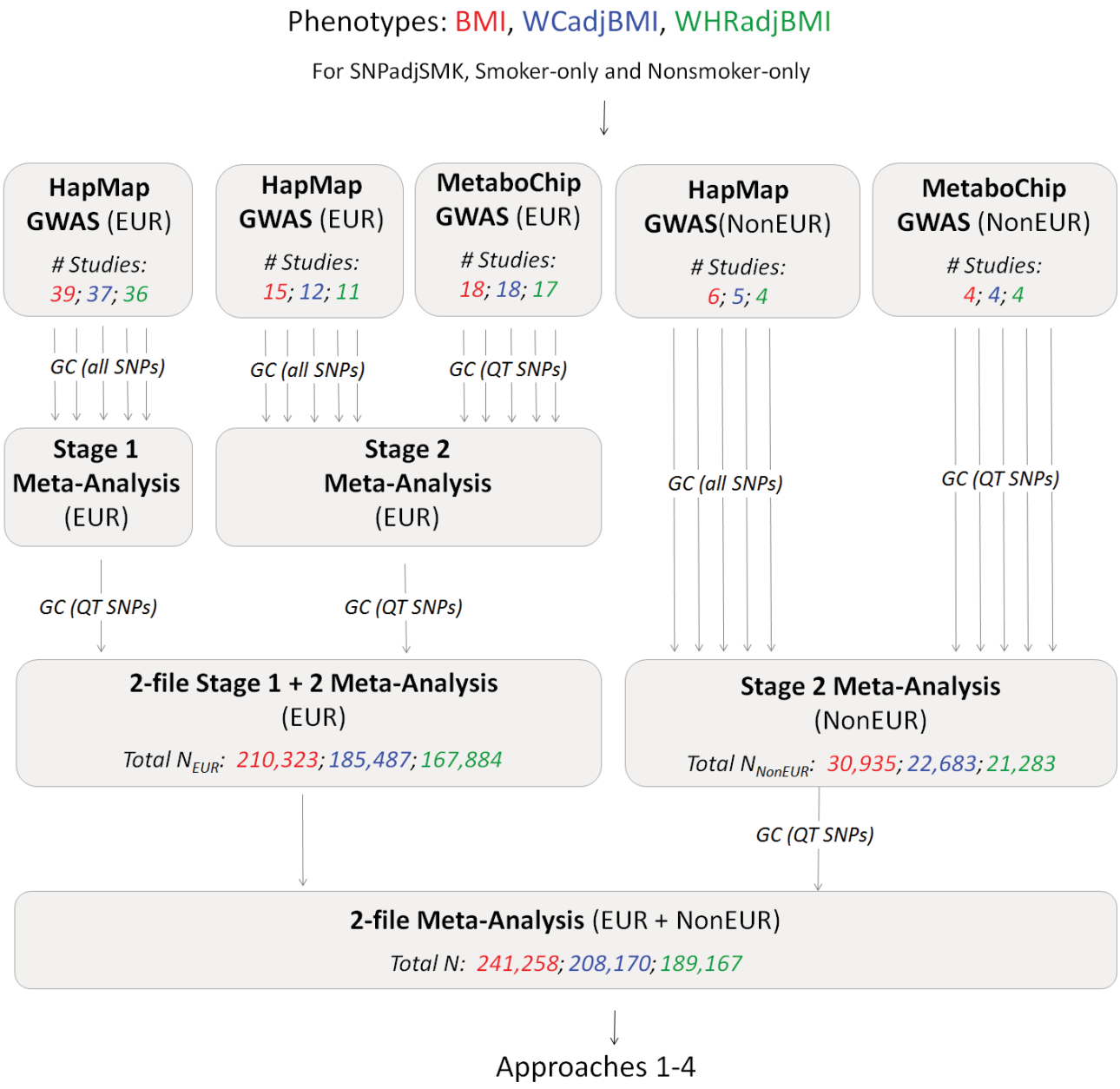
243. Landmark-Hoyvik, H. *et al.* Genome-wide association study in breast cancer survivors reveals SNPs associated with gene expression of genes belonging to MHC class I and II. *Genomics* **102**, 278-87 (2013).
244. Westra, H.J. *et al.* Systematic identification of trans eQTLs as putative drivers of known disease associations. *Nat Genet* **45**, 1238-43 (2013).
245. van Eijk, K.R. *et al.* Genetic analysis of DNA methylation and gene expression levels in whole blood of healthy human subjects. *BMC Genomics* **13**, 636 (2012).
246. Battle, A. *et al.* Characterizing the genetic basis of transcriptome diversity through RNA-sequencing of 922 individuals. *Genome Res* **24**, 14-24 (2014).
247. Benton, M.C. *et al.* Mapping eQTLs in the Norfolk Island genetic isolate identifies candidate genes for CVD risk traits. *Am J Hum Genet* **93**, 1087-99 (2013).
248. Narahara, M. *et al.* Large-scale East-Asian eQTL mapping reveals novel candidate genes for LD mapping and the genomic landscape of transcriptional effects of sequence variants. *PLoS One* **9**, e100924 (2014).
249. Quinlan, J. *et al.* Genomic architecture of sickle cell disease in West African children. *Front Genet* **5**, 26 (2014).
250. Wright, F.A. *et al.* Heritability and genomics of gene expression in peripheral blood. *Nat Genet* **46**, 430-7 (2014).
251. Schramm, K. *et al.* Mapping the genetic architecture of gene regulation in whole blood. *PLoS One* **9**, e93844 (2014).
252. Lock, E.F. *et al.* Joint eQTL assessment of whole blood and dura mater tissue from individuals with Chiari type I malformation. *BMC Genomics* **16**, 11 (2015).
253. Powell, J.E. *et al.* The Brisbane Systems Genetics Study: genetical genomics meets complex trait genetics. *PLoS One* **7**, e35430 (2012).
254. Pierce, B.L. *et al.* Mediation analysis demonstrates that trans-eQTLs are often explained by cis-mediation: a genome-wide analysis among 1,800 South Asians. *PLoS Genet* **10**, e1004818 (2014).
255. Chen, W. *et al.* Expression quantitative trait loci (eQTL) mapping in Puerto Rican children. *PLoS One* **10**, e0122464 (2015).
256. Foroughi Asl, H. *et al.* Expression quantitative trait Loci acting across multiple tissues are enriched in inherited risk for coronary artery disease. *Circ Cardiovasc Genet* **8**, 305-15 (2015).
257. Dixon, A.L. *et al.* A genome-wide association study of global gene expression. *Nat Genet* **39**, 1202-7 (2007).
258. Liang, L. *et al.* A cross-platform analysis of 14,177 expression quantitative trait loci derived from lymphoblastoid cell lines. *Genome Res* **23**, 716-26 (2013).
259. Stranger, B.E. *et al.* Population genomics of human gene expression. *Nat Genet* **39**, 1217-24 (2007).
260. Kwan, T. *et al.* Genome-wide analysis of transcript isoform variation in humans. *Nat Genet* **40**, 225-31 (2008).
261. Dimas, A.S. *et al.* Common regulatory variation impacts gene expression in a cell type-dependent manner. *Science* **325**, 1246-50 (2009).
262. Cusanovich, D.A. *et al.* The combination of a genome-wide association study of lymphocyte count and analysis of gene expression data reveals novel asthma candidate genes. *Hum Mol Genet* **21**, 2111-23 (2012).
263. Grundberg, E. *et al.* Mapping cis- and trans-regulatory effects across multiple tissues in twins. *Nat Genet* **44**, 1084-9 (2012).
264. Gutierrez-Arcelus, M. *et al.* Passive and active DNA methylation and the interplay with genetic variation in gene regulation. *Elife* **2**, e00523 (2013).

265. Mangravite, L.M. *et al.* A statin-dependent QTL for GATM expression is associated with statin-induced myopathy. *Nature* **502**, 377-80 (2013).
266. Bryois, J. *et al.* Cis and trans effects of human genomic variants on gene expression. *PLoS Genet* **10**, e1004461 (2014).
267. Huang, J. *et al.* eQTL mapping identifies insertion- and deletion-specific eQTLs in multiple tissues. *Nat Commun* **6**, 6821 (2015).
268. Naranbhai, V. *et al.* Genomic modulators of gene expression in human neutrophils. *Nat Commun* **6**, 7545 (2015).
269. Andiappan, A.K. *et al.* Genome-wide analysis of the genetic regulation of gene expression in human neutrophils. *Nat Commun* **6**, 7971 (2015).
270. Fairfax, B.P. *et al.* Genetics of gene expression in primary immune cells identifies cell type-specific master regulators and roles of HLA alleles. *Nat Genet* **44**, 502-10 (2012).
271. Murphy, A. *et al.* Mapping of numerous disease-associated expression polymorphisms in primary peripheral blood CD4+ lymphocytes. *Hum Mol Genet* **19**, 4745-57 (2010).
272. Heinzen, E.L. *et al.* Tissue-specific genetic control of splicing: implications for the study of complex traits. *PLoS Biol* **6**, e1 (2008).
273. Zeller, T. *et al.* Genetics and beyond--the transcriptome of human monocytes and disease susceptibility. *PLoS One* **5**, e10693 (2010).
274. Almlof, J.C. *et al.* Powerful identification of cis-regulatory SNPs in human primary monocytes using allele-specific gene expression. *PLoS One* **7**, e52260 (2012).
275. Kirsten, H. *et al.* Dissecting the genetics of the human transcriptome identifies novel trait-related trans-eQTLs and corroborates the regulatory relevance of non-protein coding locidagger. *Hum Mol Genet* **24**, 4746-63 (2015).
276. Almlof, J.C. *et al.* Single nucleotide polymorphisms with cis-regulatory effects on long non-coding transcripts in human primary monocytes. *PLoS One* **9**, e102612 (2014).
277. Fairfax, B.P. *et al.* Innate immune activity conditions the effect of regulatory variants upon monocyte gene expression. *Science* **343**, 1246949 (2014).
278. Barreiro, L.B. *et al.* Deciphering the genetic architecture of variation in the immune response to Mycobacterium tuberculosis infection. *Proc Natl Acad Sci U S A* **109**, 1204-9 (2012).
279. Lee, M.N. *et al.* Common genetic variants modulate pathogen-sensing responses in human dendritic cells. *Science* **343**, 1246980 (2014).
280. Huang, R.S. *et al.* Population differences in microRNA expression and biological implications. *RNA Biol* **8**, 692-701 (2011).
281. Fischer, D. *et al.* MiRNA Profiles in Lymphoblastoid Cell Lines of Finnish Prostate Cancer Families. *PLoS One* **10**, e0127427 (2015).
282. Degner, J.F. *et al.* DNase I sensitivity QTLs are a major determinant of human expression variation. *Nature* **482**, 390-4 (2012).
283. del Rosario, R.C. *et al.* Sensitive detection of chromatin-altering polymorphisms reveals autoimmune disease mechanisms. *Nat Methods* **12**, 458-64 (2015).
284. Battle, A. *et al.* Genomic variation. Impact of regulatory variation from RNA to protein. *Science* **347**, 664-7 (2015).
285. Zhang, X. *et al.* Identification of common genetic variants controlling transcript isoform variation in human whole blood. *Nat Genet* **47**, 345-52 (2015).
286. Huan, T. *et al.* Genome-wide identification of microRNA expression quantitative trait loci. *Nat Commun* **6**, 6601 (2015).
287. Greenawalt, D.M. *et al.* A survey of the genetics of stomach, liver, and adipose gene expression from a morbidly obese cohort. *Genome Res* **21**, 1008-16 (2011).

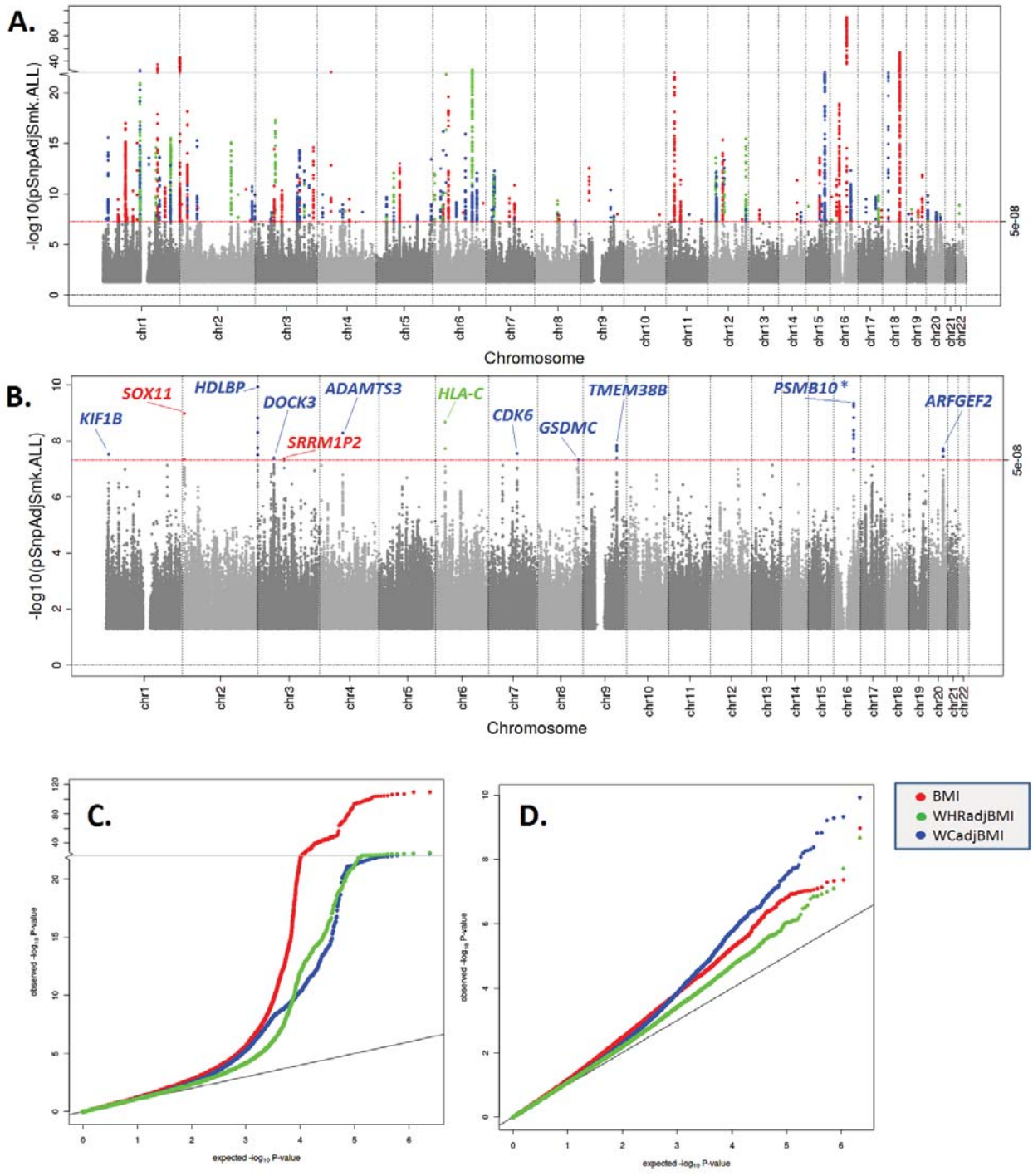
288. Kompass, K.S. & Witte, J.S. Co-regulatory expression quantitative trait loci mapping: method and application to endometrial cancer. *BMC Med Genomics* **4**, 6 (2011).
289. Zhang, B. *et al.* Integrated systems approach identifies genetic nodes and networks in late-onset Alzheimer's disease. *Cell* **153**, 707-20 (2013).
290. Schadt, E.E. *et al.* Mapping the genetic architecture of gene expression in human liver. *PLoS Biol* **6**, e107 (2008).
291. Innocenti, F. *et al.* Identification, replication, and functional fine-mapping of expression quantitative trait loci in primary human liver tissue. *PLoS Genet* **7**, e1002078 (2011).
292. Schroder, A. *et al.* Genomics of ADME gene expression: mapping expression quantitative trait loci relevant for absorption, distribution, metabolism and excretion of drugs in human liver. *Pharmacogenomics J* **13**, 12-20 (2013).
293. Wang, X. *et al.* Mapping of hepatic expression quantitative trait loci (eQTLs) in a Han Chinese population. *J Med Genet* **51**, 319-26 (2014).
294. Grundberg, E. *et al.* Population genomics in a disease targeted primary cell model. *Genome Res* **19**, 1942-52 (2009).
295. Kabakchiev, B. & Silverberg, M.S. Expression quantitative trait loci analysis identifies associations between genotype and gene expression in human intestine. *Gastroenterology* **144**, 1488-96, 1496 e1-3 (2013).
296. Ongen, H. *et al.* Putative cis-regulatory drivers in colorectal cancer. *Nature* **512**, 87-90 (2014).
297. Hulur, I. *et al.* Enrichment of inflammatory bowel disease and colorectal cancer risk variants in colon expression quantitative trait loci. *BMC Genomics* **16**, 138 (2015).
298. Keildson, S. *et al.* Expression of phosphofructokinase in skeletal muscle is influenced by genetic variation and associated with insulin sensitivity. *Diabetes* **63**, 1154-65 (2014).
299. Quigley, D.A. *et al.* The 5p12 breast cancer susceptibility locus affects MRPS30 expression in estrogen-receptor positive tumors. *Mol Oncol* **8**, 273-84 (2014).
300. Curtis, C. *et al.* The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature* **486**, 346-52 (2012).
301. Hao, K. *et al.* Lung eQTLs to help reveal the molecular underpinnings of asthma. *PLoS Genet* **8**, e1003029 (2012).
302. Gao, C. *et al.* HEFT: eQTL analysis of many thousands of expressed genes while simultaneously controlling for hidden factors. *Bioinformatics* **30**, 369-76 (2014).
303. Lamontagne, M. *et al.* Refining susceptibility loci of chronic obstructive pulmonary disease with lung eqtls. *PLoS One* **8**, e70220 (2013).
304. Luo, W. *et al.* Airway Epithelial Expression Quantitative Trait Loci Reveal Genes Underlying Asthma and Other Airway Diseases. *Am J Respir Cell Mol Biol* (2015).
305. Ding, J. *et al.* Gene expression in skin and lymphoblastoid cells: Refined statistical method reveals extensive overlap in cis-eQTL signals. *Am J Hum Genet* **87**, 779-89 (2010).
306. Wagner, J.R. *et al.* The relationship between DNA methylation, genetic and expression inter-individual variation in untransformed human fibroblasts. *Genome Biol* **15**, R37 (2014).
307. Qiu, W. *et al.* Genetics of sputum gene expression in chronic obstructive pulmonary disease. *PLoS One* **6**, e24395 (2011).
308. Fadista, J. *et al.* Global genomic and transcriptomic analysis of human pancreatic islets reveals novel genes influencing glucose metabolism. *Proc Natl Acad Sci U S A* **111**, 13924-9 (2014).
309. Larson, N.B. *et al.* Comprehensively evaluating cis-regulatory variation in the human prostate transcriptome by using gene-level allele-specific expression. *Am J Hum Genet* **96**, 869-82 (2015).
310. Singh, T. *et al.* Characterization of expression quantitative trait loci in the human colon. *Inflamm Bowel Dis* **21**, 251-6 (2015).

311. Koopmann, T.T. *et al.* Genome-wide identification of expression quantitative trait loci (eQTLs) in human heart. *PLoS One* **9**, e97380 (2014).
312. Lin, H. *et al.* Gene expression and genetic variation in human atria. *Heart Rhythm* **11**, 266-71 (2014).
313. Rantalainen, M. *et al.* MicroRNA expression in abdominal and gluteal adipose tissue is associated with mRNA expression levels and partly genetically driven. *PLoS One* **6**, e27338 (2011).
314. Gamazon, E.R. *et al.* A genome-wide integrative study of microRNAs in human liver. *BMC Genomics* **14**, 395 (2013).
315. Olsson, A.H. *et al.* Genome-wide associations between genetic and epigenetic variation influence mRNA expression and insulin secretion in human pancreatic islets. *PLoS Genet* **10**, e1004735 (2014).
316. Li, Q. *et al.* Expression QTL-based analyses reveal candidate causal genes and loci across five tumor types. *Hum Mol Genet* **23**, 5294-302 (2014).
317. Webster, J.A. *et al.* Genetic control of human brain transcript expression in Alzheimer disease. *Am J Hum Genet* **84**, 445-58 (2009).
318. Zou, F. *et al.* Brain expression genome-wide association study (eGWAS) identifies human disease-associated variants. *PLoS Genet* **8**, e1002707 (2012).
319. Kim, Y. *et al.* A meta-analysis of gene expression quantitative trait loci in brain. *Transl Psychiatry* **4**, e459 (2014).
320. Ramasamy, A. *et al.* Genetic variability in the regulation of gene expression in ten regions of the human brain. *Nat Neurosci* **17**, 1418-28 (2014).
321. Gibbs, J.R. *et al.* Abundant quantitative trait loci exist for DNA methylation and gene expression in human brain. *PLoS Genet* **6**, e1000952 (2010).
322. Gamazon, E.R. *et al.* Enrichment of cis-regulatory gene expression SNPs and methylation quantitative trait loci among bipolar disorder susceptibility variants. *Mol Psychiatry* **18**, 340-6 (2013).
323. Kim, S., Cho, H., Lee, D. & Webster, M.J. Association between SNPs and gene expression in multiple regions of the human brain. *Transl Psychiatry* **2**, e113 (2012).
324. Shpak, M. *et al.* An eQTL analysis of the human glioblastoma multiforme genome. *Genomics* **103**, 252-63 (2014).
325. Colantuoni, C. *et al.* Temporal dynamics and genetic control of transcription in the human prefrontal cortex. *Nature* **478**, 519-23 (2011).
326. Liu, C. *et al.* Whole-genome association mapping of gene expression in the human prefrontal cortex. *Mol Psychiatry* **15**, 779-84 (2010).
327. Irizarry, R.A. *et al.* Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* **4**, 249-64 (2003).

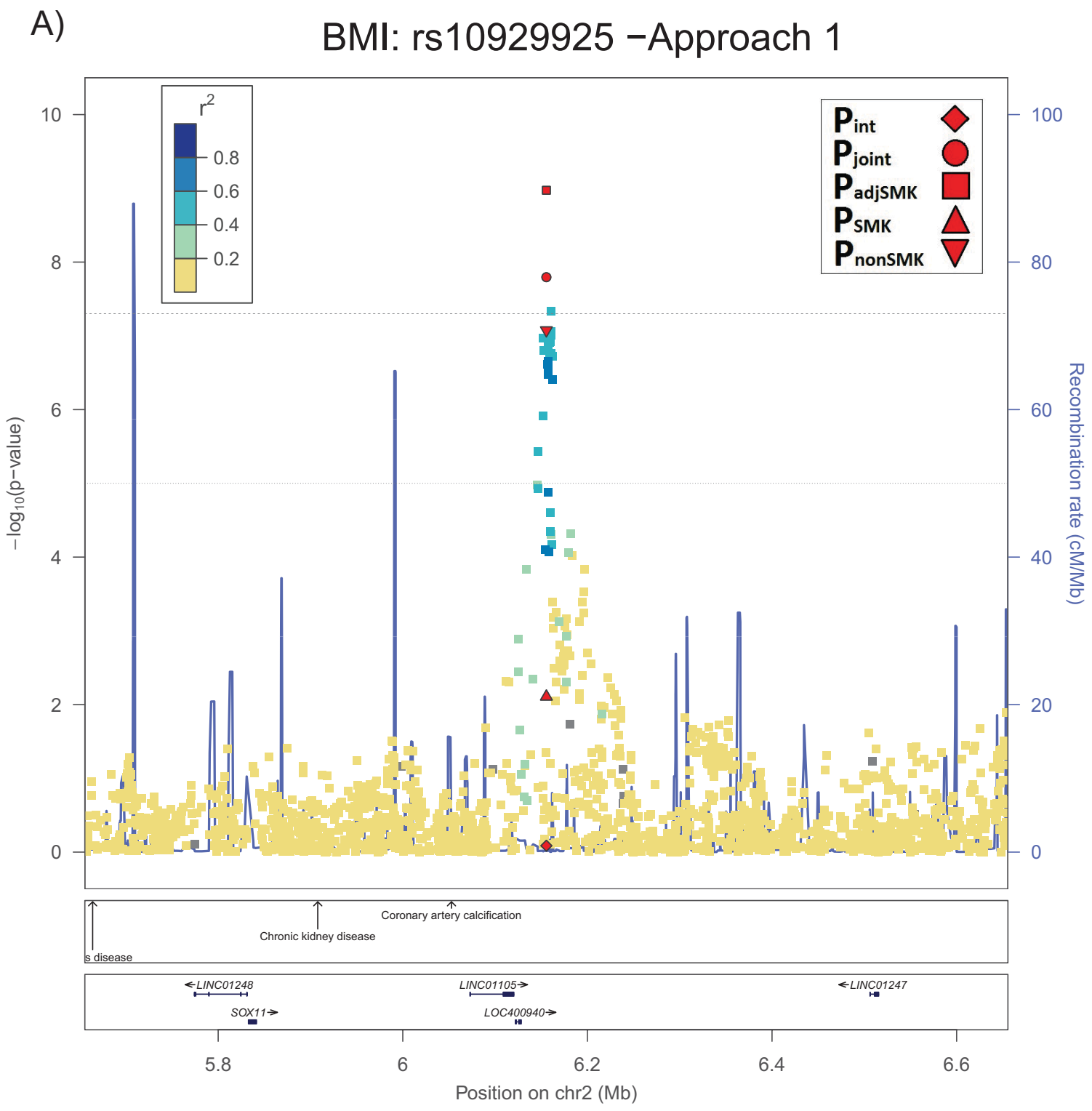
Supplementary Figure 1. Summary of overall study design and workflow for meta-analyses. All numbers provided represent the maximum number specific for that trait (BMI-red, WCadjBMI-blue, and WHRadjBMI-green) and strata (EUR-European descent participants, nonEUR-excluding European descent participants). Three studies provided GWAS data for EUR and nonEUR participants.

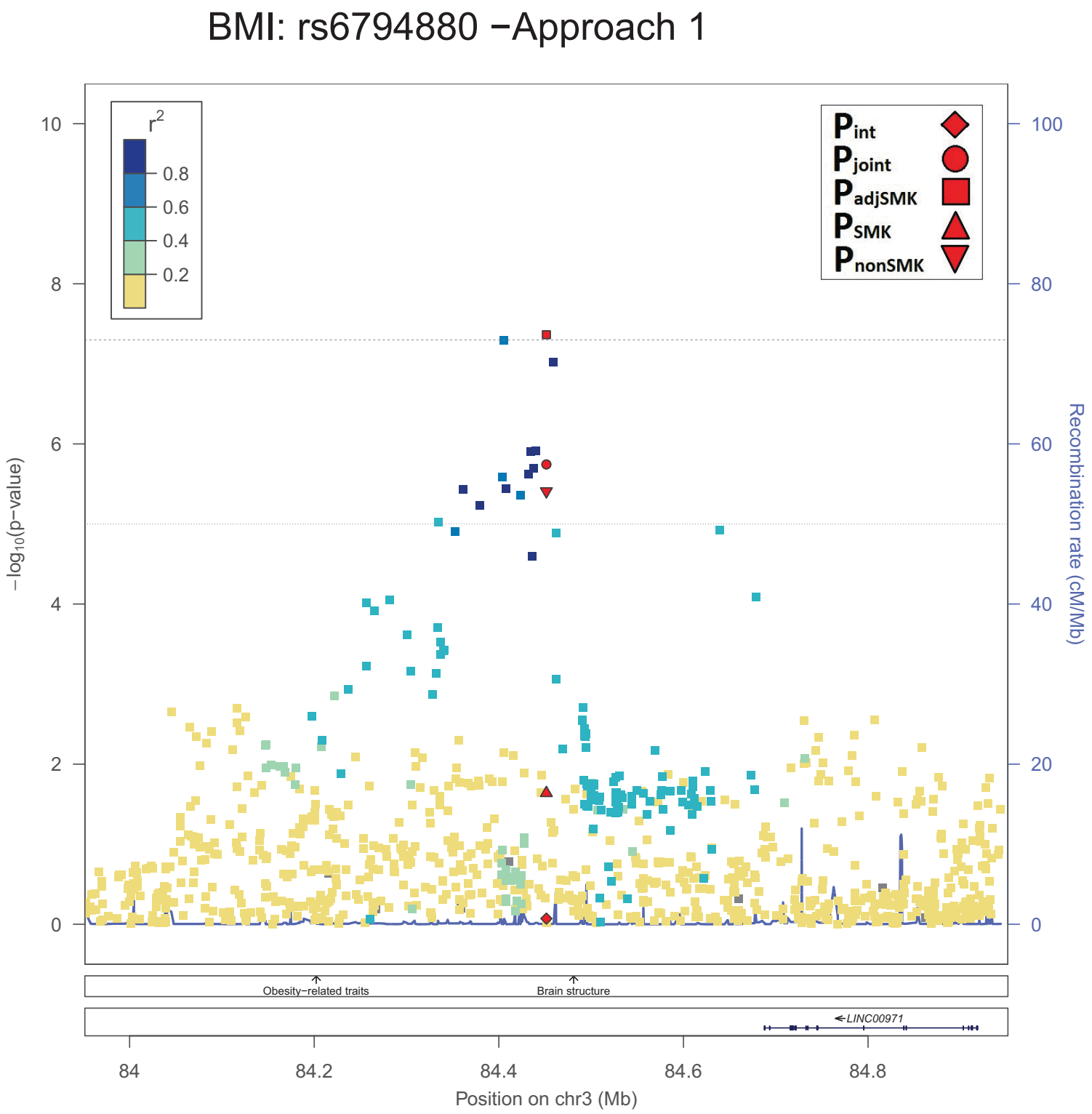


Supplementary Figure 2. Summary plots of discovery meta-analysis for Approach 1 primary meta-analyses. (A) Manhattan plot showing the loci identified in Approach 1 in primary meta-analyses, used to identify significant main effects loci (SNPadjSMK), in the primary meta-analyses association –log10P-values for BMI-red, WCadjBMI-blue, and WHRadjBMI-green; (B) Manhattan plot showing the loci identified in Approach 1 excluding known regions +/- 500 kb and labeled with the nearest gene to the index SNP; (C) QQ-plot showing the Approach 1 P-values as observed against those expected under the null for each phenotypes separately (colored); (D) QQ-plot for Approach 1 after excluding known association regions. **PSMB10* locus is >500 +/- kb from previously identified index SNPs, but is not independent of known GWAS signals.

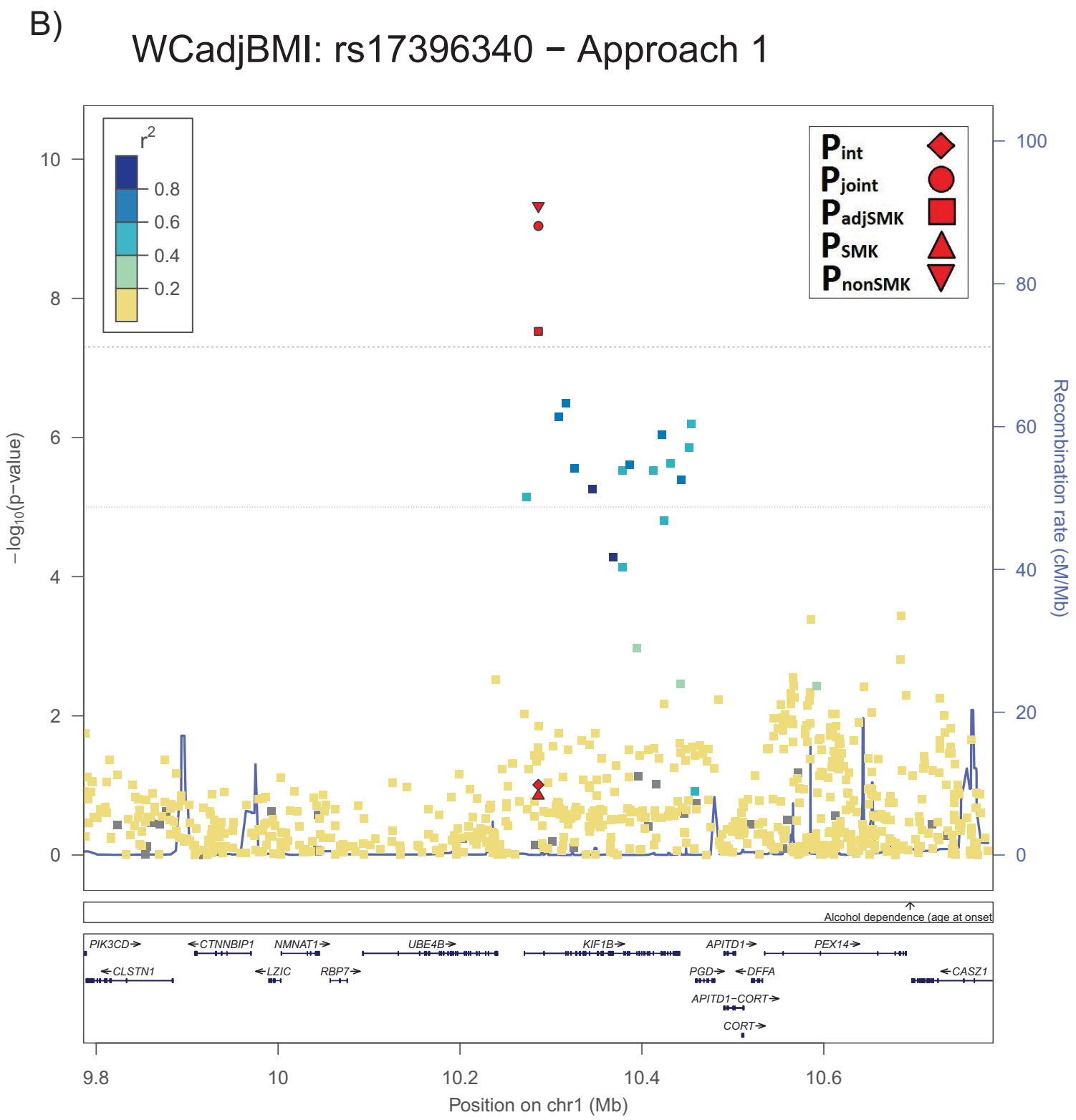


Supplementary Figure 3. Regional association plot for all loci identified in Approach 1 in primary meta-analyses, used to identify significant interaction (SNPadjSMK), in the primary meta-analyses for A) BMI, B) WCadjBMI, and C) WHRadjBMI, and ordered as they appear in Table 1. LD has been calculated using the combined ancestries from the 1000 Genomes Phase 1 reference panel. For comparison, each plot highlights the p-value for the tag SNP in Approach 1 (P_{adjSMK}), Approach 2 (P_{joint}), Approach 3 (P_{int}), current smokers (P_{SMK}), and in nonsmokers (P_{nonSMK}). EUR-European-only meta-analysis.

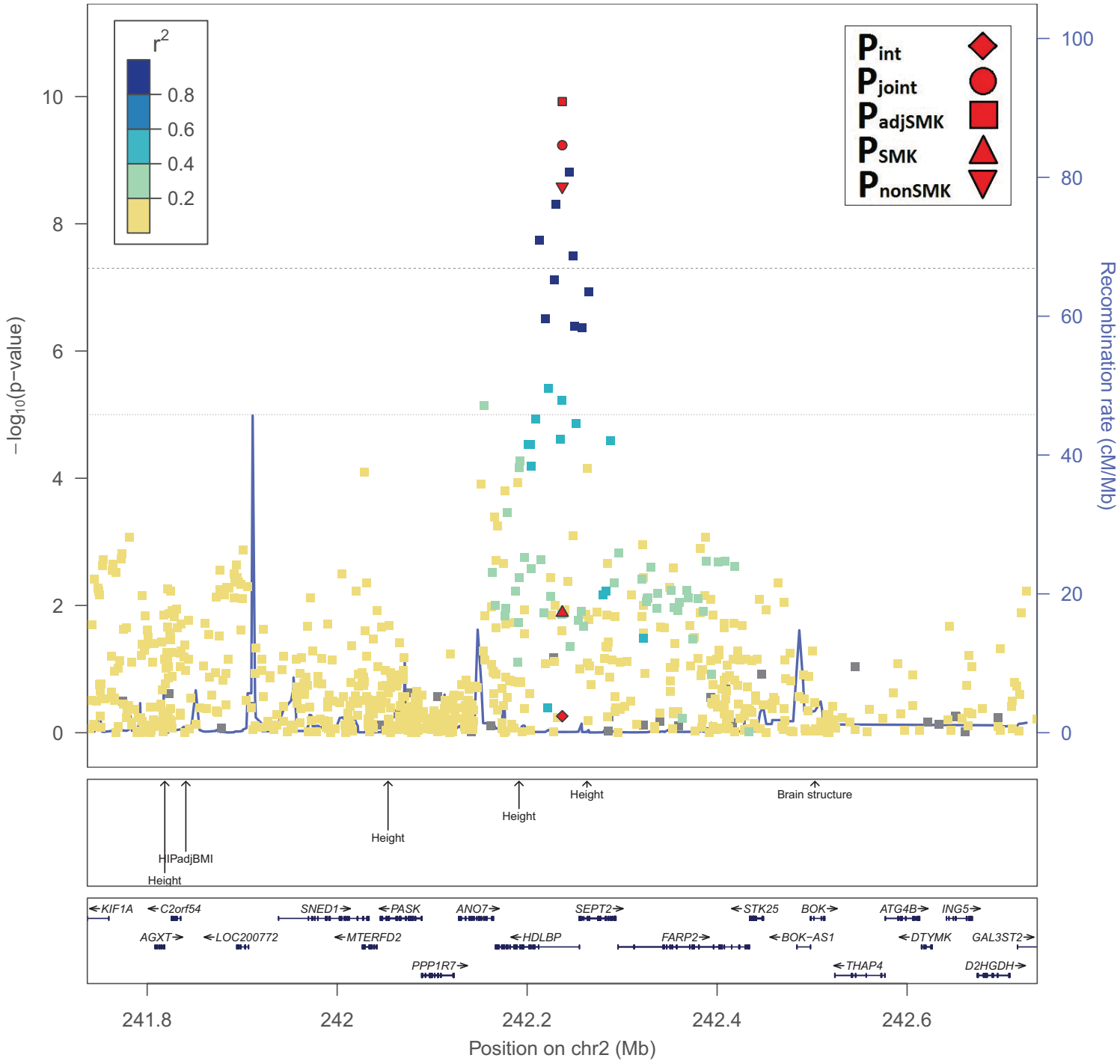

A)
BMI: rs10929925 −Approach 1
r2
0.8
0.6
0.4
0.2
Pint
Pjoint
PadjSMK
PSMK
PnonSMK
−log10(p−value)
Recombination rate (cM/Mb)
Coronary artery calcification
Chronic kidney disease
s disease
LINC01248
LINC01105
LINC01247
SOX11
LOC400940
Position on chr2 (Mb)

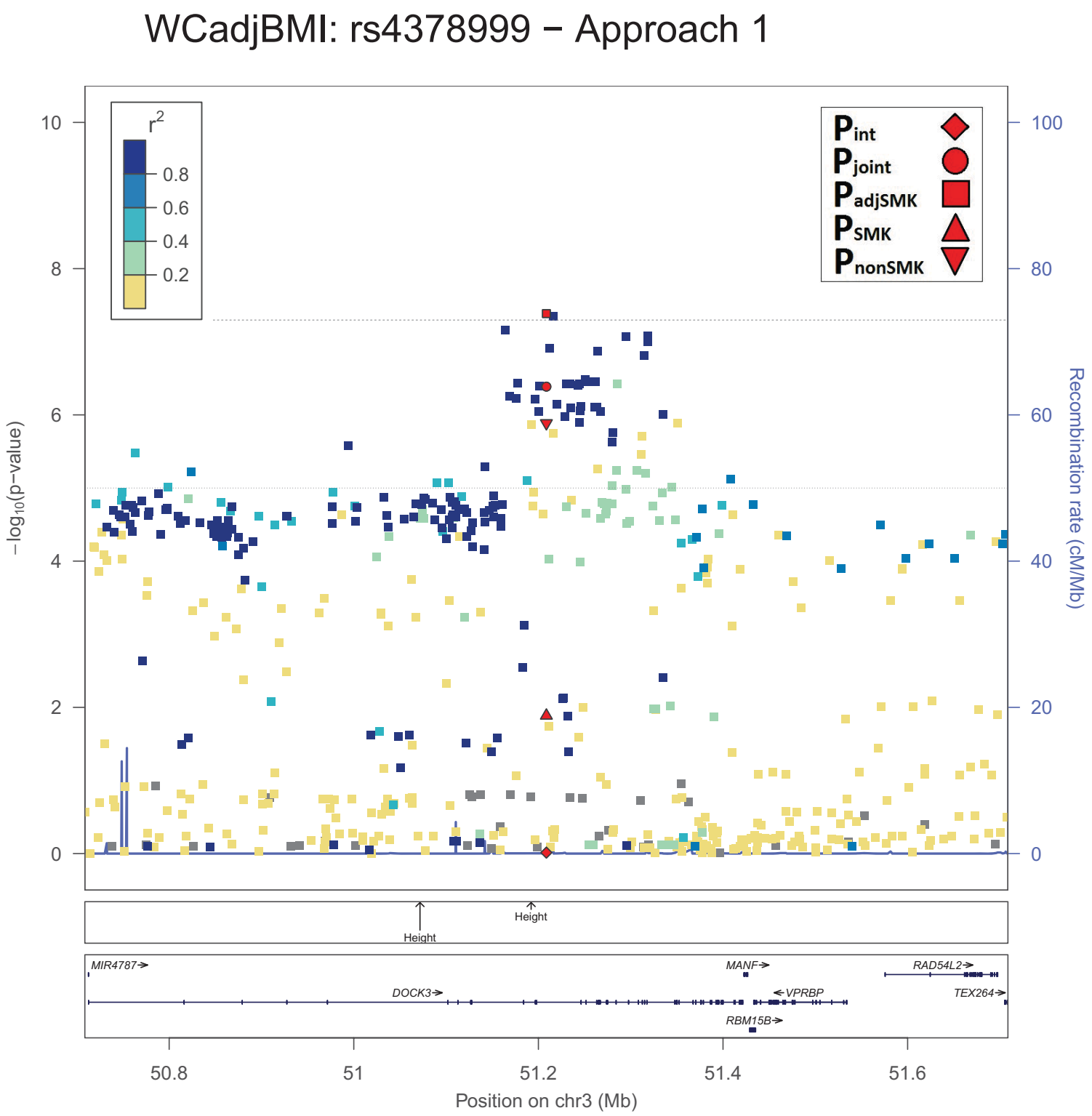

BMI: rs6794880 −Approach 1
r2
0.8
0.6
0.4
0.2
Pint
Pjoint
PadjSMK
PSMK
PnonSMK
10
8
6
4
2
0
100
80
60
40
20
0
−log10(p−value)
Recombination rate (cM/Mb)
Obesity−related traits
Brain structure
LINC00971
84
84.2
84.4
84.6
84.8
Position on chr3 (Mb)

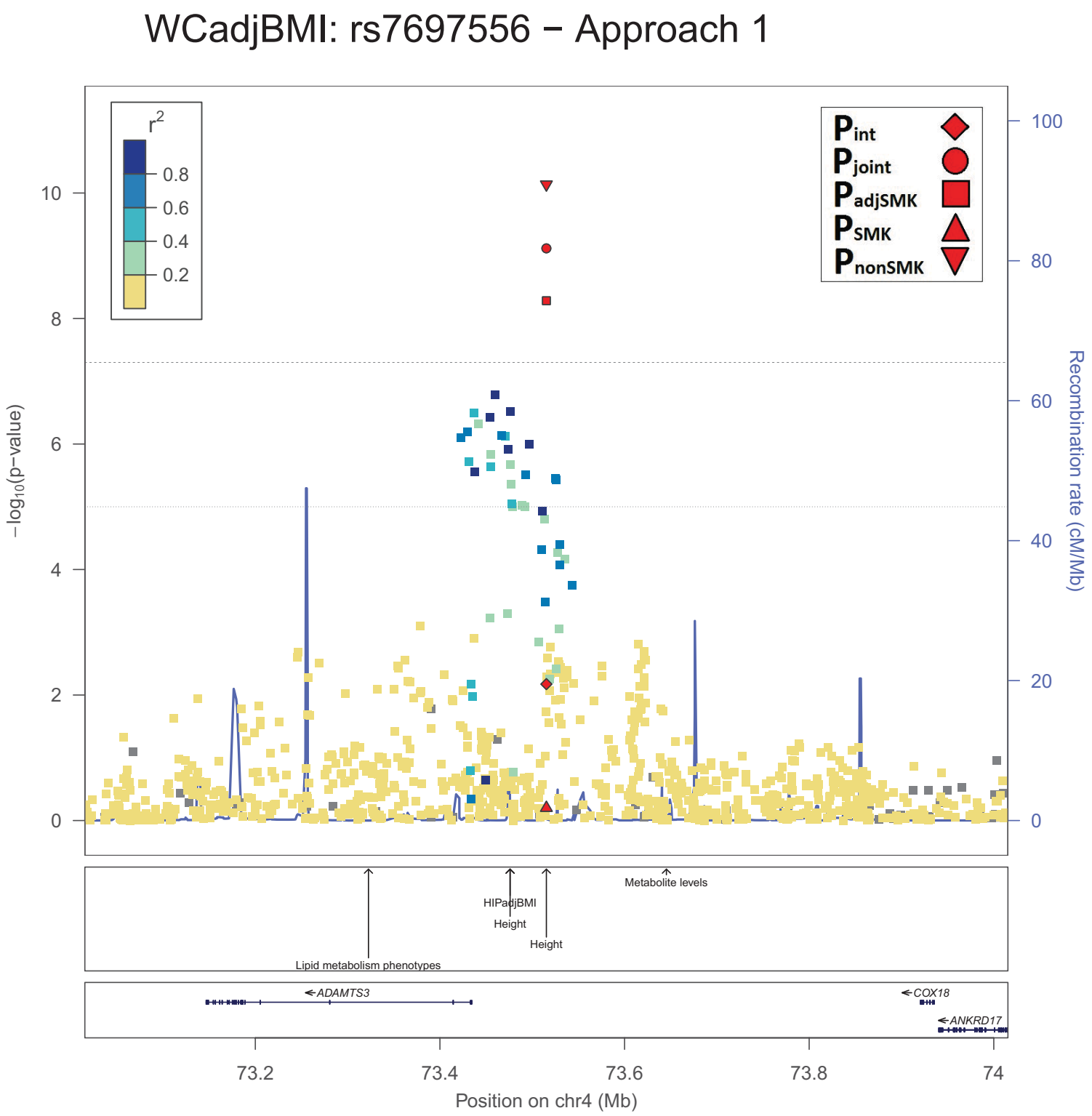
B)


WCadjBMI: rs17396340 − Approach 1
r2
0.8
0.6
0.4
0.2
Pint
Pjoint
PadjSMK
PSMK
PnonSMK
−log10(p−value)
Recombination rate (cM/Mb)
Alcohol dependence (age at onset)
PIK3CD
CTNNBIP1
NMNAT1
UBE4B
KIF1B
APITD1
PEX14
CLSTN1
LZIC
RBP7
PGD
DFFA
CASZ1
APITD1−CORT
CORT
Position on chr1 (Mb)

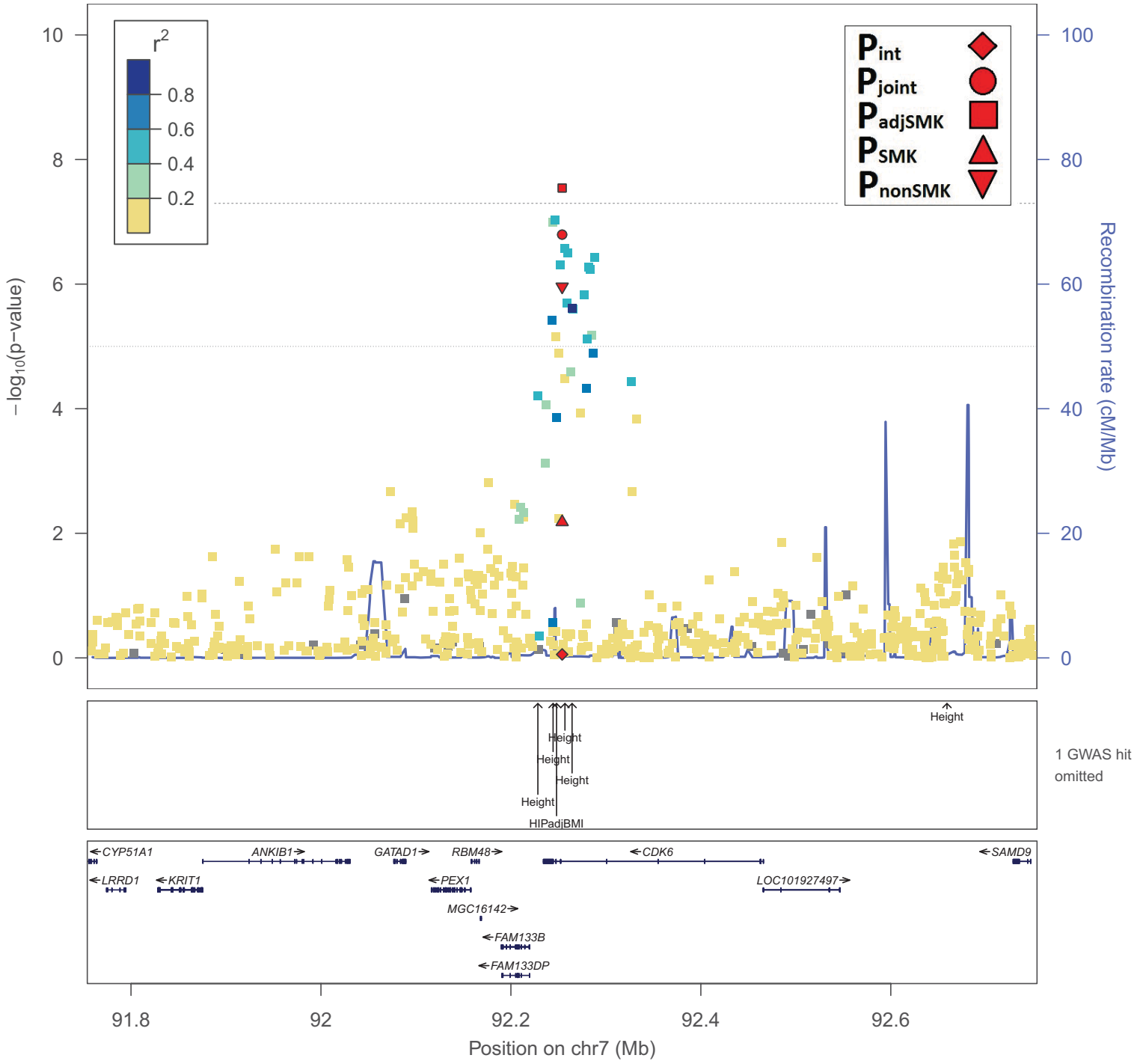
WCadjBMI: rs6743226 − Approach 1

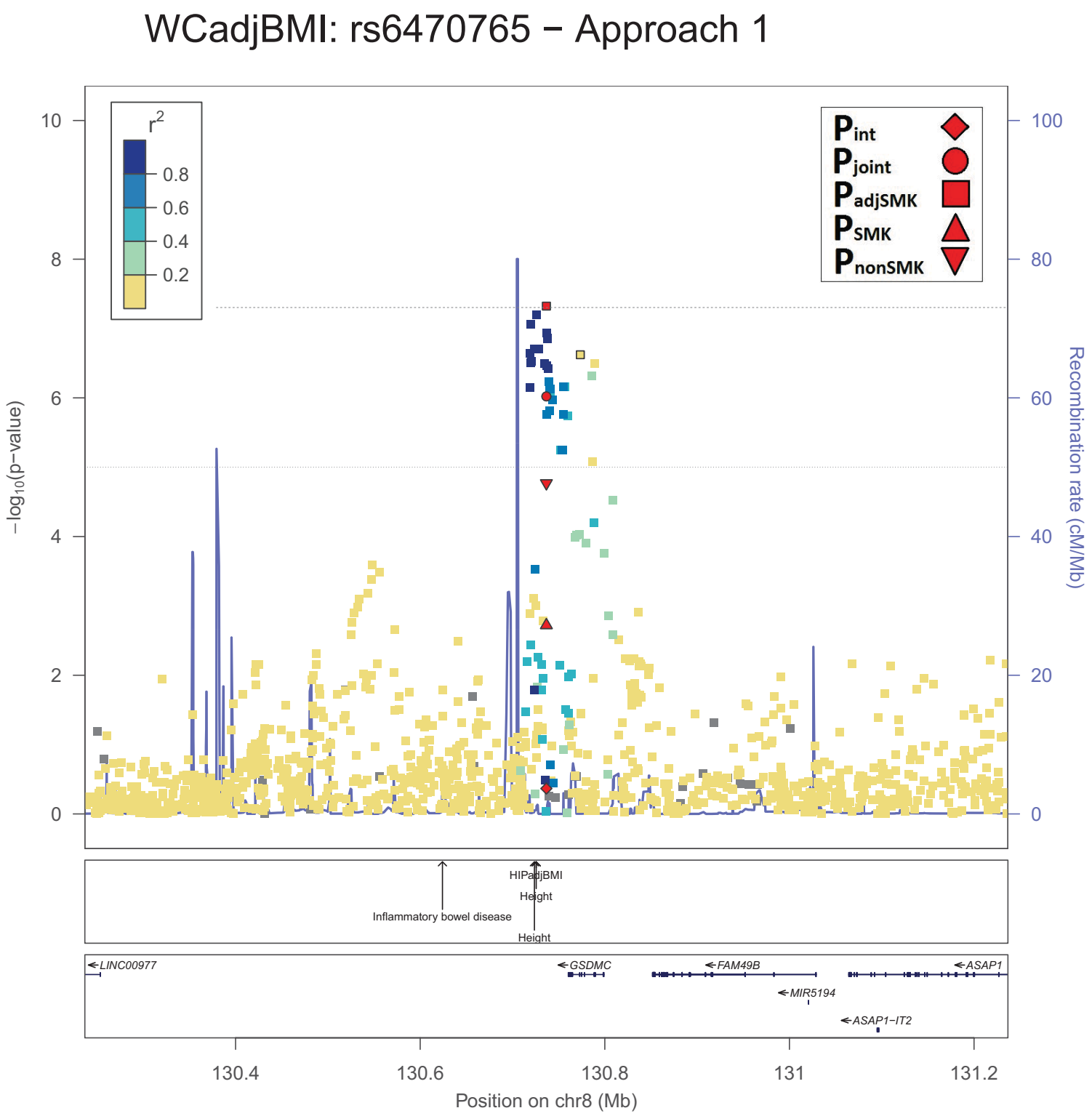



WCadjBMI: rs4378999 − Approach 1
r2
0.8
0.6
0.4
0.2
Pint
Pjoint
PadjSMK
PSMK
PnonSMK
−log10(p−value)
Recombination rate (cM/Mb)
Height
Height
MIR4787
DOCK3
MANF
VPRBP
RBM15B
RAD54L2
TEX264
Position on chr3 (Mb)

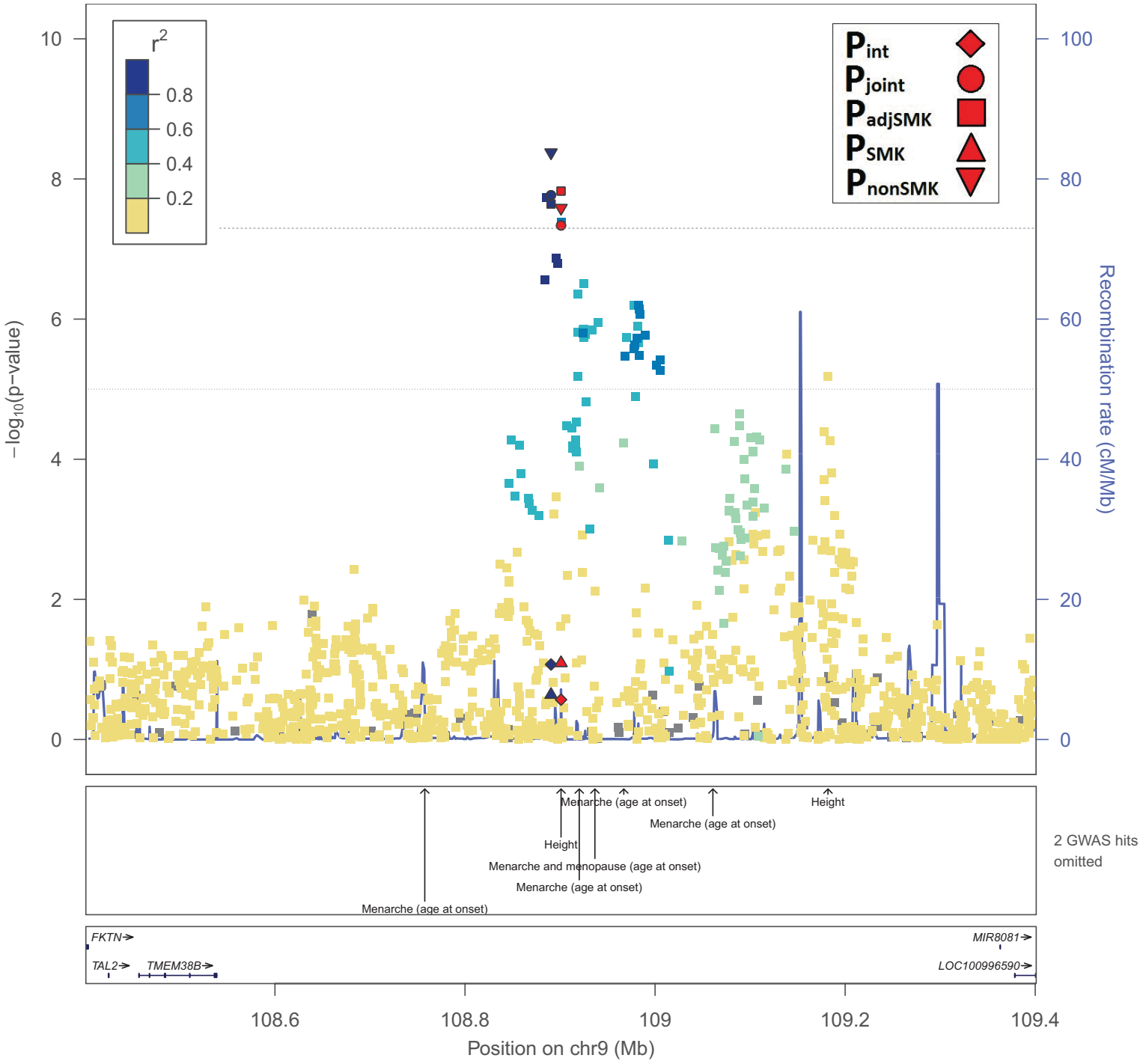

WCadjBMI: rs7697556 − Approach 1
r2
0.8
0.6
0.4
0.2
Pint
Pjoint
PadjSMK
PSMK
PnonSMK
−log10(p−value)
Recombination rate (cM/Mb)
Metabolite levels
HIPadjBMI
Height
Height
Lipid metabolism phenotypes
ADAMTS3
COX18
ANKRD17
Position on chr4 (Mb)

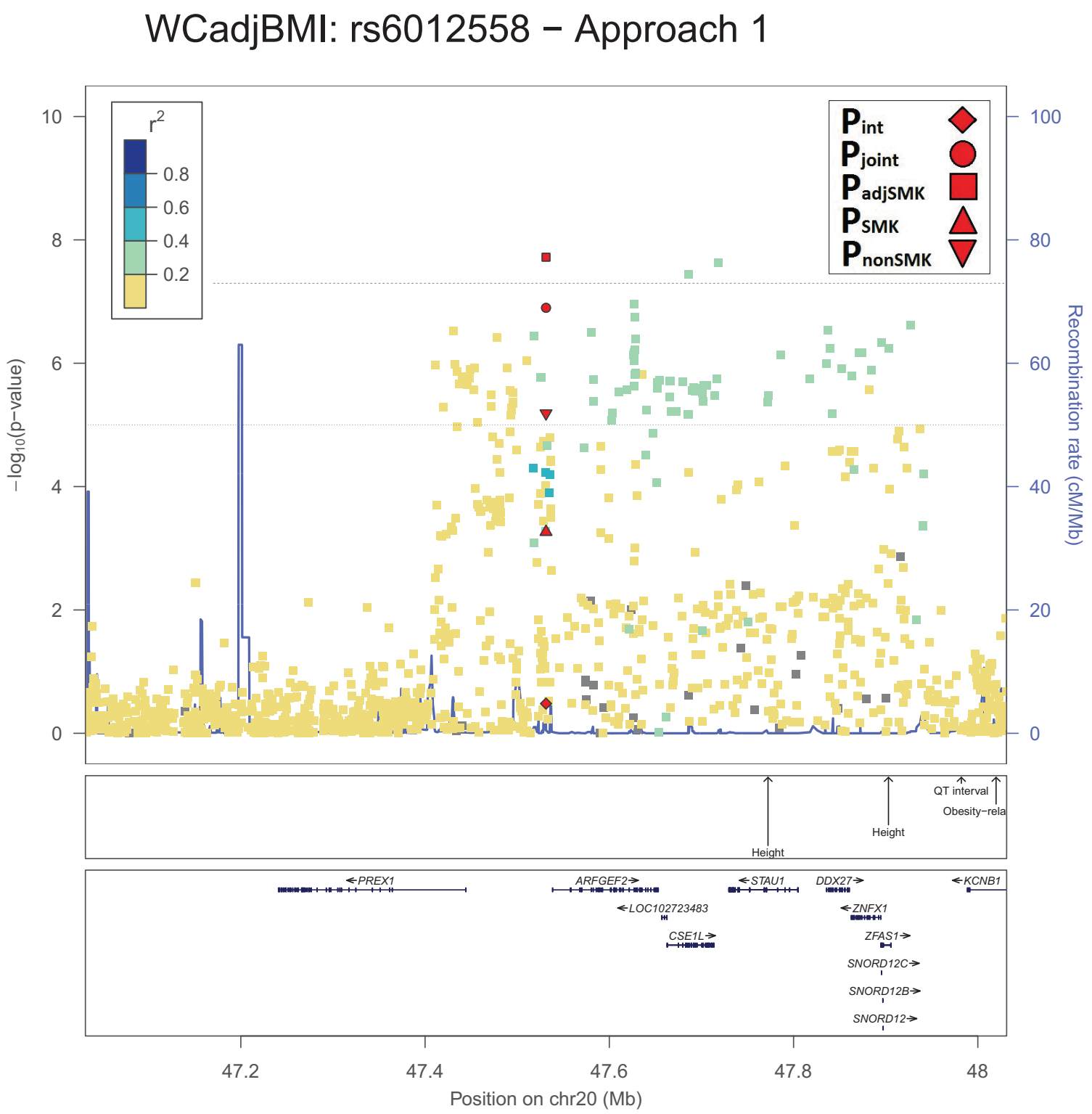
WCadjBMI: rs10269774 − Approach 1



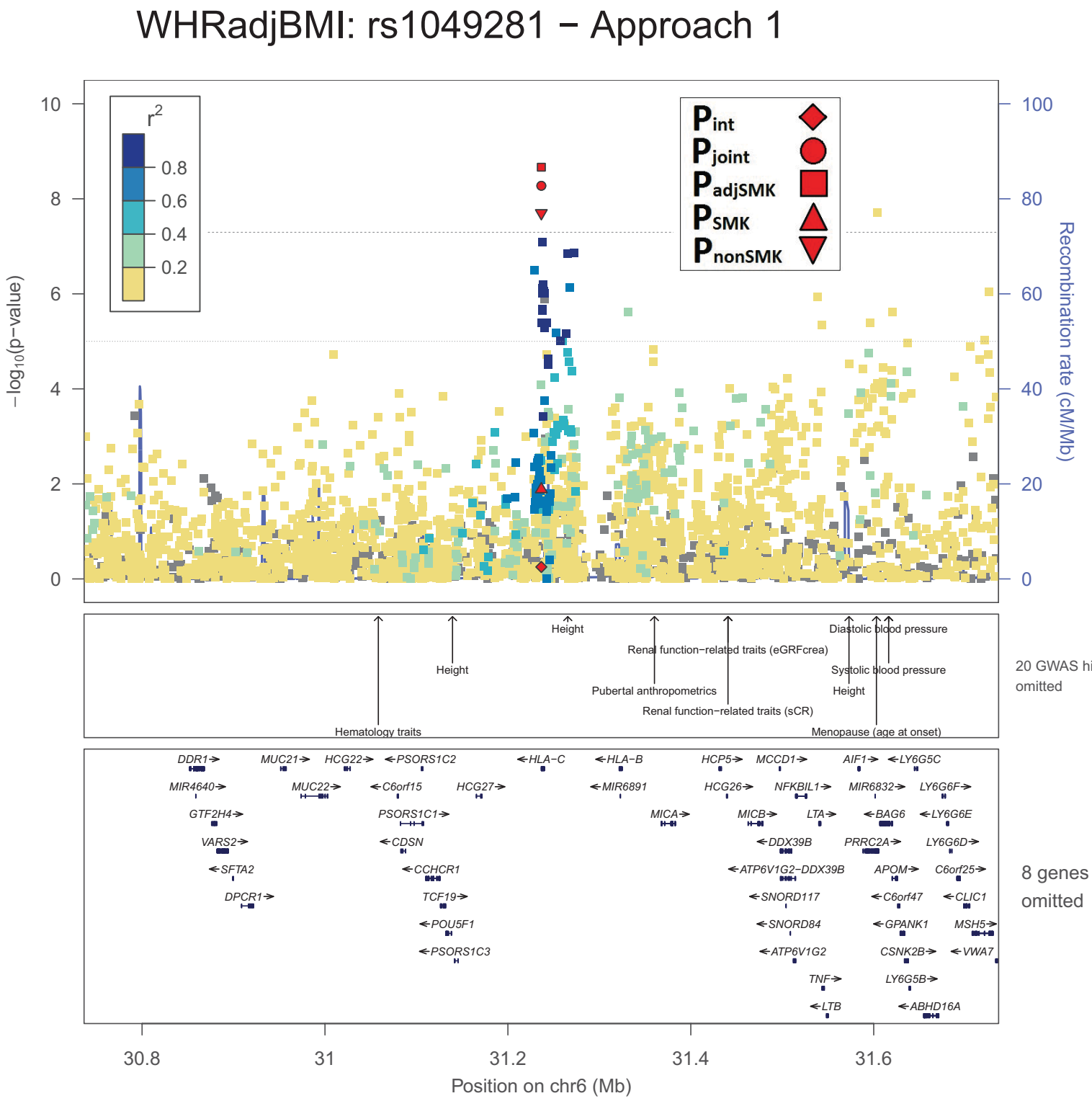

WCadjBMI: rs6470765 − Approach 1
r2
0.8
0.6
0.4
0.2
Pint
Pjoint
PadjSMK
PSMK
PnonSMK
−log10(p−value)
Recombination rate (cM/Mb)
HIPadjBMI
Height
Inflammatory bowel disease
Height
LINC00977
GSDMC
FAM49B
ASAP1
MIR5194
ASAP1−IT2
Position on chr8 (Mb)

WCadjBMI: rs9409082 − Approach 1

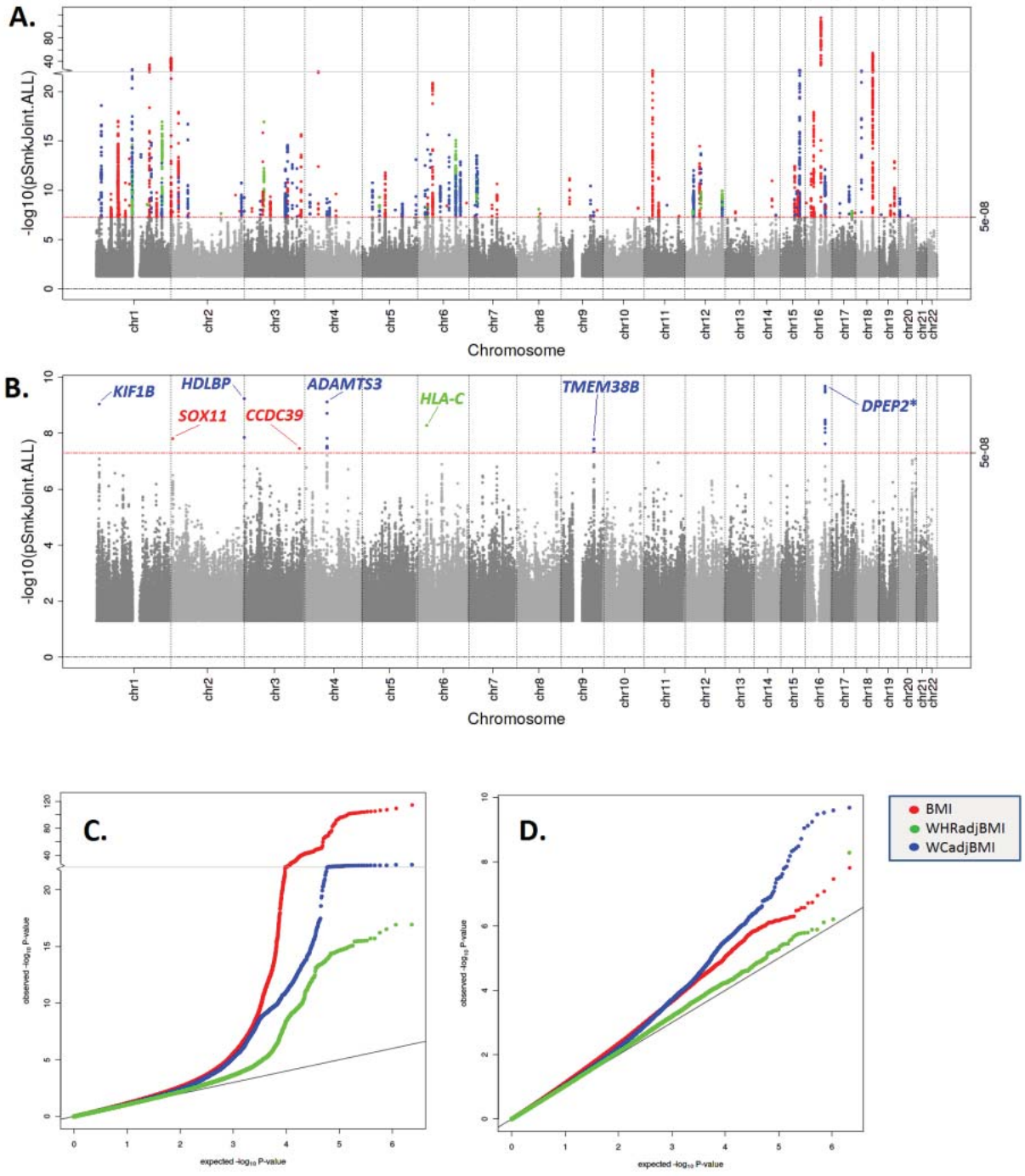



WCadjBMI: rs6012558 − Approach 1
r2
0.8
0.6
0.4
0.2
Pint
Pjoint
PadjSMK
PSMK
PnonSMK
−log10(p−value)
Recombination rate (cM/Mb)
10
8
6
4
2
0
100
80
60
40
20
0
Height
Height
QT interval
Obesity−rela
PREX1
ARFGEF2
STAU1
DDX27
KCNB1
LOC102723483
ZNFX1
CSE1L
ZFAS1
SNORD12C
SNORD12B
SNORD12
47.2
47.4
47.6
47.8
48
Position on chr20 (Mb)

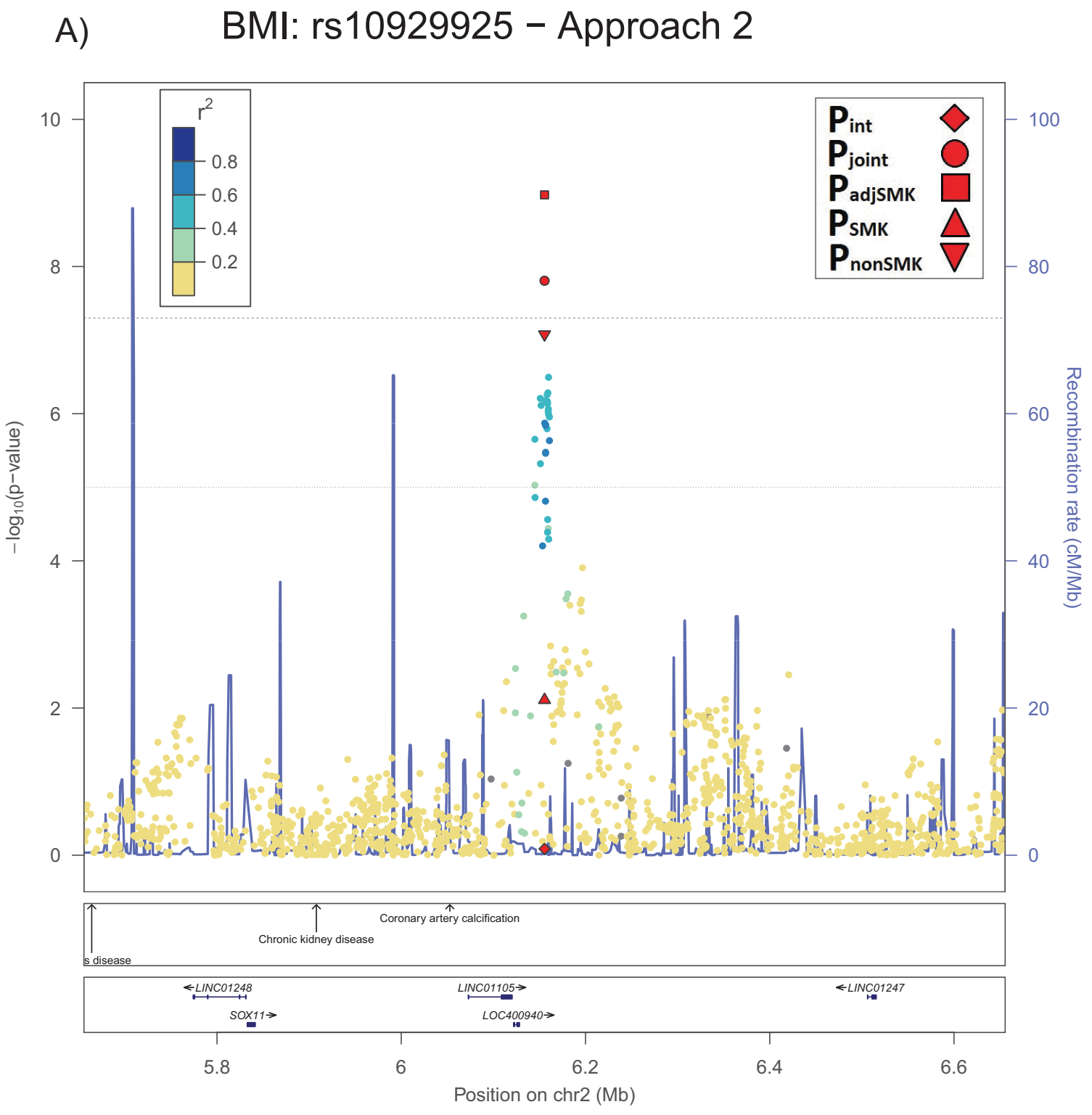
C)


WHRadjBMI: rs1049281 − Approach 1
r2
0.8
0.6
0.4
0.2
Pint
Pjoint
PadjSMK
PSMK
PnonSMK
−log10(p−value)
Recombination rate (cM/Mb)
Height
Height
Height
Renal function−related traits (eGRFcrea)
Diastolic blood pressure
Systolic blood pressure
Pubertal anthropometrics
Height
Renal function−related traits (sCR)
Hematology traits
Menopause (age at onset)
20 GWAS hits omitted
8 genes omitted
Position on chr6 (Mb)

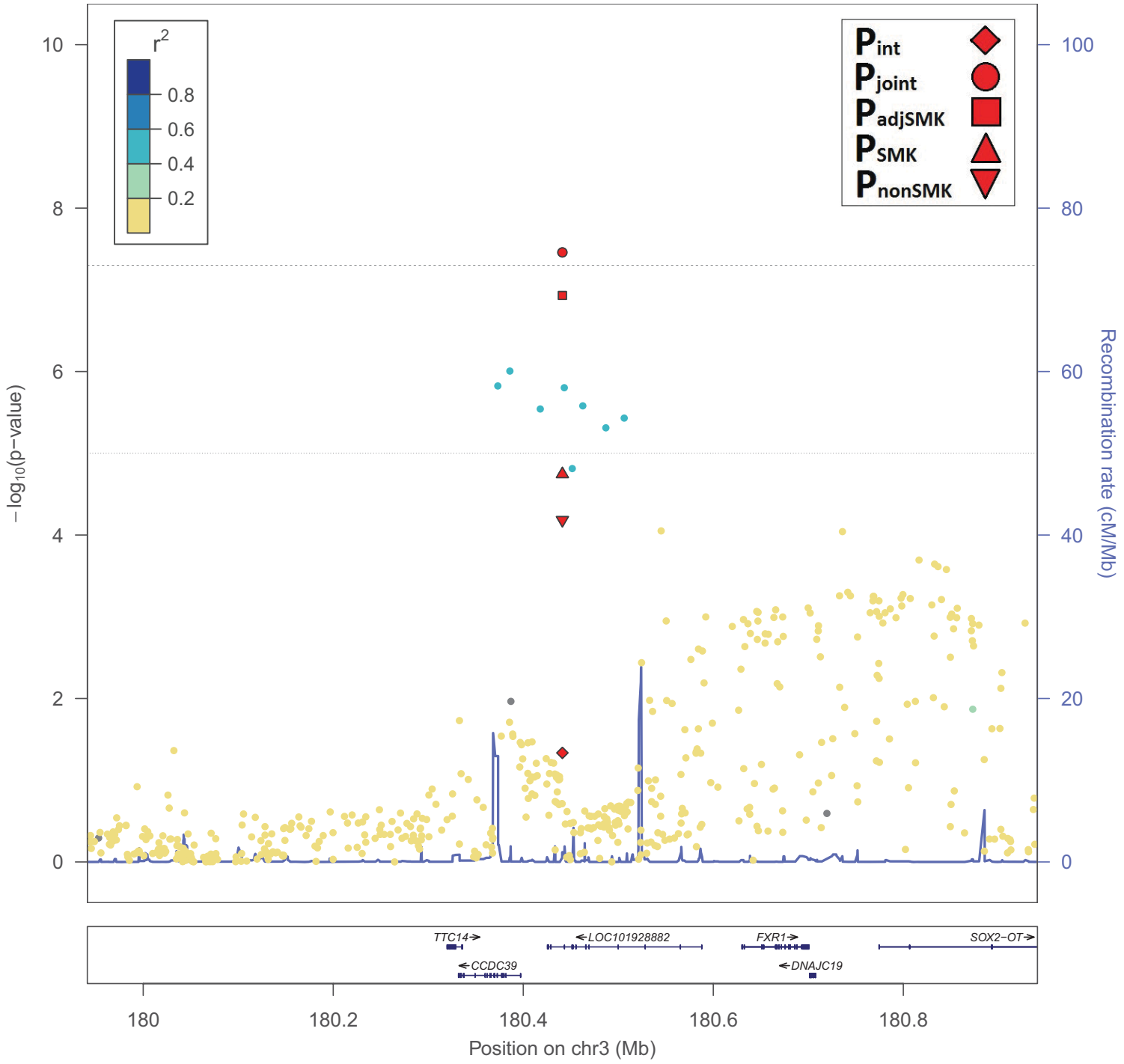
Supplementary Figure 4. Summary plots of discovery meta-analysis for Approach 2 primary meta-analyses. (A) Manhattan plot showing the loci identified in Approach 2 in primary meta-analyses, used to identify significant joint main+interaction effects loci (SNPjoint), in the primary meta-analyses association –log10P-values for BMI-red, WCadjBMI-blue, and WHRadjBMI-green; (B) Manhattan plot showing the loci identified in Approach 2 excluding known regions +/- 500 kb and labeled with the nearest gene to the index SNP; (C) QQ-plot showing the Approach 2 P-values as observed against those expected under the null for each phenotypes separately (colored); (D) QQ-plot for Approach 2 after excluding known association regions.



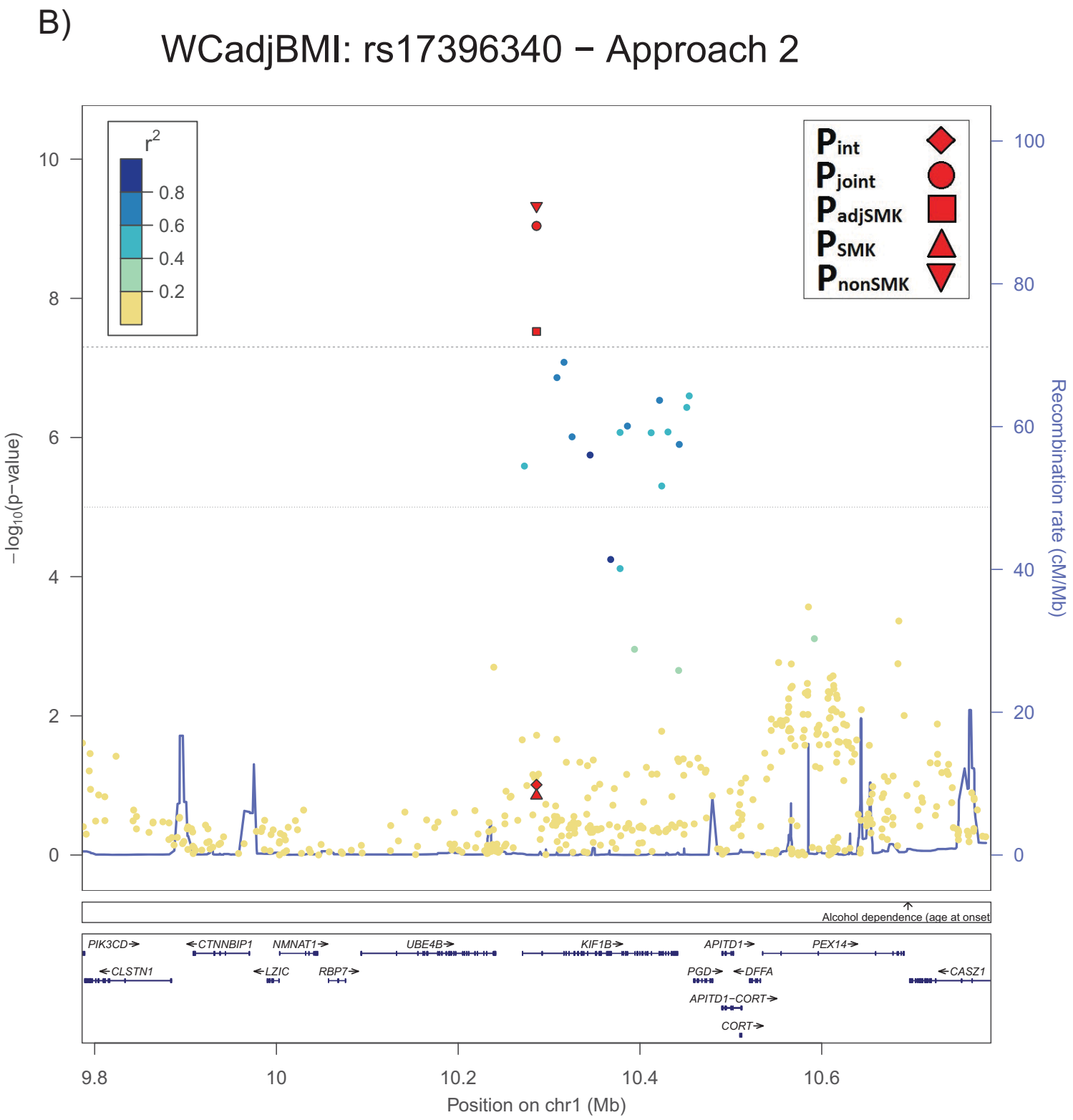
Supplementary Figure 5. Regional association plot for all loci identified in Approach 2 in primary meta-analyses, used to identify significant interaction (SNPint), in the primary meta-analyses for A) BMI and B) WCadjBMI, and ordered as they appear in Table 1. LD has been calculated using the combined ancestries from the 1000 Genomes Phase 1 reference panel. For comparison, each plot highlights the p-value for the tag SNP in Approach 1 (P_{adjSMK}), Approach 2 (P_{joint}), Approach 3 (P_{int}), current smokers (P_{SMK}), and in nonsmokers (P_{nonSMK}). EUR-European-only meta-analysis.


A)
BMI: rs10929925 − Approach 2
r2
0.8
0.6
0.4
0.2
Pint
Pjoint
PadjSMK
PSMK
PnonSMK
−log10(p−value)
Recombination rate (cM/Mb)
Coronary artery calcification
Chronic kidney disease
s disease
LINC01248
SOX11
LINC01105
LOC400940
LINC01247
Position on chr2 (Mb)

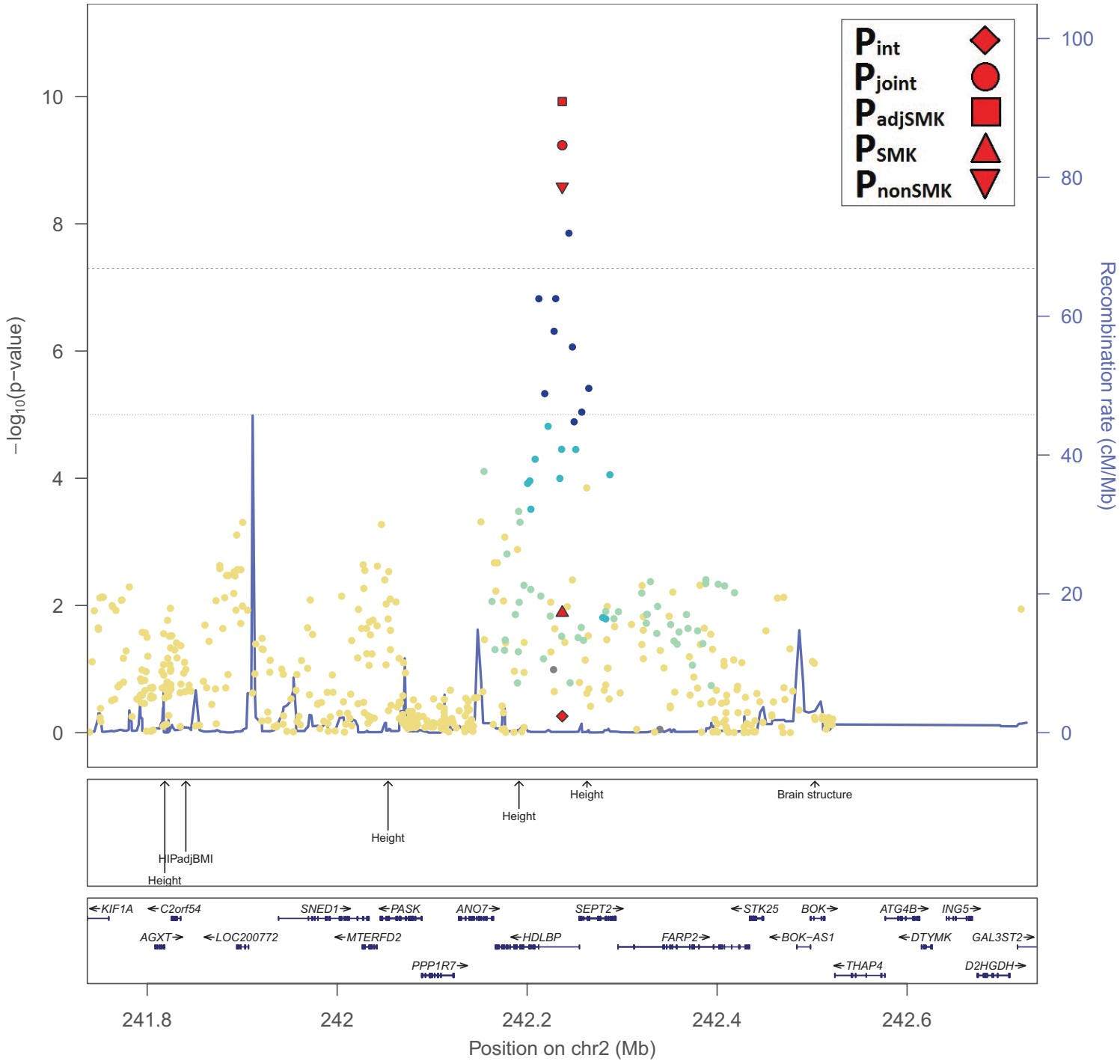
BMI: rs13069244 − Approach 2



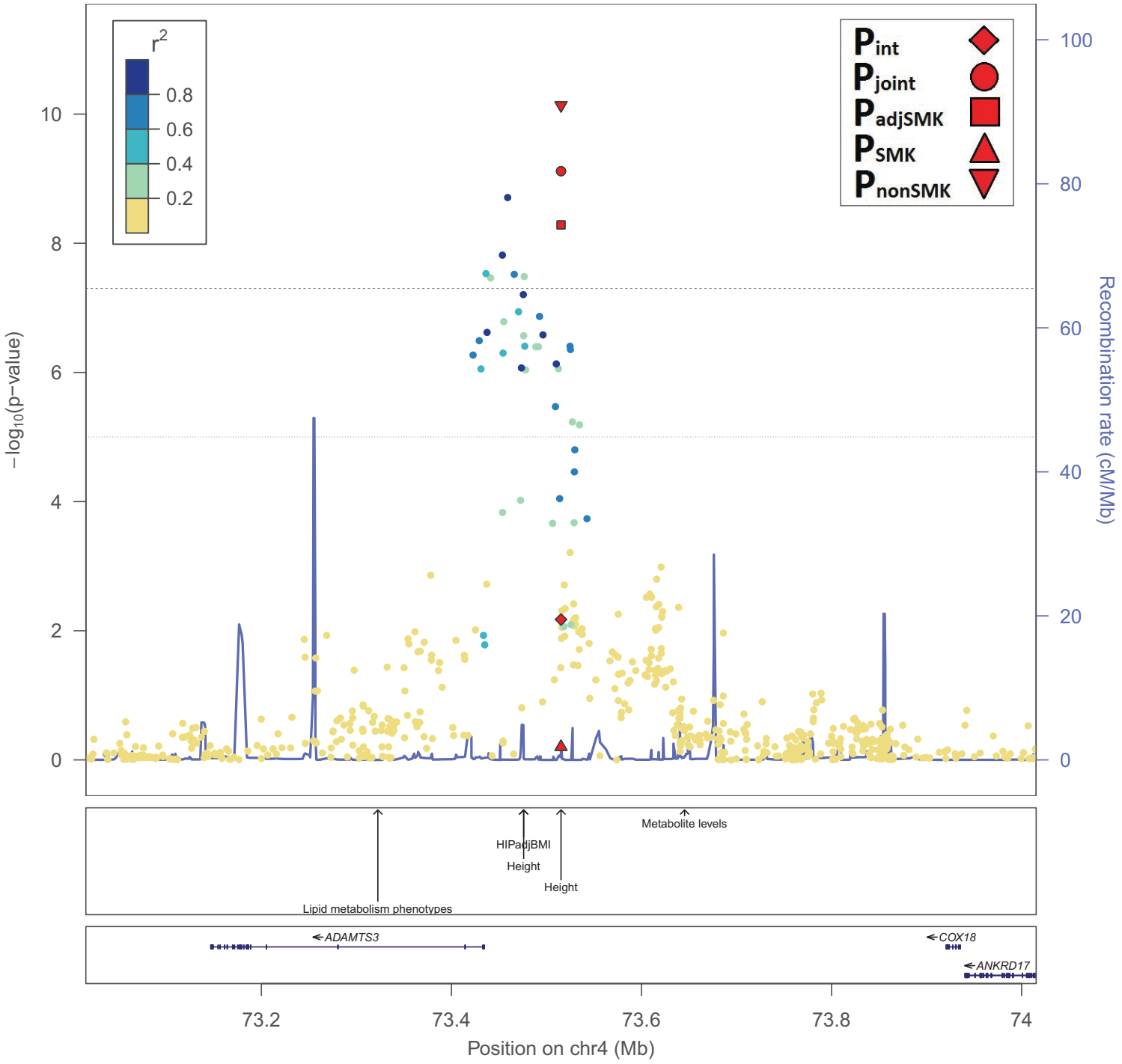
B)


WCadjBMI: rs17396340 − Approach 2
r2
0.8
0.6
0.4
0.2
Pint
Pjoint
PadjSMK
PSMK
PnonSMK
−log10(p−value)
Recombination rate (cM/Mb)
Alcohol dependence (age at onset)
PIK3CD
CTNNBIP1
NMNAT1
UBE4B
KIF1B
APITD1
PEX14
CLSTN1
LZIC
RBP7
PGD
DFFA
CASZ1
APITD1−CORT
CORT
Position on chr1 (Mb)

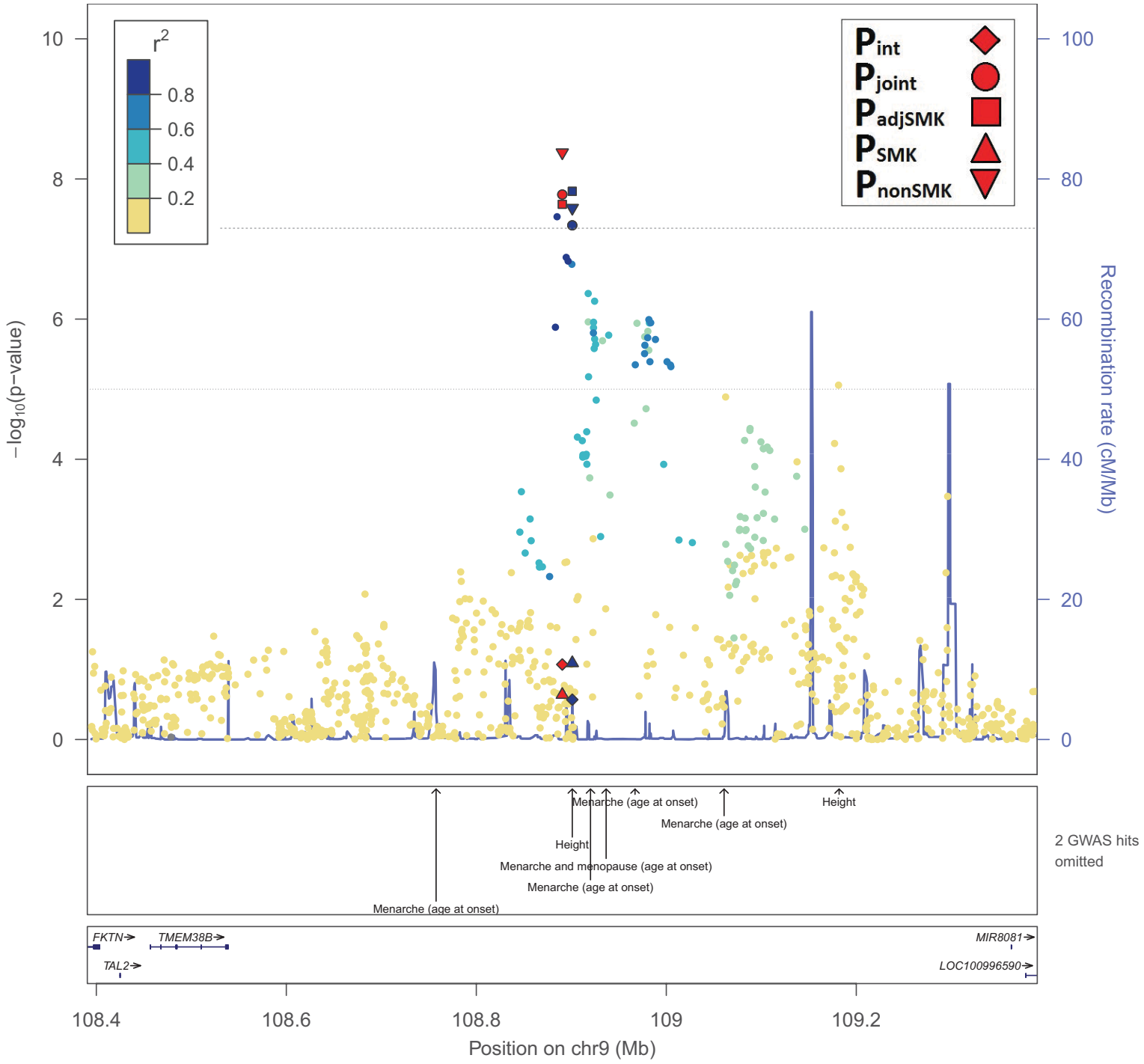
WCadjBMI: rs6743226 − Approach 2



WCadjBMI: rs7697556 − Approach 2

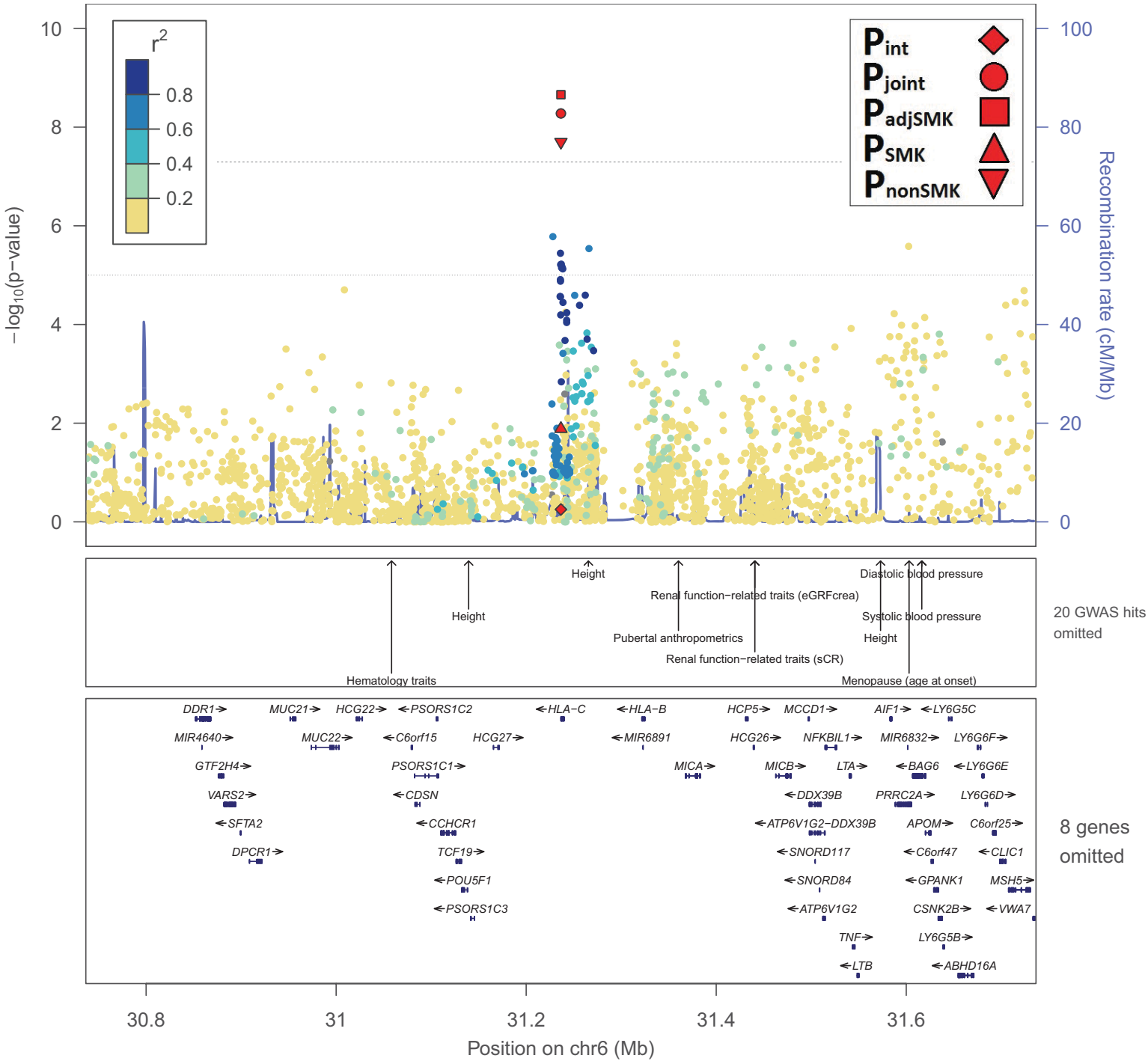


WCadjBMI: rs9408815 − Approach 2



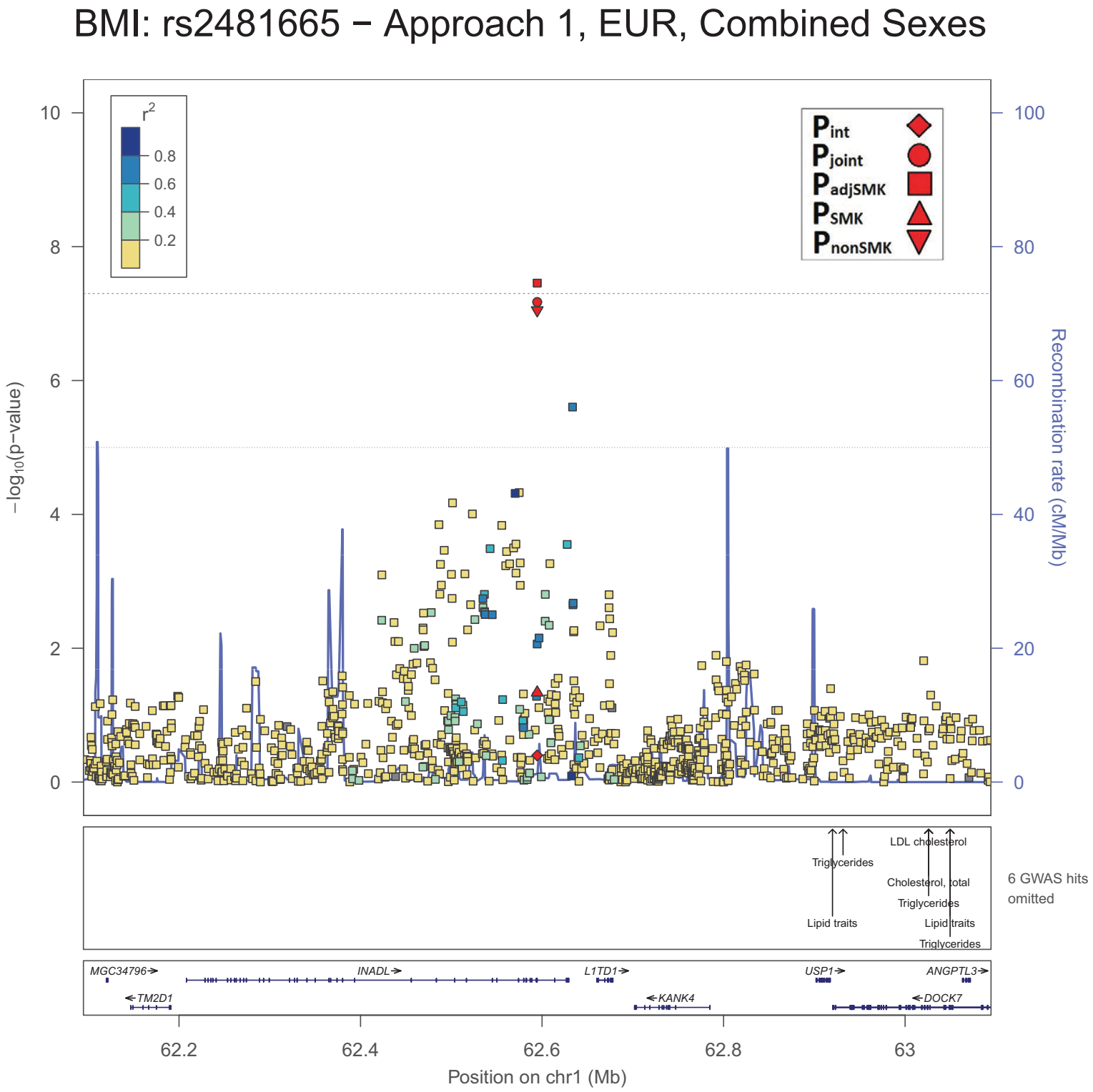
C)

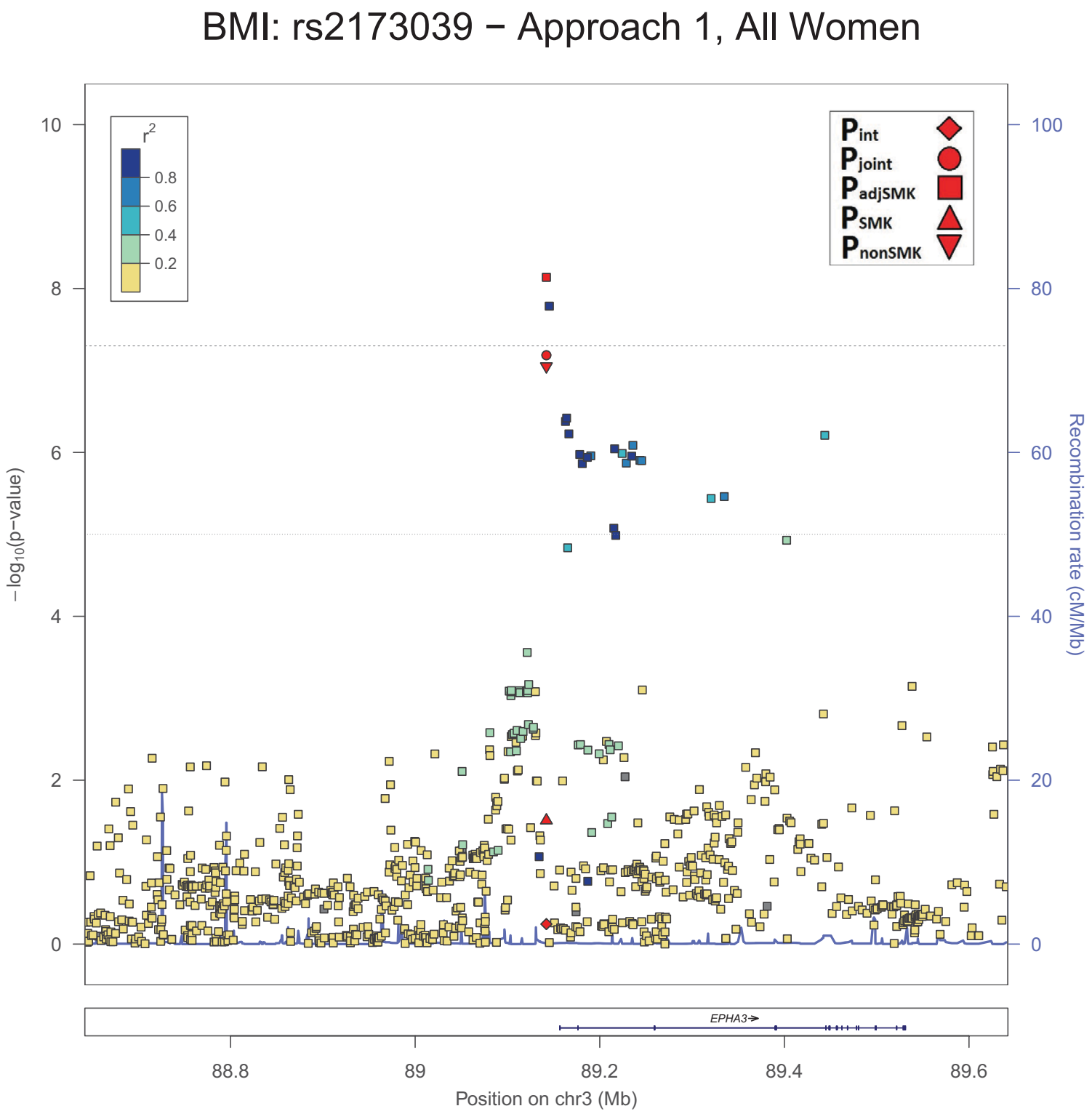
WHRadjBMI: rs1049281 − Approach 2



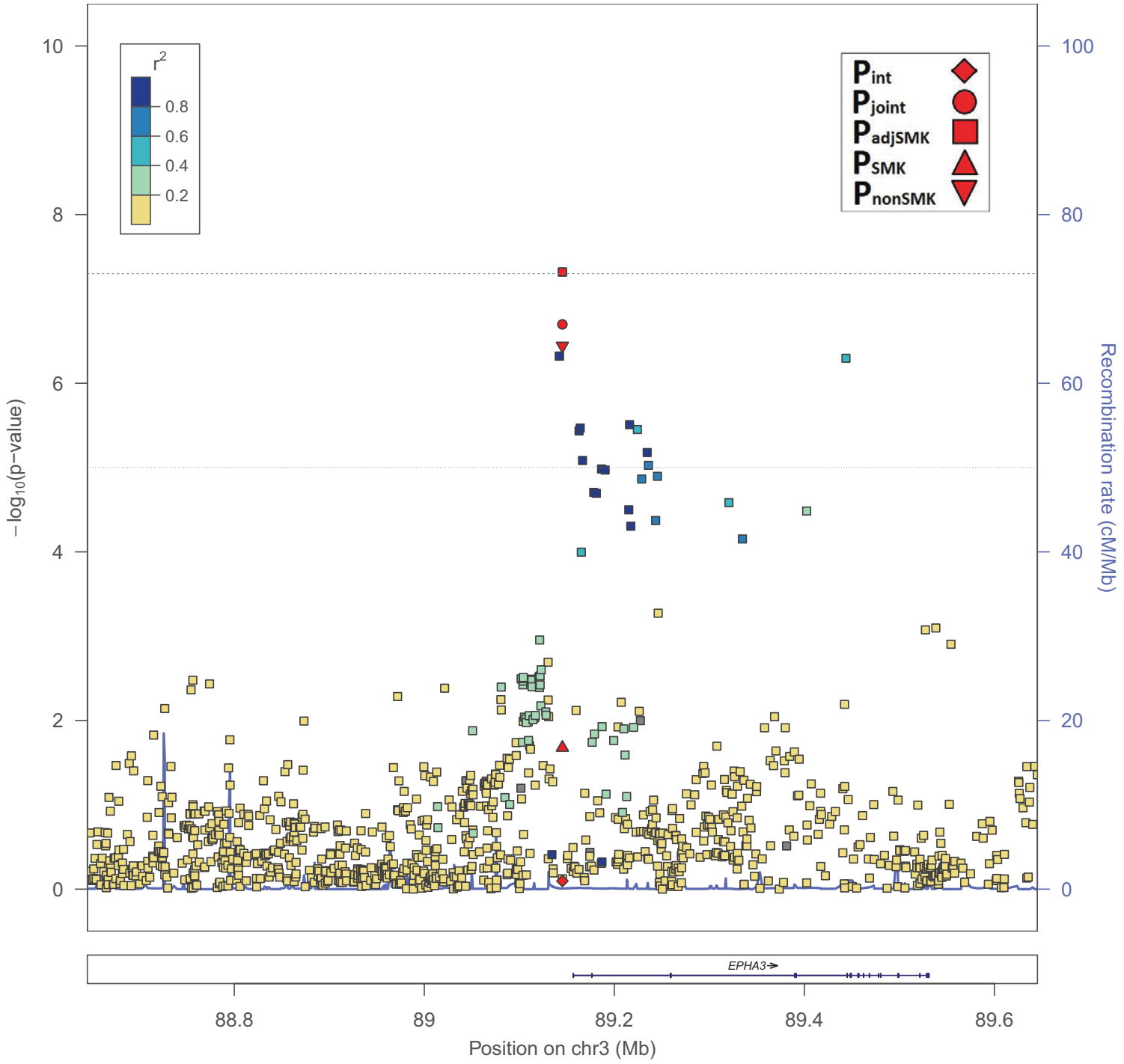
Supplementary Figure 6. Regional association plot for all loci identified in Secondary meta-analyses, and ordered as they appear in Tables 2. LD has been calculated using the combined ancestries from the 1000 Genomes Phase 1 reference panel. For comparison, each plot highlights the p-value for the tag SNP in Approach 1 (P_{adjSMK}), Approach 2 (P_{joint}), Approach 3 (P_{int}), current smokers (P_{SMK}), and in nonsmokers (P_{nonSMK}). P-values are shown from the strata in which the signal was identified (e.g. European-only women).

A.

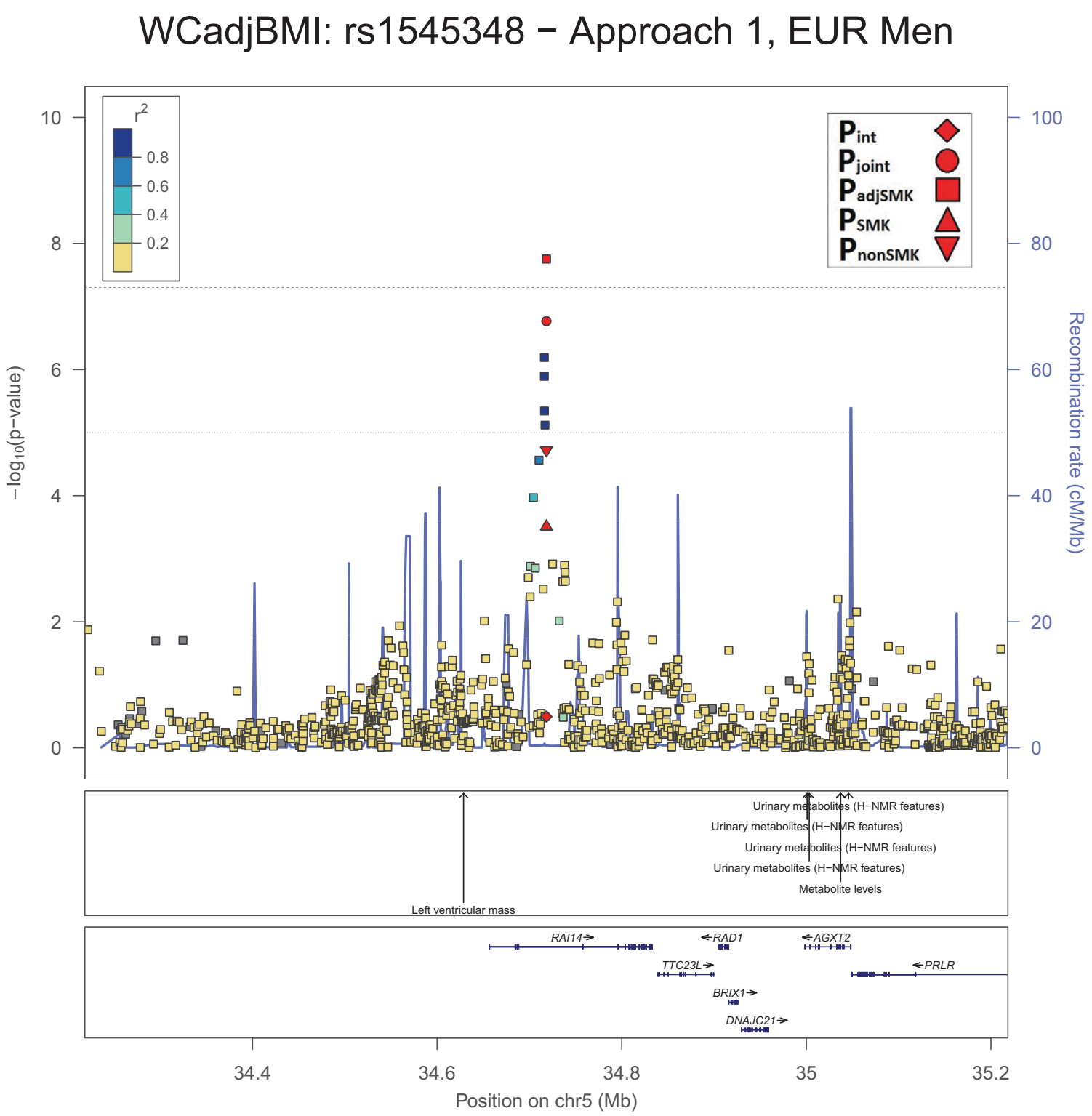

BMI: rs2481665 − Approach 1, EUR, Combined Sexes
r2
0.8
0.6
0.4
0.2
Pint
Pjoint
PadjSMK
PSMK
PnonSMK
−log10(p−value)
Recombination rate (cM/Mb)
10
8
6
4
2
0
100
80
60
40
20
0
Triglycerides
Lipid traits
LDL cholesterol
Cholesterol, total
Triglycerides
Lipid traits
Triglycerides
6 GWAS hits omitted
MGC34796
INADL
L1TD1
USP1
ANGPTL3
TM2D1
KANK4
DOCK7
62.2
62.4
62.6
62.8
63
Position on chr1 (Mb)


BMI: rs2173039 − Approach 1, All Women
r2
0.8
0.6
0.4
0.2
Pint
Pjoint
PadjSMK
PSMK
PnonSMK
10
8
6
4
2
0
100
80
60
40
20
0
−log10(p−value)
Recombination rate (cM/Mb)
EPHA3
88.8
89
89.2
89.4
89.6
Position on chr3 (Mb)

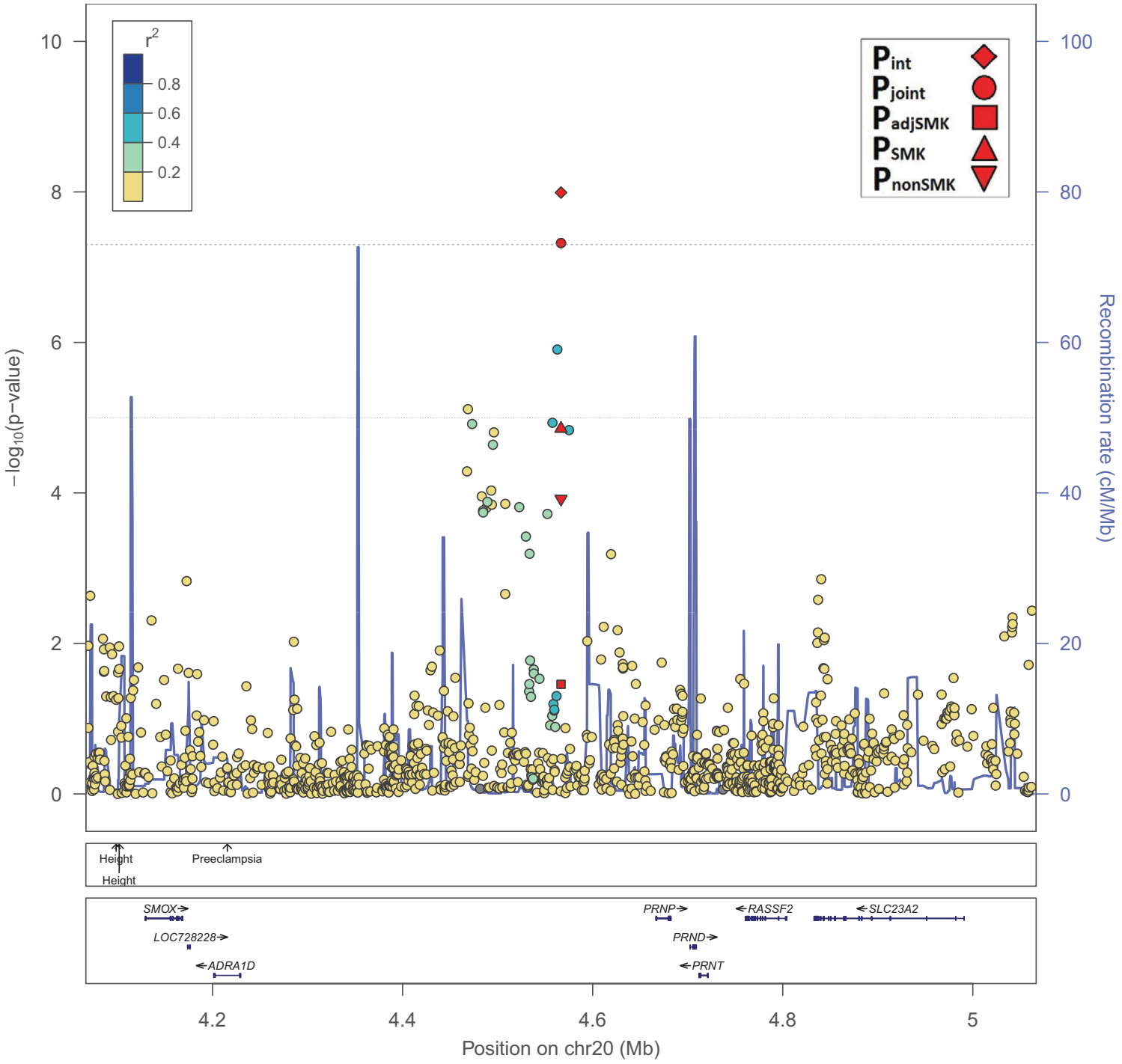
BMI: rs12629427 − Approach 1, EUR Women



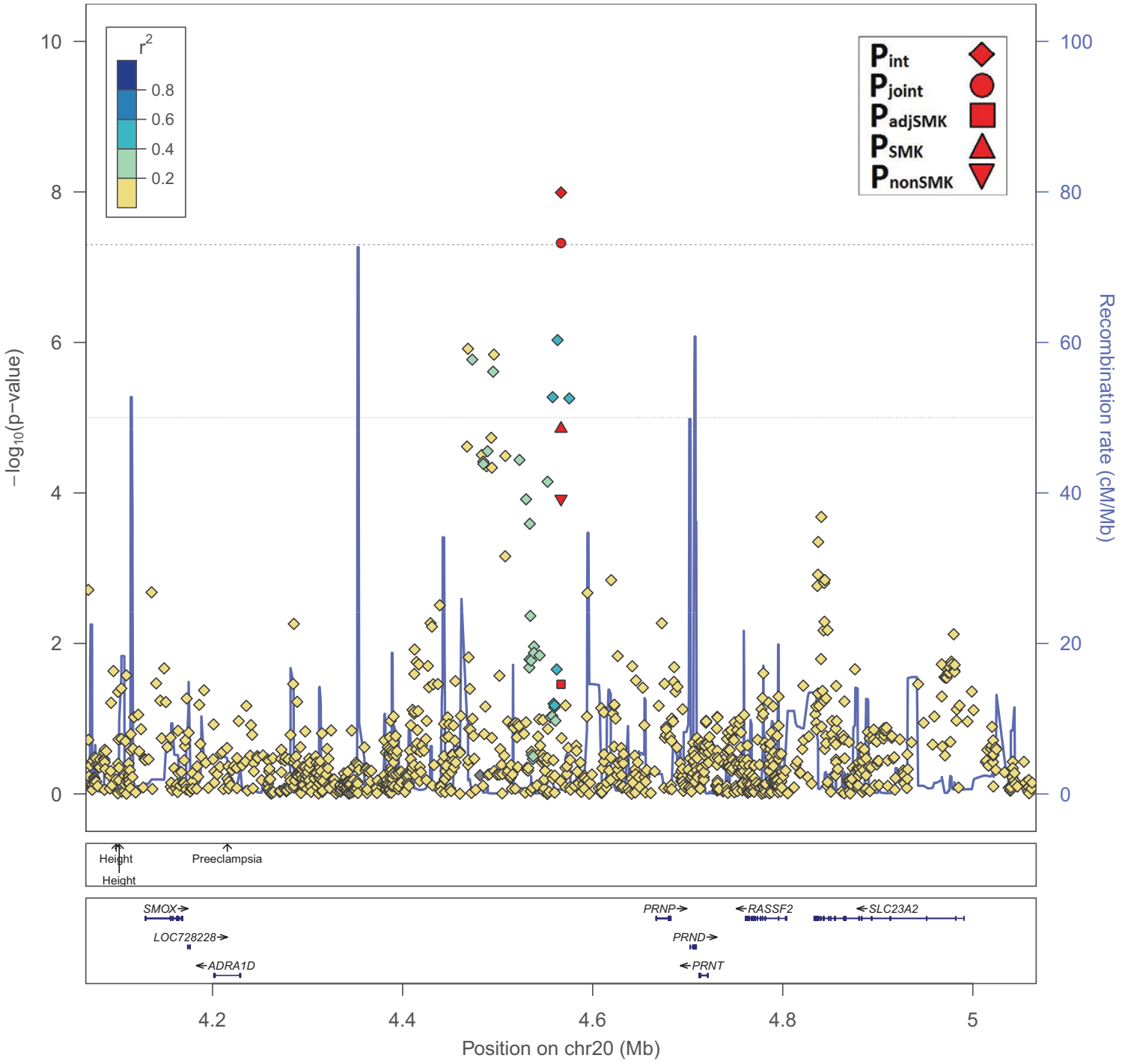
B.


WCadjBMI: rs1545348 − Approach 1, EUR Men
r2
0.8
0.6
0.4
0.2
Pint
Pjoint
PadjSMK
PSMK
PnonSMK
−log10(p−value)
Recombination rate (cM/Mb)
Urinary metabolites (H−NMR features)
Urinary metabolites (H−NMR features)
Urinary metabolites (H−NMR features)
Urinary metabolites (H−NMR features)
Metabolite levels
Left ventricular mass
RAI14
RAD1
AGXT2
TTC23L
PRLR
BRIX1
DNAJC21
Position on chr5 (Mb)

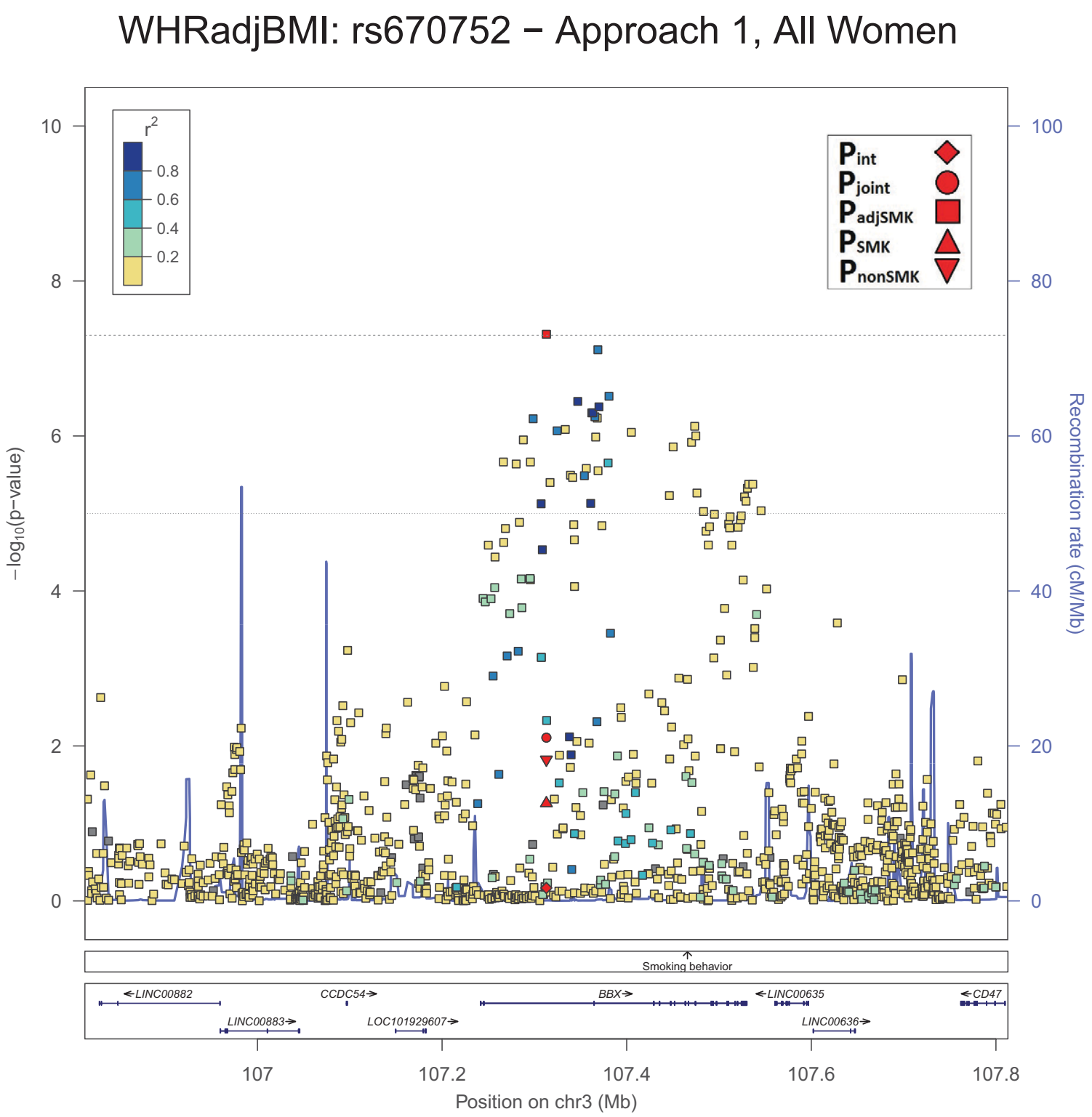
WCadjBMI: rs6076699 − Approach 2, EUR Women



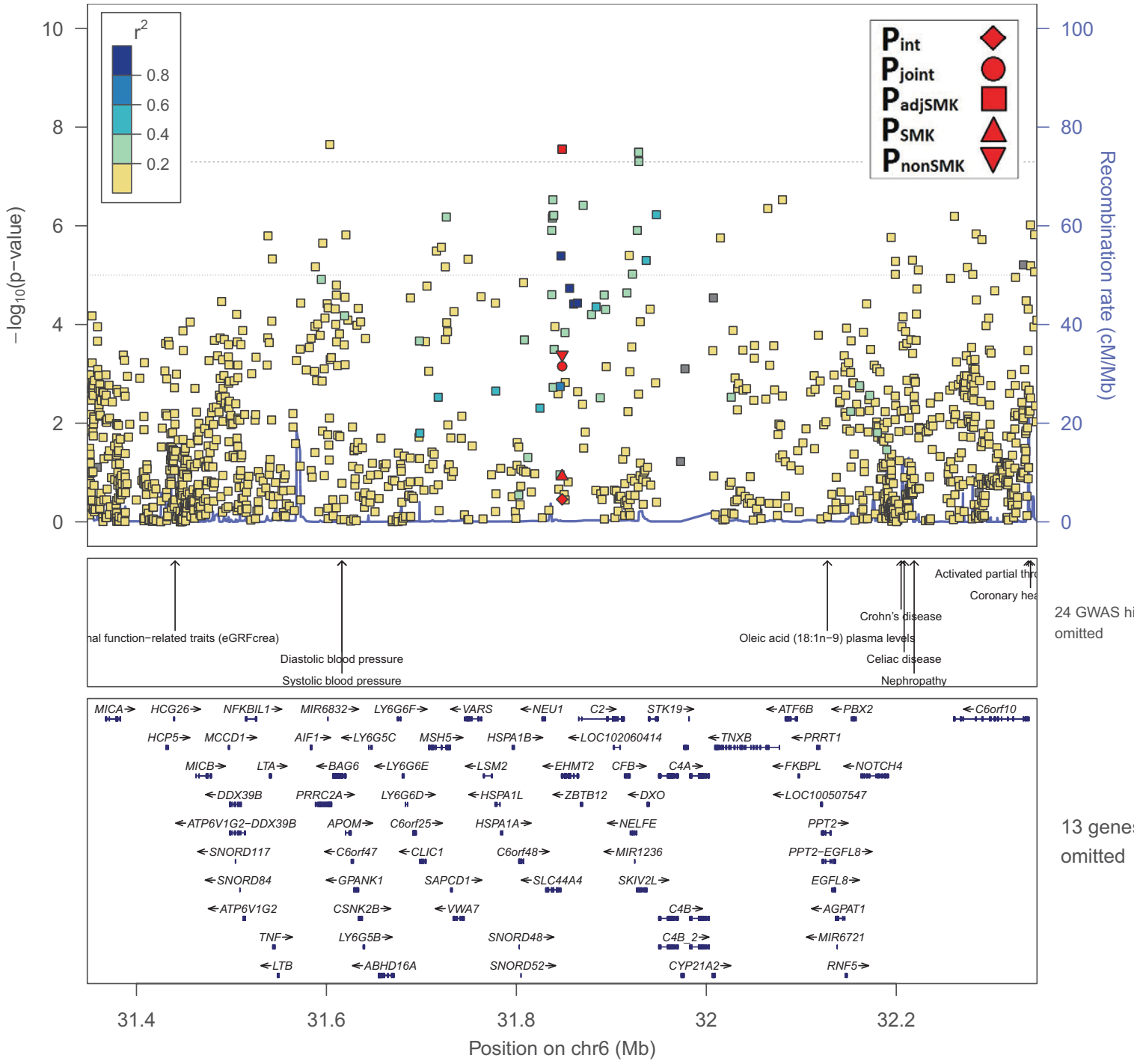
WCadjBMI: rs6076699 − Approach 3, EUR Women



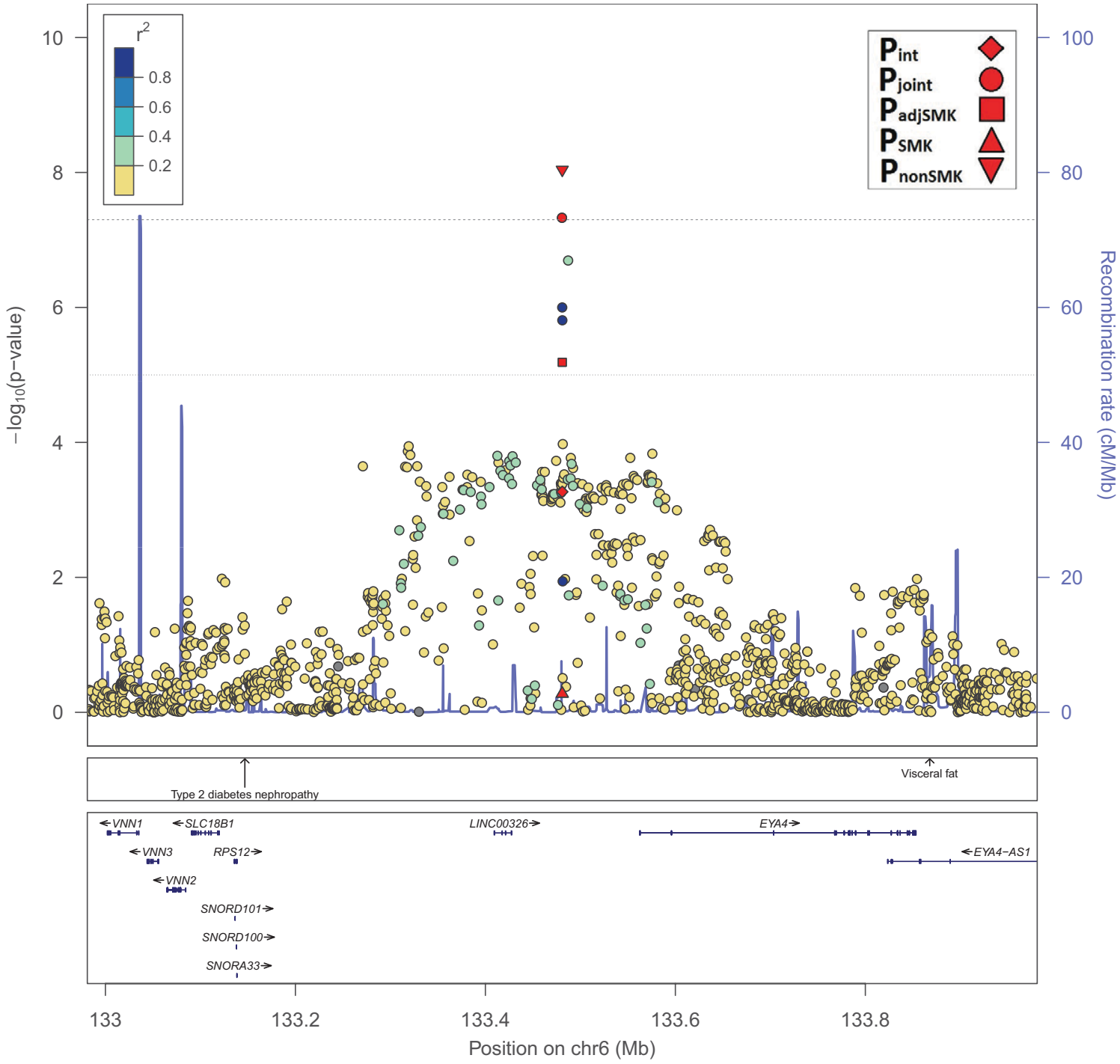
C.


WHRadjBMI: rs670752 − Approach 1, All Women
r2
0.8
0.6
0.4
0.2
Pint
Pjoint
PadjSMK
PSMK
PnonSMK
−log10(p−value)
Recombination rate (cM/Mb)
Smoking behavior
LINC00882
CCDC54
BBX
LINC00635
CD47
LINC00883
LOC101929607
LINC00636
Position on chr3 (Mb)

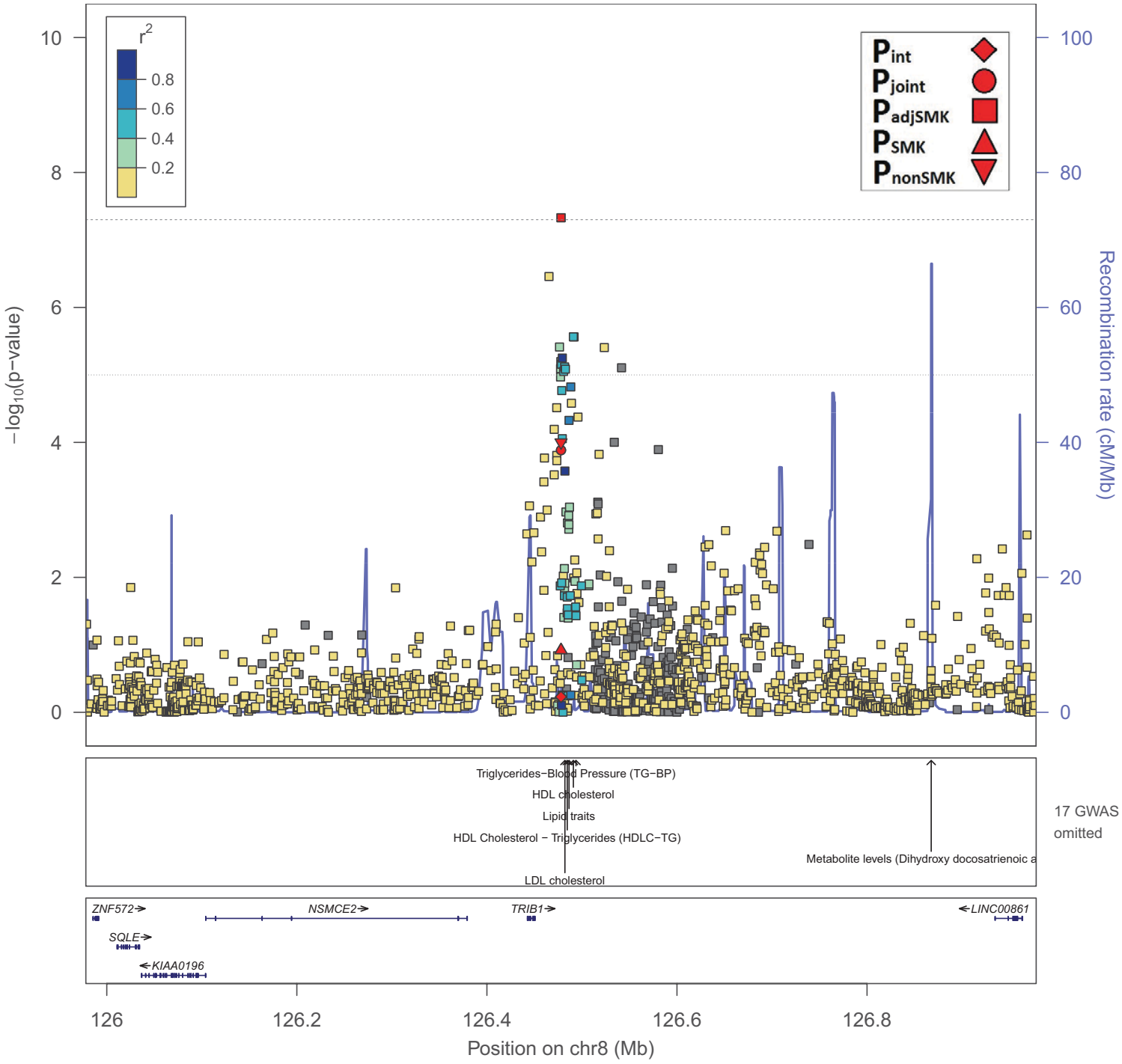
WHRadjBMI: rs589428 − Approach 1, EUR Combined Sexes



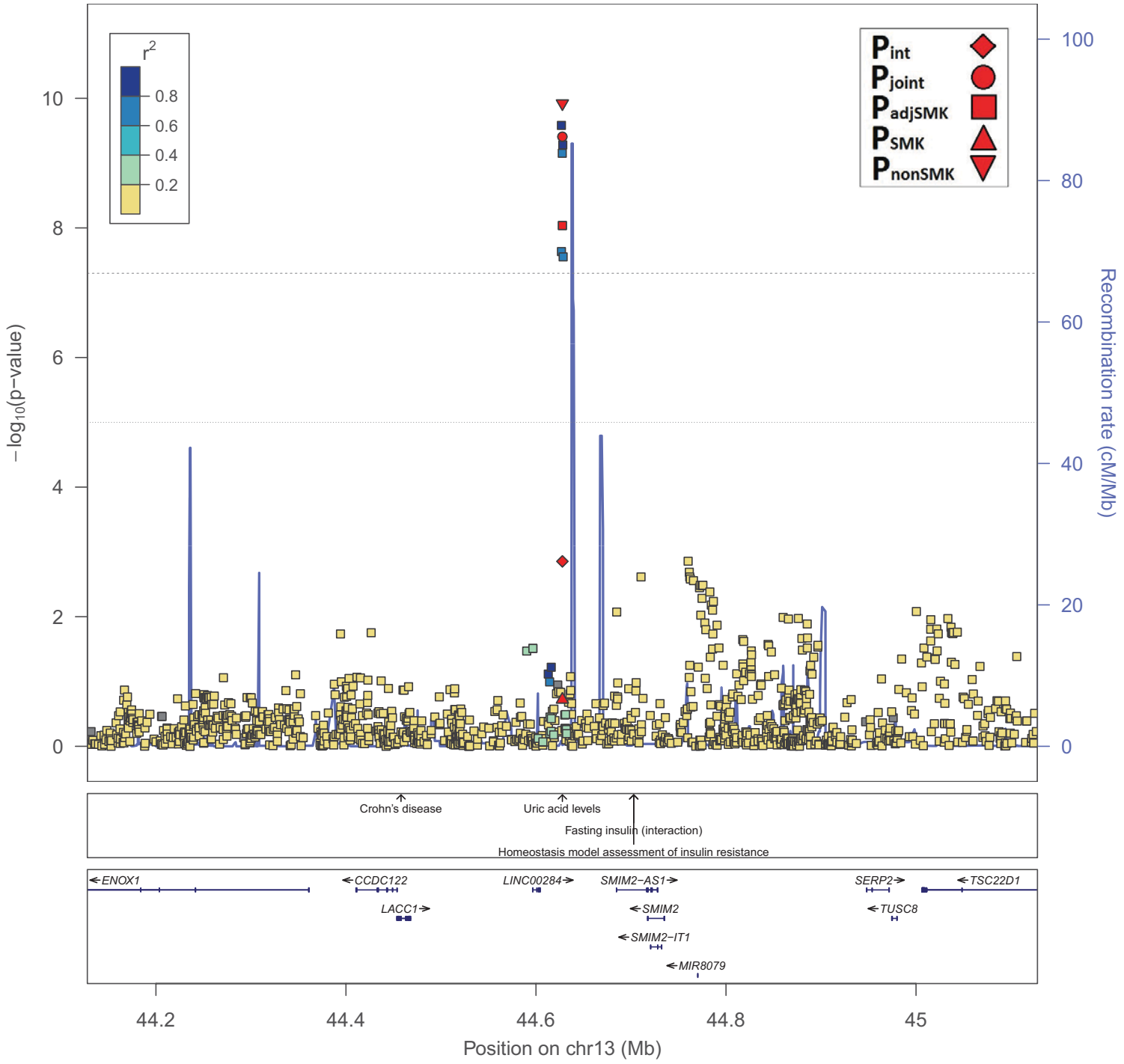
WHRadjBMI: rs1856293 − Approach 2 EUR Combined Sexes



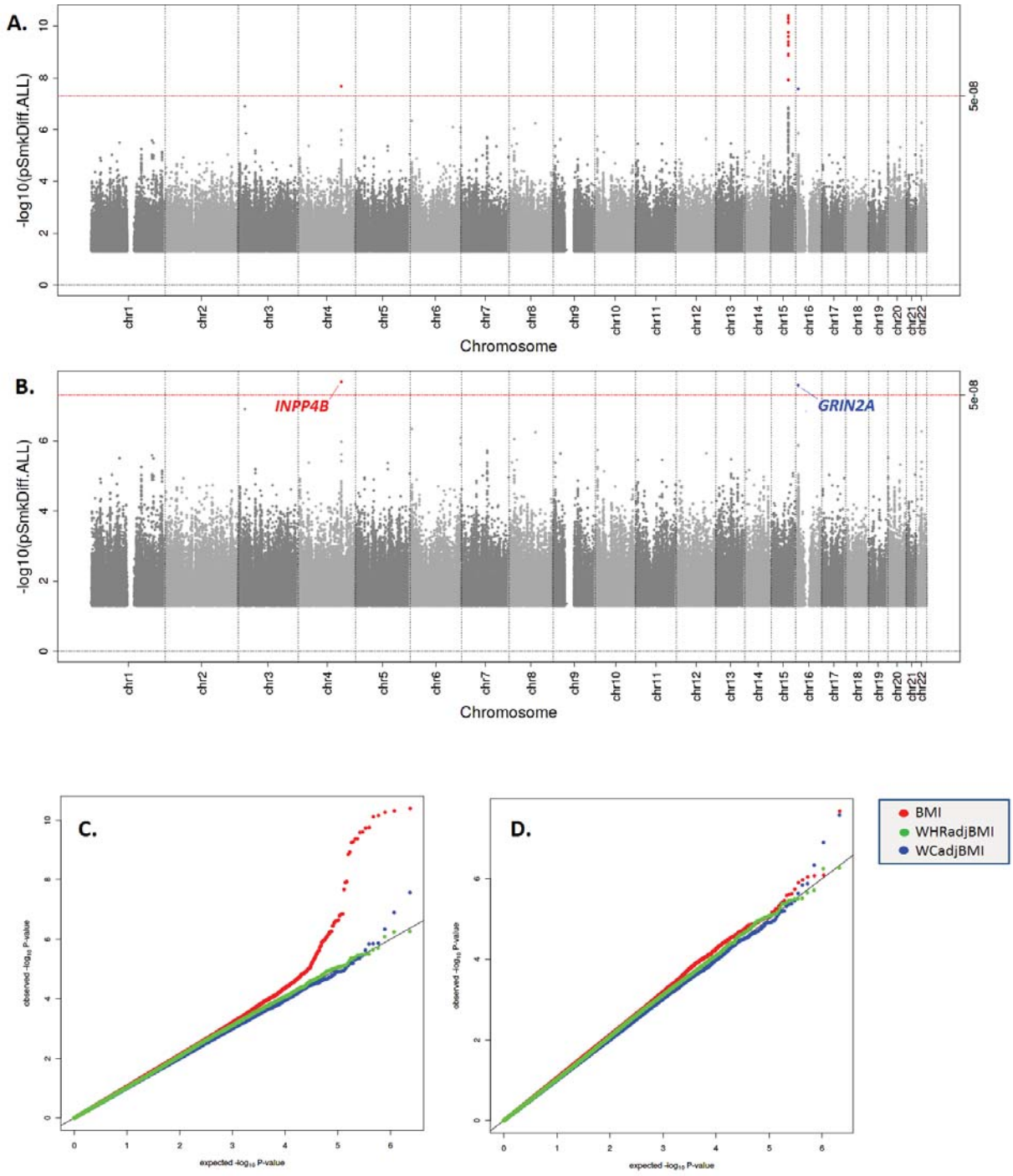
WHRadjBMI: rs2001945 − Approach 1, All Women



WHRadjBMI: rs17065323 − Approach 1, EUR Combined Sexes

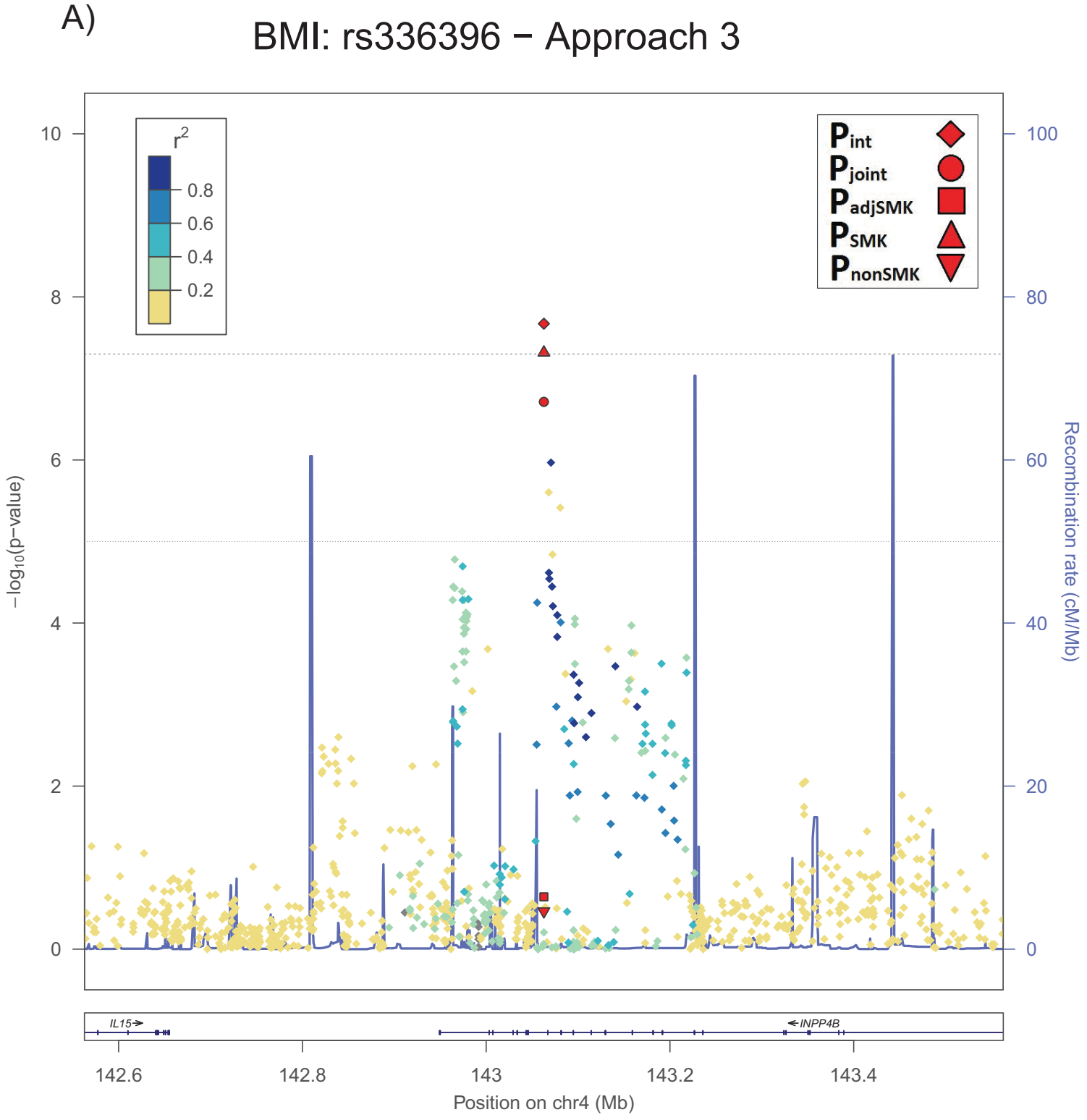


Supplementary Figure 7. Summary plots of discovery meta-analysis for Approach 3 primary meta-analyses. (A) Manhattan plot showing the loci identified in Approach 2 in primary meta-analyses, used to identify significant interaction effects loci (SNPint), in the primary meta-analyses association –log10P-values for BMI-red, WCadjBMI-blue, and WHRadjBMI-green; (B) Manhattan plot showing the loci identified in Approach 2 excluding known regions +/- 500 kb and labeled with the nearest gene to the index SNP; (C) QQ-plot showing the Approach 2 P-values as observed against those expected under the null for each phenotypes separately (colored); (D) QQ-plot for Approach 2 after excluding known association regions.

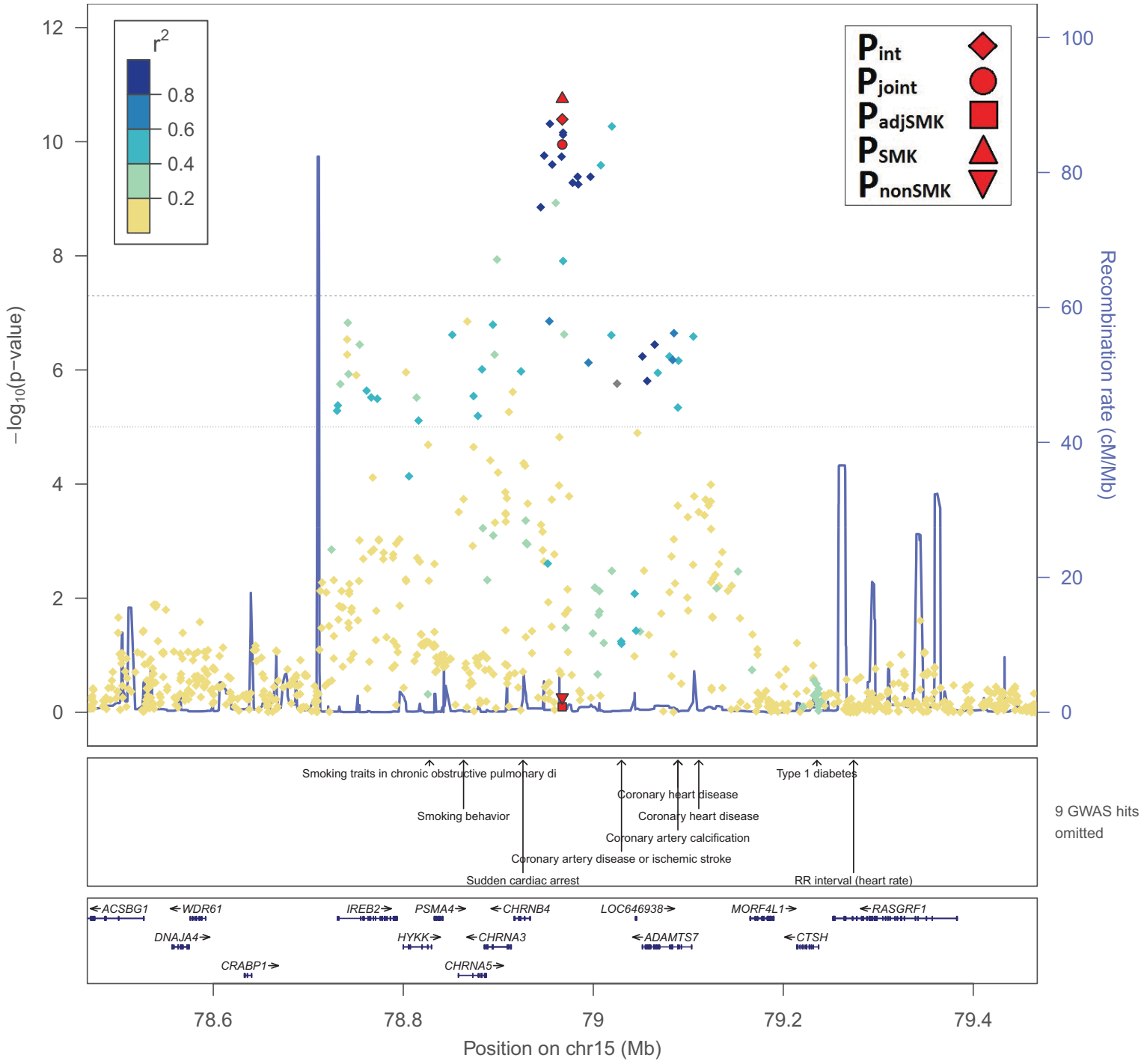


Supplementary Figure 8. Regional association plot for all loci identified in Approach 3 in primary meta-analyses, used to identify significant interaction (SNPint), in the primary meta-analyses for A) BMI and B) WCadjBMI, and ordered as they appear in Table 3. LD has been calculated using the combined ancestries from the 1000 Genomes Phase 1 reference panel. For comparison, each plot highlights the p-value for the tag SNP in Approach 1 (P_{adjSMK}), Approach 2 (P_{joint}), Approach 3 (P_{int}), current smokers (P_{SMK}), and in nonsmokers (P_{nonSMK}). EUR-European-only meta-analysis.

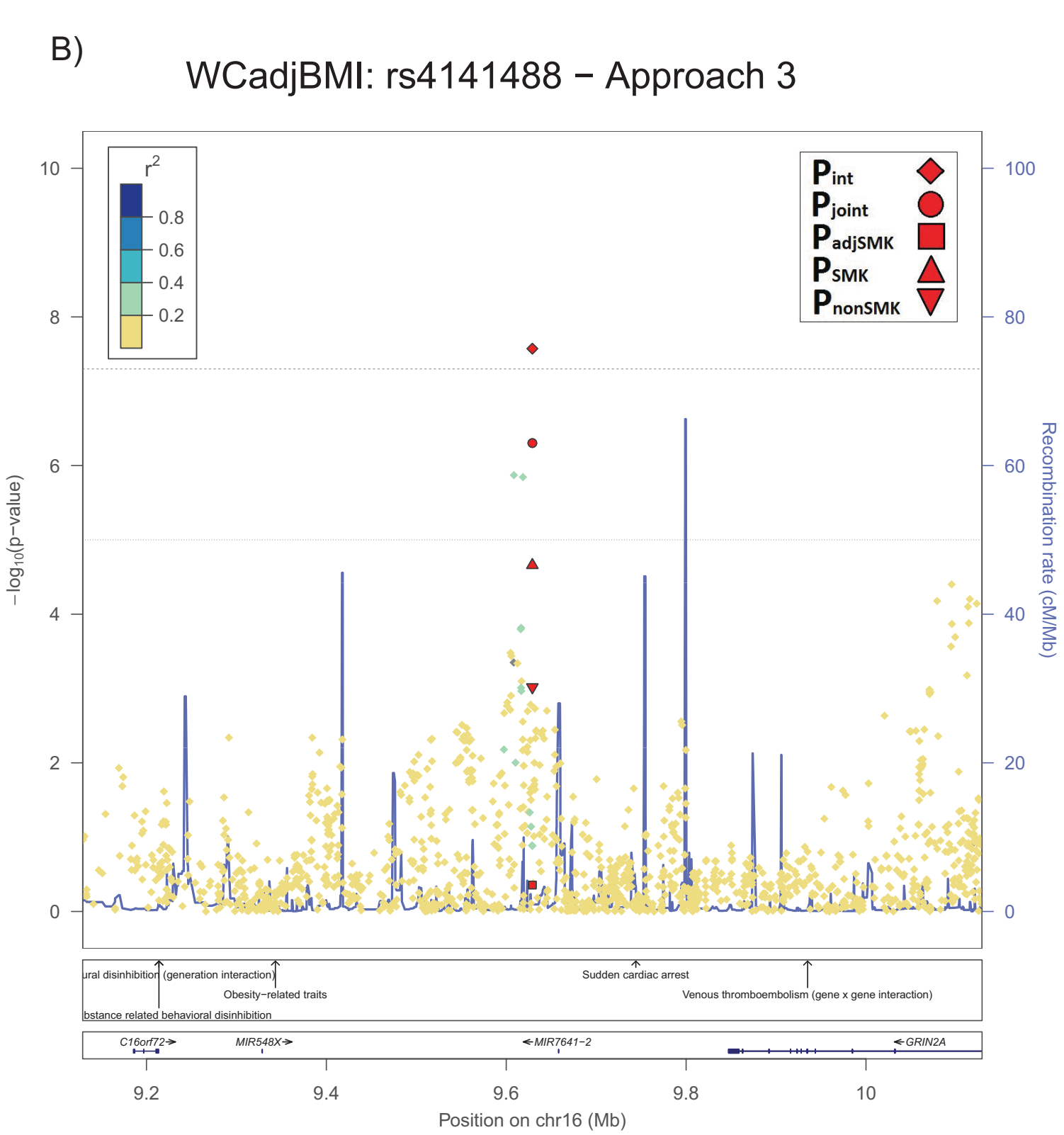
A)


BMI: rs336396 − Approach 3
r2
0.8
0.6
0.4
0.2
Pint
Pjoint
PadjSMK
PSMK
PnonSMK
10
8
6
4
2
0
100
80
60
40
20
0
−log10(p−value)
Recombination rate (cM/Mb)
IL15
INPP4B
142.6
142.8
143
143.2
143.4
Position on chr4 (Mb)

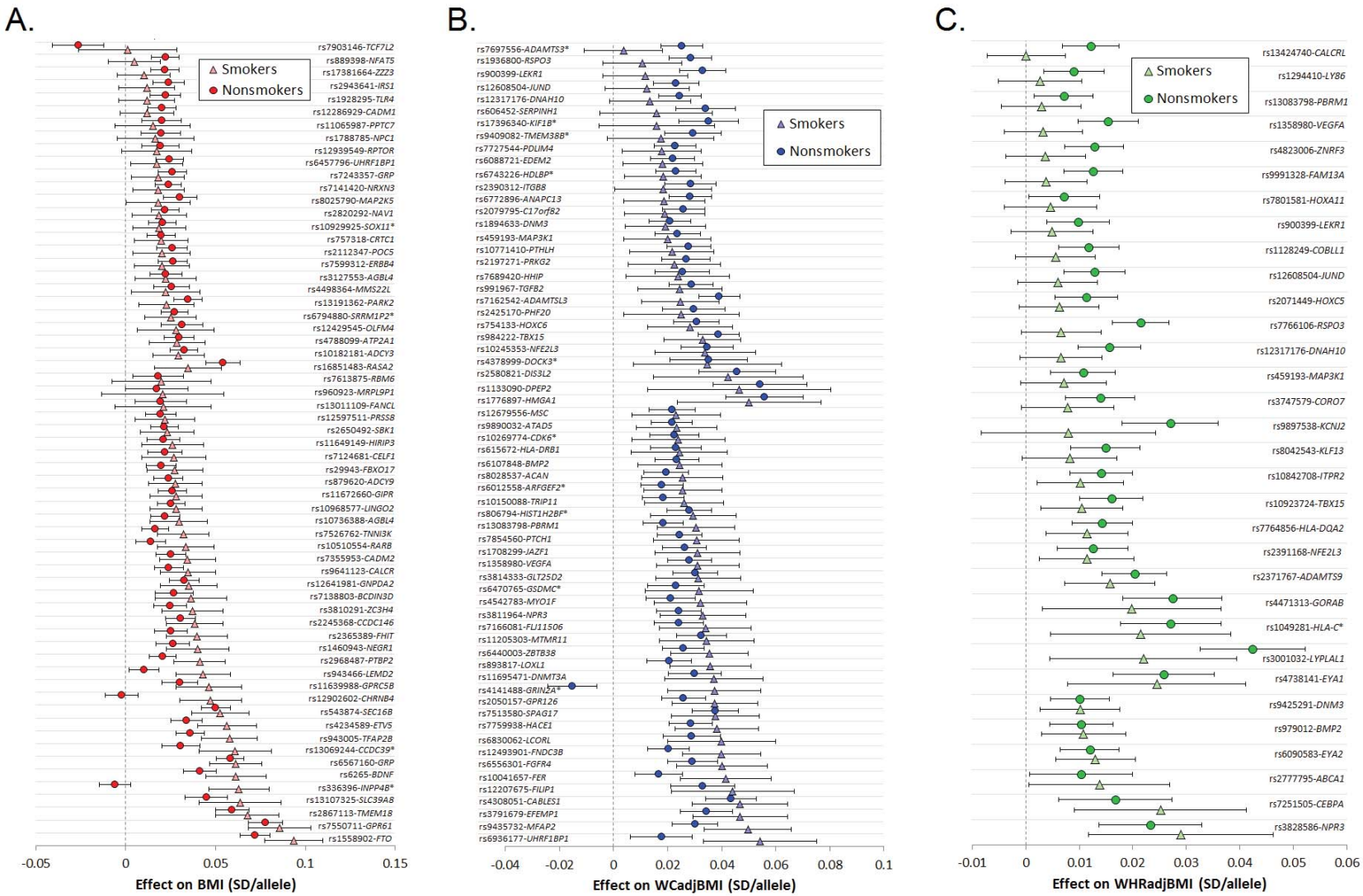
BMI: rs12902602 − Approach 3



B)


WCadjBMI: rs4141488 − Approach 3
r2
0.8
0.6
0.4
0.2
Pint
Pjoint
PadjSMK
PSMK
PnonSMK
−log10(p−value)
Recombination rate (cM/Mb)
ural disinhibition
(generation interaction)
Obesity−related traits
Sudden cardiac arrest
Venous thromboembolism (gene x gene interaction)
bstance related behavioral disinhibition
C16orf72
MIR548X
MIR7641−2
GRIN2A
Position on chr16 (Mb)

Supplementary Figure 9. Estimated effects (β ± 95% CI) per risk allele for A) BMI, B) WCadjBMI, and C) WHRadjBMI for the most significant variant for each locus identified in the primary meta-analyses (combined ancestries and combined sexes) for Approaches 1 (SNPadjSMK), 2 (SNPjoint) and 3 (SNPint). Loci are ordered by greater magnitude of effect in smokers compared to nonsmokers and labeled with the nearest gene.



Supplementary Figure 10. Estimated effect estimates (β ± 95% CI) per risk allele for A) BMI, B) WCadjBMI, and C) WHRadjBMI for the most significant variant for each locus identified in the secondary meta-analyses (sex-stratified and European-only analyses) for Approaches 1 (SNPadjSMK), 2 (SNPjoint) and 3 (SNPint). Loci are ordered by greater magnitude of effect in smokers compared to nonsmokers and labeled with the nearest gene.