Genetic Architecture of Subcortical Brain Structures in 38,851 Individuals

Claudia L Satizabal*[1-4], Hieab HH Adams*[5-7], Derrek P Hibar*[8], Charles C White*[9,10], Maria J Knol[5], Jason L Stein[8,11,12], Markus Scholz[13,14], Muralidharan Sargurupremraj[15], Neda Jahanshad[8], Gennady V Roshchupkin[5,6,16], Albert V Smith[17-19], Joshua C Bis[20], Xueqiu Jian[21], Michelle Luciano[22], Edith Hofer[23,24], Alexander Teumer[25], Sven J van der Lee[5], Jingyun Yang[26,27], Lisa R Yanek[28], Tom V Lee[29], Shuo Li[30], Yanhui Hu[31], Jia Yu Koh[32], John D Eicher[33], Sylvane Desrivières[34], Alejandro Arias-Vasquez[35-38], Ganesh Chauhan[39,40], Lavinia Athanasiu[41,42], Miguel E Rentería[43], Sungeun Kim[44-46], David Hoehn[47], Nicola J Armstrong[48], Qiang Chen[49], Avram J Holmes[50,51], Anouk den Braber[52-55], Iwona Kloszewska[56], Micael Andersson[57,58], Thomas Espeseth[59,60], Oliver Grimm[61], Lucija Abramovic[62], Saud Alhusaini[63,64], Yuri Milaneschi[65], Martina Papmeyer[66,67], Tomas Axelsson[68], Stefan Ehrlich[51,69,70], Roberto Roiz-Santiañez[71-73], Bernd Kraemer[74], Asta K Håberg[75,76], Hannah J Jones[77-79], G Bruce Pike[80,81], Dan J Stein[82,83], Allison Stevens[70], Janita Bralten[36,38], Meike W Vernooij[5,6], Tamara B Harris[84], Irina Filippi[85], A Veronica Witte[86,87], Tulio Guadalupe[88,89], Katharina Wittfeld[90,91], Thomas H Mosley[92], James T Becker[93-95], Nhat Trung Doan[96], Saskia P Hagenaars[22], Yasaman Saba[97], Gabriel Cuellar-Partida[98], Najaf Amin[5], Saima Hilal[99,100], Kwangsik Nho[44-46], Nazanin Mirza-Schreiber[47,101], Konstantinos Arfanakis[26,102,103], Diane M Becker[28], David Ames[104,105], Aaron L Goldman[49], Phil H Lee[51,106-109], Dorret I Boomsma[52-54,110], Simon Lovestone[111,112], Sudheer Giddaluru[113,114], Stephanie Le Hellard[113,114], Manuel Mattheisen[115-119], Marc M Bohlken[62], Dalia Kasperaviciute[120,121], Lianne Schmaal[122,123], Stephen M Lawrie[66], Ingrid Agartz[96,118,124], Esther Walton[69,125], Diana Tordesillas-

Gutierrez[73,126], Gareth E Davies[127], Jean Shin[128], Jonathan C Ipser[82], Louis N Vinke[129], Martine Hoogman[36,38], Tianye Jia[34], Ralph Burkhardt[14,130], Marieke Klein[36,38], Fabrice Crivello[131], Deborah Janowitz[90], Owen Carmichael[132], Unn K Haukvik[133,134], Benjamin S Aribisala[135,136], Helena Schmidt[97], Lachlan T Strike[98,137], Ching-Yu Cheng[32,138], Shannon L Risacher[45,46], Benno Pütz[47], Debra A Fleischman[26,27,139], Amelia A Assareh[140], Venkata S Mattay[49,141,142], Randy L Buckner[51,143], Patrizia Mecocci[144], Anders M Dale[145-149], Sven Cichon[150-152], Marco P Boks[62], Mar Matarin[120,153,154], Brenda WJH Penninx[65], Vince D Calhoun[155-157], M Mallar Chakravarty[158,159], Andre Marquand[38,160], Christine Macare[34], Shahrzad Kharabian Masouleh[86,161], Jaap Oosterlaan[162-164], Philippe Amouyel[165-168], Katrin Hegenscheid[169], Jerome I Rotter[170], Andrew J Schork[171,172], David CM Liewald[22], Greig I De Zubicaray[173,174], Tien Yin Wong[32,175], Li Shen[175,176], Philipp G Sämann[47], Henry Brodaty[140,177], Joshua L Roffman[51], Eco JC De Geus[52-54,110], Magda Tsolaki[178], Susanne Erk[179], Kristel R Van Eijk[180], Gianpiero L Cavalleri[181], Nic JA Van der Wee[182,183], Andrew M McIntosh[22,66], Randy L Gollub[51,70,106], Kazima B Bulayeva[184], Manon Bernard[128], Jennifer S Richards[35,38,185], Jayandra J Himali[3,4,30], Markus Loeffler[13,14], Nanda Rommelse[37,38,186], Wolfgang Hoffmann[91,187], Lars T Westlye[59,188], Maria C Valdés Hernández[135,189], Narelle K Hansell[98,137], Theo GM Van Erp[190,191], Christiane Wolf[192], John BJ Kwok[193-195], Bruno Vellas[196,197], Andreas Heinz[198], Loes M Olde Loohuis[199], Norman Delanty[63,200], Beng-Choon Ho[201], Christopher RK Ching[8,202], Elena Shumskaya[36,38,160], Baljeet Singh[203], Albert Hofman[5,204], Dennis Van der Meer[205,206], Georg Homuth[207], Bruce M Psaty[20,208-210], Mark E Bastin[135,189], Grant W Montgomery[211], Tatiana M Foroud[46,212], Simone Reppermund[140,213], Jouke-Jan Hottenga[52-54,110], Andrew Simmons[214-216], Andreas Meyer-Lindenberg[61], Wiepke

Cahn[62], Christopher D Whelan[8,63], Marjolein MJ Van Donkelaar[36,38], Qiong Yang[30], Norbert Hosten[169], Robert C Green[106,217], Anbupalam Thalamuthu[140], Sebastian Mohnke[179], Hilleke E Hulshoff Pol[62], Honghuang Lin[3,218], Clifford R Jack Jr[219], Peter R Schofield[194,220], Thomas W Mühleisen[152,221,222], Pauline Maillard[203], Steven G Potkin[223], Wei Wen[140], Evan Fletcher[203], Arthur W Toga[224], Oliver Gruber[74], Matthew Huentelman[172], George Davey Smith[78], Lenore J Launer[84], Lars Nyberg[57,225,226], Erik G Jönsson[96,118], Benedicto Crespo-Facorro[72,73], Nastassja Koen[82,83], Douglas N Greve[70,227], André G Uitterlinden[5,228], Daniel R Weinberger[49,141,229-231], Vidar M Steen[113,114], Iryna O Fedko[52,53,110], Nynke A Groenewold[82], Wiro J Niessen[6,16,232], Roberto Toro[233], Christophe Tzourio[40], William T Longstreth Jr[209,234], M Kamran Ikram[5,235], Jordan W Smoller[51,106,108,109], Marie-Jose Van Tol[236], Jessika E Sussmann[66], Tomas Paus[237-239], Hervé Lemaître[85], Matthias L Schroeter[14,86,240], Bernard Mazoyer[131], Ole A Andreassen[59,96], Florian Holsboer[47,241], Chantal Depondt[242], Dick J Veltman[65], Jessica A Turner[156,157,243], Zdenka Pausova[128], Gunter Schumann[34], Daan Van Rooij[35,38,185], Srdjan Djurovic[113,244], Ian J Deary[22], Katie L McMahon[173,174], Bertram Müller-Myhsok[47,245,246], Rachel M Brouwer[62], Hilkka Soininen[247,248], Massimo Pandolfo[242], Thomas H Wassink[201], Joshua W Cheung[8], Thomas Wolfers[36,38], Jean-Luc Martinot[249], Marcel P Zwiers[38,160], Matthias Nauck[250,251], Ingrid Melle[59,96], Nicholas G Martin[98], Ryota Kanai[252-254], Eric Westman[255], René S Kahn[62], Sanjay M Sisodiya[120], Tonya White[6,256], Arvin Saremi[8], Hans van Bokhoven[36,38], Han G Brunner[36,38,257,258], Henry Völzke[25,251], Margaret J Wright[137,259], Dennis Van 't Ent[52-54,110], Markus M Nöthen[151,260], Roel A Ophoff[62,199], Jan K Buitelaar[35,38,186], Guillén Fernández[35,38], Perminder S Sachdev[140,261], Marcella Rietschel[61], Neeltje EM Van Haren[62,256], Simon E Fisher[38,89], Alexa S Beiser[3,4,30], Clyde Francks[38,89],

Andrew J Saykin[45,46,212], Karen A Mather[140,194], Nina Romanczuk-Seiferth[198], Catharina A Hartman[185], Anita L DeStefano[3,30], Dirk J Heslenfeld[262], Michael W Weiner[263,264], Henrik Walter[179], Pieter J Hoekstra[185], Paul A Nyquist[28], Barbara Franke[36-38], David A Bennett[26,27], Hans J Grabe[90,91], Andrew D Johnson[33], Christopher Chen[99,100], Cornelia M van Duijn[5,265], Oscar L Lopez[94,95], Myriam Fornage[21,266], Joanna M Wardlaw[22,189,267], Reinhold Schmidt[23], Charles DeCarli[268], Philip L De Jager[269,270], Arno Villringer[86,87], Stéphanie Debette[40], Vilmundur Gudnason[18,19], Sarah E Medland[98]**, Joshua M Shulman[29,271-273]**, Paul M Thompson[8]**, Sudha Seshadri[1,3,4]**, M Arfan Ikram[5,6]**

1.   Glenn Biggs Institute for Alzheimer's & Neurodegenerative Diseases, UT Health San Antonio, San Antonio, Texas, USA.

2.   Department of Population Health Sciences, UT Health San Antonio, San Antonio, Texas, USA.

3.   The Framingham Heart Study, Framingham, Massachusetts, USA.

4.   Department of Neurology, Boston University School of Medicine, Boston, Massachusetts, USA.

5.   Department of Epidemiology, Erasmus MC, Rotterdam, The Netherlands.

6.   Department of Radiology and Nuclear Medicine, Erasmus MC, Rotterdam, The Netherlands.

7.   Department of Clinical Genetics, Erasmus MC, Rotterdam, The Netherlands.

8.   Imaging Genetics Center, USC Mark and Mary Stevens Neuroimaging and Informatics Institute, Keck School of Medicine, University of Southern California, Los Angeles, USA.

9.   Cell Circuits Program, Broad Institute, Cambridge, Massachusetts, USA.

10.   Center for Translational & Computational Neuroimmunology, Department of Neurology, Columbia University Medical Center, New York City, New York, USA.

11.   Department of Genetics, University of North Carolina (UNC), Chapel Hill, North Carolina, USA.

12.   UNC Neuroscience Center, University of North Carolina, Chapel Hill, North Carolina, USA.

13.   Institute for Medical Informatics, Statistics and Epidemiology, University of Leipzig, Leipzig, Germany.

14.   LIFE - Leipzig Research Center for Civilization Diseases, University of Leipzig, Leipzig, Germany.

15.   INSERM U1219, VINTAGE team, University of Bordeaux, Bordeaux, France.

16.   Department of Medical Informatics, Erasmus MC, Rotterdam, The Netherlands.

17.   Department of Biostatistics, University of Michigan, Ann Arbor, Michigan, USA.

18.   Faculty of Medicine, University of Iceland, Iceland.

19.   Icelandic Heart Association, Iceland.

20.   Cardiovascular Health Research Unit, Department of Medicine, University of Washington, Seattle, Washington, USA.

21.   The Brown Foundation Institute of Molecular Medicine, University of Texas Health Science Center at Houston, Houston, Texas, USA.

22.   Centre for Cognitive Ageing and Cognitive Epidemiology, Psychology, University of Edinburgh, Edinburgh, UK.

23.   Clinical Division of Neurogeriatrics, Department of Neurology, Medical University of Graz, Graz, Austria.

24.   Institute for Medical Informatics, Statistics and Documentation, Medical University of Graz, Graz, Austria.

25.   Institute for Community Medicine, University Medicine Greifswald, Greifswald, Germany.

26.   Rush Alzheimer's Disease Center, Rush University Medical Center, Chicago, Illinois, USA.

27.   Department of Neurological Sciences, Rush University Medical Center, Chicago, Illinois, USA.

28.   GeneSTAR Research Program, Department of Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA.

29.   Department of Neurology, Baylor College of Medicine, Houston, Texas, USA.

30.   Department of Biostatistics, Boston University School of Public Health, Boston, Massachusetts, USA.

31.   Department of Genetics, Harvard Medical School, Boston, Massachusetts, USA.

32.   Singapore Eye Research Institute, Singapore National Eye Centre, Singapore.

33.   National Heart, Lung and Blood Institute's Framingham Heart Study, Division of Intramural Research, Population Sciences Branch, Framingham, Massachusetts, USA.

34.   MRC-SGDP Centre, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK.

35.   Department of Cognitive Neuroscience, Radboud University Medical Center, Nijmegen, The Netherlands.

36.   Department of Human Genetics, Radboud University Medical Center, Nijmegen, The Netherlands.

37.   Department of Psychiatry, Radboud University Medical Center, Nijmegen, The Netherlands.

38.   Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, The Netherlands.

39.   Centre for Brain Research, Indian Institute of Science, Bangalore, India.

40.   Univ. Bordeaux, Inserm, Bordeaux Population Health Research Center, UMR 1219, CHU Bordeaux, Bordeaux, France.

41.   CoE NORMENT, Division of Mental Health and Addiction, Oslo University Hospital, Oslo, Norway.

42. CoE NORMENT, Institute of Clinical Medicine, University of Oslo, Oslo, Norway.

43. Department of Genetics and Computational Biology, QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia.

44. Center for Computational Biology and Bioinformatics, Indiana University School of Medicine, Indianapolis, Indiana, USA.

45. Center for Neuroimaging, Radiology and Imaging Sciences, Indiana University School of Medicine, Indianapolis, Indiana, USA.

46. Indiana Alzheimer Disease Center, Indiana University School of Medicine, Indianapolis, Indiana, USA.

47. Max Planck Institute of Psychiatry, Munich, Germany.

48. Mathematics and Statistics, Murdoch University, Perth, Western Australia, Australia.

49. Lieber Institute for Brain Development, Baltimore, Maryland, USA.

50. Department of Psychology, Yale University, New Haven, Connecticut, USA.

51. Department of Psychiatry, Massachusetts General Hospital, Boston, Massachusetts, USA.

52. Department of Biological Psychology, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands.

53. Netherlands Twin Register, Vrije Universiteit, Amsterdam, The Netherlands.

54. Amsterdam Neuroscience, Amsterdam, The Netherlands.

55. Alzheimer Center Amsterdam, Department of Neurology, Amsterdam Neuroscience, VU Amsterdam, Amsterdam UMC, Amsterdam, The Netherlands.

56. Medical University of Lodz, Lodz, Poland.

57. Department of Integrative Medical Biology, Umeå University, Umeå, Sweden.

58. Umeå Center for Functional Brain Imaging, Umeå University, Umeå, Sweden.

59. NORMENT - KG Jebsen Centre for Psychosis Research, Division of Mental Health and Addiction, Oslo University Hospital, Oslo, Norway.

60. Department of Psychology, University of Oslo, Oslo, Norway.

61. Central Institute of Mental Health, Medical Faculty Mannheim, University Heidelberg, Mannheim, Germany.

62. Department of Psychiatry, UMC Brain Center, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands.

63. The Royal College of Surgeons in Ireland, Dublin, Ireland.

64. Department of Neurology, Yale School of Medicine, New Haven, Connecticut, USA.

65. Department of Psychiatry, Amsterdam Neuroscience, VU University Medical Center, Amsterdam, The Netherlands

66. Division of Psychiatry, Royal Edinburgh Hospital, University of Edinburgh, Edinburgh, UK.

67. Division of Systems Neuroscience of Psychopathology, Translational Research Center, University Hospital of Psychiatry, University of Bern, Bern, Switzerland.

68. Department of Medical Sciences, Molecular Medicine and Science for Life Laboratory, Uppsala University, Uppsala, Sweden.

69. Division of Psychological and Social Medicine and Developmental Neurosciences, Faculty of Medicine, TU Dresden, Dresden, Germany.

70. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Charlestown, Massachusetts, USA.

71. Department of Psychiatry, University Hospital Marqués de Valdecilla, School of Medicine, University of Cantabria-IDIVAL, Santander, Spain.

72. Department of Medicine, University Hospital Marqués de Valdecilla, School of Medicine, University of Cantabria-IDIVAL, Santander, Spain.

73. CIBERSAM (Centro Investigación Biomédica en Red Salud Mental), Santander, Spain.

74. Section for Experimental Psychopathology and Neuroimaging, Dept of General Psychiatry, Heidelberg University, Heidelberg, Germany.

75. Department of Neuroscience, Faculty of Medicine, Norwegian University of Science and Technology (NTNU), Trondheim, Norway.

76. Department of Radiology, St. Olav's Hospital, Trondheim University Hospital, Trondheim, Norway.

77. Centre for Academic Mental Health, Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK.

78. MRC Integrative Epidemiology Unit, Bristol Medical School, University of Bristol, Bristol, UK.

79. NIHR Biomedical Research Centre at the University Hospitals Bristol NHS Foundation Trust and the University of Bristol, Bristol UK.

80. Department of Radiology, University of Calgary, Calgary, Alberta, Canada.

81. Department of Clinical Neurosciences, University of Calgary, Calgary, Alberta, Canada.

82. Department of Psychiatry and Mental Health, University of Cape Town, Cape Town, South Africa.

83. South African Medical Research Council (SAMRC) Unit on Risk & Resilience in Mental Disorders, Cape Town, South Africa.

84. Laboratory of Epidemiology and Population Sciences, National Institute on Aging, Intramural Research Program, National Institutes of Health, Bethesda, Maryland, USA.

85. INSERM UMR 1000 ''Neuroimaging and Psychiatry'', University Paris-Sud, University Paris-Saclay, University Paris Descartes, Paris, France.

86. Department of Neurology, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany.

87.   Faculty of Medicine, CRC 1052 Obesity Mechanisms, University of Leipzig, Leipzig, Germany.

88.   International Max Planck Research School for Language Sciences, Nijmegen, The Netherlands.

89.   Language and Genetics Department, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands.

90.   Department of Psychiatry, University Medicine Greifswald, Greifswald, Germany.

91.   German Center for Neurodegenerative Diseases (DZNE), Rostock/Greifswald, Greifswald, Germany.

92.   Department of Medicine, University of Mississippi Medical Center, Jackson, Mississippi, USA.

93.   Department of Psychology, University of Pittsburgh, Pittsburgh, Pennsylvania, USA.

94.   Department of Neurology, University of Pittsburgh, Pittsburgh, Pennsylvania, USA.

95.   Department of Psychiatry, University of Pittsburgh, Pittsburgh, Pennsylvania, USA.

96.   NORMENT - KG Jebsen Centre for Psychosis Research, Institute of Clinical Medicine, University of Oslo, Oslo, Norway.

97.   Institute of Molecular Biology and Biochemistry, Centre for Molecular Medicine, Medical University of Graz, Graz, Austria.

98.   QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia.

99.   Department of Pharmacology, National University of Singapore, Singapore.

100.   Memory Aging and Cognition Center, National University Health System, Singapore.

101.   Institute of Neurogenomics, Helmholtz Zentrum München, German Research Centre for Environmental Health, Neuherberg, Germany.

102.   Department of Biomedical Engineering, Illinois Institute of Technology, Chicago, Illinois, USA.

103. Department of Diagnostic Radiology and Nuclear Medicine, Rush University Medical Center, Chicago, Illinois, USA.

104. Academic Unit for Psychiatry of Old Age, University of Melbourne, Victoria, Australia.

105. National Ageing Research Institute, Royal Melbourne Hospital, Melbourne, Victoria, Australia.

106. Harvard Medical School, Boston, Massachusetts, USA.

107. Lurie Center for Autism, Massachusetts General Hospital, Harvard Medical School, Lexington, Massachusetts, USA.

108. Psychiatric and Neurodevelopmental Genetics Unit, Center for Genomic Medicine, Massachusetts General Hospital, Boston, Massachusetts, USA.

109. Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard, Boston, Massachusetts, USA.

110. Amsterdam Public Health Research Institute, VU Medical Center, Amsterdam, The Netherlands.

111. Department of Psychiatry, University of Oxford, Oxford, UK.

112. NIHR Dementia Biomedical Research Unit, King's College London, London, UK.

113. NORMENT - KG Jebsen Centre for Psychosis Research, Department of Clinical Science, University of Bergen, Bergen, Norway.

114. Dr Einar Martens Research Group for Biological Psychiatry, Center for Medical Genetics and Molecular Medicine, Haukeland University Hospital, Bergen, Norway.

115. Center for integrated Sequencing, iSEQ, Aarhus University, Aarhus, Denmark.

116. Department of Biomedicine, Aarhus University, Aarhus, Denmark.

117. The Lundbeck Foundation Initiative for Integrative Psychiatric Research, iPSYCH, Aarhus and Copenhagen, Denmark.

118. Department of Clinical Neuroscience, Centre for Psychiatric Research, Karolinska Institutet, Stockholm, Sweden.

119. Stockholm Health Care Services, Stockholm County Council, Stockholm, Sweden.

120. UCL Queen Square Institute of Neurology, London, UK

121. Epilepsy Society, Bucks, UK.

122. Centre for Youth Mental Health, The University of Melbourne, Melbourne, Victoria, Australia.

123. Orygen, The National Centre of Excellence in Youth Mental Health, Melbourne, Victoria, Australia.

124. Department of Research and Development, Diakonhjemmet Hospital, Oslo, Norway.

125. Department of Psychology, University of Bath, Bath, United Kingdom.

126. Neuroimaging Unit,Technological Facilities Valdecilla Biomedical Research Institute IDIVAL, Santander, Cantabria, Spain.

127. Avera Institute for Human Genetics, Sioux Falls, South Dakota, USA.

128. Hospital for Sick Children, University of Toronto, Toronto, Ontario, Canada

129. Center for Systems Neuroscience, Boston University, Boston, Massachusetts, USA.

130. Institute of Clinical Chemistry and Laboratory Medicine, University Hospital Regensburg, Regensburg, Germany.

131. Neurodegeneratives Diseases Institute, CNRS UMR 5293, Université de Bordeaux , Bordeaux, France.

132. Pennington Biomedical Research Center, Baton Rouge, Louisiana, USA.

133. Department of Adult Psychiatry, Institute for Clinical Medicine, University of Oslo, Norway.

134. NORMENT - KG Jebsen Centre for Psychosis Research, Oslo University Hospital, Norway.

135. Brain Research Imaging Centre, University of Edinburgh, Edinburgh, UK.

136. Department of Computer Science, Lagos State University, Ojo, Nigeria.

137. Queensland Brain Institute, University of Queensland, Brisbane, Queensland, Australia.

138. Ophthalmology & Visual Sciences Academic Clinical Program (Eye ACP), Duke-NUS Medical School, Singapore.

139. Department of Behavioral Sciences, Rush University Medical Center, Chicago, Illinois, USA.

140. Centre for Healthy Brain Ageing, School of Psychiatry, University of New South Wales, Sydney, New South Wales, Australia.

141. Department of Neurology, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA.

142. Department of Radiology, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA.

143. Department of Psychology, Center for Brain Science, Harvard University, Cambridge, Massachusetts, USA.

144. Section of Gerontology and Geriatrics, Department of Medicine, University of Perugia, Perugia, Italy.

145. Center for Multimodal Imaging and Genetics, University of California, San Diego, California, USA.

146. Department of Cognitive Sciences, University of California, San Diego, California, USA.

147. Department of Neurosciences, University of California, San Diego, California, USA.

148. Department of Psychiatry, University of California, San Diego, California, USA.

149. Department of Radiology, University of California, San Diego, California, USA.

150. Division of Medical Genetics, Department of Biomedicine, University of Basel, Basel, Switzerland.

151. Institute of Human Genetics, University of Bonn, Bonn, Germany.

152.   Institute of Neuroscience and Medicine (INM-1), Research Centre Jülich, Jülich, Germany.

153.   Reta Lila Weston Institute, UCL Institute of Neurology, London, UK.

154.   Department of Molecular Neuroscience, UCL Institute of Neurology, London, UK.

155.   Department of ECE, University of New Mexico, Albuquerque, New Mexico, USA.

156.   The Mind Research Network & LBERI, Albuquerque, New Mexico, USA.

157.   Tri-institutional Center for Translational Research in Neuroimaging and Data Science (TReNDS), Georgia State University, Atlanta, Georgia, USA.

158.   Cerebral Imaging Centre, Douglas Mental Health University Institute, Montreal, Québec, Canada.

159.   Departments of Psychiatry and Biological and Biomedical Engineering, McGill University, Montreal, QC, Canada.

160.   Donders Centre for Cognitive Neuroimaging, Radboud University, Nijmegen, The Netherlands.

161.   Institute of Neuroscience and Medicine, Brain & Behaviour (INM-7), Research Centre Jülich, Jülich, Germany.

162.   Department of Clinical Neuropsychology, VU University Amsterdam, Amsterdam, The Netherlands.

163.   Department of Pediatrics, VU Medical Center, Amsterdam, The Netherlands.

164.   Emma Children's Hospital Amsterdam Medical Center, Amsterdam, The Netherlands.

165.   Univ. Lille, Labex DISTALZ - U1167 - RID-AGE - Risk factors and molecular determinants of aging-related diseases, Lille, France.

166.   Inserm, U1167, Lille, France.

167.   Centre Hospitalier Universitaire Lille, Lille, France.

168.   Institut Pasteur de Lille, Lille, France.

169. Institute of Diagnostic Radiology and Neuroradiology, University Medicine Greifswald, Greifswald, Germany.

170. Institute for Translational Genomics and Population Sciences, Los Angeles Biomedical Research Institute and Pediatrics at Harbor-UCLA Medical Center, Torrance, California, USA.

171. Institute of Biological Psychiatry, Mental Health Center Sct. Hans, Roskilde, Denmark.

172. Neurogenomics Division, The Translational Genomics Research Institute (TGEN), Phoenix, Arizona, USA.

173. Faculty of Health, Queensland University of Technology (QUT), Brisbane, Queensland, Australia.

174. Institute of Health and Biomedical Innovation, Queensland University of Technology (QUT), Brisbane, Queensland, Australia.

175. Academic Medicine Research Institute, Duke-NUS Medical School, Singapore.

176. Department of Biostatistics, Epidemiology and Informatics, University of Pennsylvania, Philadelphia, Pennsylvania, USA.

177. Dementia Centre for Research Collaboration, UNSW, Sydney, New South Wales, Australia.

178. Department of Neurology, Aristotle University of Thessaloniki, Thessaloniki, Greece.

179. Division of Mind and Brain Research, Department of Psychiatry and Psychotherapy CCM, Charité - Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, Berlin, Germany.

180. Brain Center Rudolf Magnus, Human Neurogenetics Unit, UMC Utrecht, Utrecht, The Netherlands.

181. Department of Molecular and Cellular Therapeutics, the Royal College of Surgeons in Ireland, Dublin, Ireland.

182. Department of Psychiatry, Leiden University Medical Center, Leiden, The Netherlands.

183. Leiden Institute for Brain and Cognition, Leiden University Medical Center, Leiden, The Netherlands.

184. Department of Evolution and Genetics, Dagestan State University, Makhachkala, Dagestan, Russia.

185. Department of Psychiatry, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands.

186. Karakter Child and Adolescent Psychiatry University Center, Nijmegen, The Netherlands.

187. Institute for Community Medicine, Section Epidemiology of Health Care and Community Health, University Medicine Greifswald, Greifswald, Germany.

188. NORMENT - KG Jebsen Centre for Psychosis Research, Department of Psychology, University of Oslo, Oslo, Norway.

189. Centre for Clinical Brain Sciences, University of Edinburgh, Edinburgh, UK.

190. Translational Neuroscience Laboratory, Department of Psychiatry and Human Behavior, University of California-Irvine, Irvine, California, USA.

191. Center for the Neurobiology of Learning and Memory, University of California Irvine, Irvine, California, USA.

192. University of Wuerzburg, Department of Psychiatry, Psychosomatics and Psychotherapy, Wuerzburg, Germany.

193. Brain and Mind Centre, University of Sydney, Sydney, New South Wales, Australia.

194. Neuroscience Research Australia, Sydney, New South Wales, Australia.

195. University of New South Wales, Sydney, New South Wales, Australia.

196. Department of Internal Medicine, INSERM U 558, University of Toulouse, Toulouse, France.

197. Department of Geriatric Medicine, INSERM U 558, University of Toulouse, Toulouse, France.

198. Department of Psychiatry and Psychotherapy, Charité Universitätsmedizin Berlin, CCM, Berlin, Germany.

199. Center for Neurobehavioral Genetics, University of California, Los Angeles, California, USA.

200. Neurology Division, Beaumont Hospital, Dublin, Ireland.

201. Department of Psychiatry, Carver College of Medicine, University of Iowa, Iowa City, Iowa, USA.

202. Interdepartmental Neuroscience Graduate Program, UCLA School of Medicine, Los Angeles, California, USA.

203. Imaging of Dementia and Aging (IDeA) Laboratory, Department of Neurology, University of California Davis, Davis, California, USA.

204. Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, Massachusetts, USA.

205. NORMENT, KG Jebsen Centre for Psychosis Research, Division of Mental Health and Addiction, Oslo University Hospital & Institute of Clinical Medicine, University of Oslo, Oslo, Norway.

206. School of Mental Health and Neuroscience, Faculty of Health, Medicine and Life Sciences, Maastricht University, Maastricht, The Netherlands.

207. Interfaculty Institute for Genetics and Functional Genomics, University Medicine Greifswald, Greifswald, Germany.

208. Kaiser Permanent Washington Health Research Institute, Seattle, Washington, USA

209. Department of Epidemiology, University of Washington, Seattle, Washington, USA.

210. Department of Health Services, University of Washington, Seattle, Washington, USA.

211. Institute for Molecular Bioscience, The University of Queensland, Brisbane, Queensland, Australia.

212.   Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, Indiana, USA.

213.   Department of Developmental Disability Neuropsychiatry, School of Psychiatry, UNSW Medicine, Sydney, New South Wales, Australia.

214.   Biomedical Research Unit for Dementia, King's College London, London, UK.

215.   Department of Neuroimaging, Institute of Psychiatry, King's College London, London, UK.

216.   Division of Clinical Geriatrics, Department of Neurobiology, Care Sciences and Society, Karolinska Institute, Stockholm, Sweden.

217.   Division of Genetics, Department of Medicine, Brigham and Women's Hospital, Boston, Massachusetts, USA.

218.   Section of Computational Biomedicine, Department of Medicine, Boston University School of Medicine, Boston, Massachusetts, USA.

219.   Department of Radiology, Mayo Clinic, Rochester, Minnesota, USA.

220.   School of Medical Sciences, UNSW, Sydney, New South Wales, Australia.

221.   Department of Biomedicine, University of Basel, Basel, Switzerland.

222.   C. & O. Vogt Institute for Brain Research, Heinrich Heine University Düsseldorf, Düsseldorf, Germany.

223.   Department of Psychiatry and Human Behavior, University of California-Irvine, Irvine, California, USA.

224.   Laboratory of Neuro Imaging, USC Mark and Mary Stevens Neuroimaging and Informatics Institute, Keck School of Medicine of the University of Southern California, Los Angeles, California, USA.

225.   Umeå Centre for Functional Brain Imaging (UFBI), Umeå University, Umeå 901 87, Sweden.

226.   Radiation Sciences, Umeå University, Umeå, Sweden.

227. Department of Radiology, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts, USA.

228. Department of Internal Medicine, Erasmus Medical Center, Rotterdam, The Netherlands.

229. Department of Neuroscience, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA.

230. Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA.

231. Department of Psychiatry, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA.

232. Imaging Physics, Faculty of Applied Sciences, Delft University of Technology, The Netherlands.

233. Institut Pasteur, Paris, France.

234. Department of Neurology, University of Washington, Seattle, Washington, USA.

235. Department of Neurology, Erasmus MC, Rotterdam, The Netherlands.

236. Cognitive Neuroscience Center, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands.

237. Bloorview Research Institute, Holland Bloorview Kids Rehabilitation Hospital, Toronto, Ontario, Canada.

238. Department of Psychology, University of Toronto, Toronto, Ontario, Canada.

239. Department of Psychiatry, University of Toronto, Toronto, Ontario, Canada.

240. Clinic for Cognitive Neurology, University Clinic Leipzig, Leipzig, Germany.

241. HMNC Brain Health, Munich, Germany.

242. Department of Neurology, Hopital Erasme, Universite Libre de Bruxelles, Brussels, Belgium.

243. Department of Psychology, Georgia State University, Atlanta, Georgia, USA.

244.   Department of Medical Genetics, Oslo University Hospital, Oslo, Norway.

245.   Munich Cluster for Systems Neurology (SyNergy), Munich, Germany.

246.   Institute of Translational Medicine, University of Liverpool, Liverpool, UK.

247.   Institute of Clinical Medicine, Neurology, University of Eastern Finland, Kuopio, Finland.

248.   Neurocentre Neurology, Kuopio University Hospital, Kuopio, Finland.

249.   INSERM, Research Unit 1000 "Neuroimaging and Psychiatry", Paris Saclay University, Paris Descartes University; DIGITEO Labs, Gif sur Yvette, France.

250.   Institute of Clinical Chemistry and Laboratory Medicine, University Medicine Greifswald, Greifswald, Germany.

251.   German Center for Cardiovascular Research (DZHK eV), partner site Greifswald, Greifswald, Germany.

252.   Department of Neuroinformatics, Araya, Inc., Tokyo, Japan.

253.   Institute of Cognitive Neuroscience, University College London, London, UK.

254.   School of Psychology, University of Sussex, Brighton, UK.

255.   Department of Neurobiology, Care Sciences and Society, Karolinska Institutet, Stockholm, Sweden.

256.   Department of Child and Adolescent Psychiatry/Psychology, Erasmus MC-Sophia Children's Hospital, Rotterdam, The Netherlands.

257.   Department of Clinical Genetics, Maastricht University Medical Center, Maastricht, The Netherlands.

258.   GROW School for Oncology and Developmental Biology, Maastricht, The Netherlands.

259.   Centre for Advanced Imaging, University of Queensland, Brisbane, Queensland, Australia.

260.   Department of Genomics, Life & Brain Center, University of Bonn, Germany.

261.   Neuropsychiatric Institute, Prince of Wales Hospital, Sydney, New South Wales, Australia.

262.   Department of Psychology, VU University Amsterdam, Amsterdam, The Netherlands.

263.   Center for Imaging of Neurodegenerative Disease, San Francisco VA Medical Center, University of California, San Francisco, USA.

264.   Department of Radiology and Biomedical Imaging, University of California, San Francisco, California, USA.

265.   Leiden Academic Centre for Drug Research (LACDR), Leiden University, The Netherlands.

266.   Human Genetics Center, University of Texas Health Science Center at Houston, Houston, Texas, USA.

267.   UK Dementia Research Institute, University of Edinburgh, Edinburgh, UK.

268.   Department of Neurology, Center for Neuroscience, University of California at Davis, Sacramento, California, USA.

269.   Medical and Population Genetics, Broad Institute, Cambridge, Massachusetts, USA.

270.   Center for Translational & Computational Neuroimmunology, Department of Neurology, Columbia University Medical Center, New York, New York, USA.

271.   Department of Neuroscience, Baylor College of Medicine, Houston, Texas, USA.

272.   Department of Molecular & Human Genetics, Baylor College of Medicine, Houston, Texas, USA.

273.   Jan and Dan Duncan Neurological Research Institute, Texas Children's Hospital, Houston, Texas, USA.

* CLS, HHHA, DPH, and CCW contributed equally.

** SEM, JMS, PMT, SS, and MAI jointly supervised this work.

Corresponding Authors:

Dr. Claudia L. Satizabal

Glenn Biggs Institute for Alzheimer's & Neurodegenerative Diseases

7703 Floyd Curl Drive, MSC 8070, San Antonio, TX 78229

Phone: +1 210-450-8417

Email: satizabal@uthscsa.edu

Dr. M. Arfan Ikram

Erasmus University Medical Center Rotterdam

P.O. Box 2040, 3000 CA, Rotterdam, The Netherlands

Phone: +31 10 7043930

E-mail: m.a.ikram@erasmusmc.nl

1 **Abstract**

2   Subcortical brain structures are integral to motion, consciousness, emotions, and

3 learning. We identified common genetic variation related to the volumes of nucleus

4 accumbens, amygdala, brainstem, caudate nucleus, globus pallidus, putamen, and thalamus,

5 using genome-wide association analyses in almost 40,000 individuals from CHARGE,

6 ENIGMA and the UK-Biobank. We show that variability in subcortical volumes is heritable,

7 and identify 48 significantly associated loci (40 novel at the time of analysis). Annotation of

8 these loci utilizing gene expression, methylation, and neuropathological data identified 199

9 genes putatively implicated in neurodevelopment, synaptic signaling, axonal transport,

10 apoptosis, inflammation/infection, and susceptibility to neurological disorders. This set of

11 genes is significantly enriched for *Drosophila* orthologs associated with

12 neurodevelopmental phenotypes, suggesting evolutionarily conserved mechanisms. Our

13 findings uncover novel biology and potential drug targets underlying brain development

14 and disease.

15

16   Subcortical brain structures are essential for the control of autonomic and

17   sensorimotor functions[1,2], modulation of processes involved in learning, memory, and

18   decision-making[3,4], as well as in emotional reactivity[5,6] and consciousness[7]. They often act

19   through networks influencing input to and output from the cerebral cortex[8,9]. The

20   pathology of many cognitive, psychiatric, and movement disorders is restricted to, begins

21   in, or predominantly involves subcortical brain structures and related circuitries[10]. For

22   instance, tau pathology has shown to manifest itself early in the brainstem of individuals

23   with Alzheimer's disease before spreading to cortical areas through efferent networks[11].

24   Similarly, the formation of Lewy bodies and Lewy neurites in Parkinson's disease appears

25   early in the lower brainstem (and olfactory structures) before affecting the substantia

26   nigra[12].

27   Recent investigations have identified genetic loci influencing the volumes of the

28   putamen, caudate, and pallidum, which pointed to genes controlling neurodevelopment

29   and learning, apoptosis, and transport of metals[13,14]. However, a larger study combining

30   these samples, which include individuals of a broad age-range across diverse studies,

31   would enable increased power to identify additional novel genetic variants contributing to

32   variability in subcortical structures, and further improve our understanding of brain

33   development and disease.

34   We sought to identify novel genetic variants influencing the volumes of seven subcortical

35   structures (nucleus accumbens, amygdala, caudate nucleus, putamen, globus pallidus,

36   thalamus, and brainstem – including mesencephalon, pons, and medulla oblongata),

37   through genome-wide association (GWA) analyses in almost 40,000 individuals from 53

38 study samples (Supplementary Table 1-3) from the Cohorts of Heart and Aging Research in

39 Genomic Epidemiology (CHARGE) consortium, the Enhancing Neuro Imaging Genetics

40 through Meta-Analysis (ENIGMA) consortium, and the United Kingdom Biobank (UKBB).

41

42 **RESULTS**

43

44 ***Heritability***

45 To examine the extent to which genetic variation accounts for variation in subcortical

46 brain volumes, we estimated their heritability in two family-based cohorts: the

47 Framingham Heart Study (FHS) and the Austrian Stroke Prevention Study (ASPS-Fam). Our

48 analyses are in line with previous studies conducted in twins[15], suggesting that variability

49 in subcortical volumes is moderately to highly heritable. The structures with highest

50 heritability in the FHS and the ASPS-Fam are the brainstem (ranging from 79-86%),

51 caudate nucleus (71-85%), putamen (71-79%) and nucleus accumbens (66%); followed by

52 the globus pallidus (55-60%), thalamus (47-54%), and amygdala (34-59%) (Figure 1,

53 Supplementary Table 4). We additionally estimated SNP-based heritability using GCTA in

54 the Rotterdam Study, and LD score regression (LDSC) in the full European sample. As

55 expected, SNP-based heritability estimates were somewhat lower, ranging from 47% for

56 the thalamus to 17% for the amygdala using GCTA, and ranging from 33% for the

57 brainstem to 9% for the amygdala using LDSC. These values are consistent with heritability

58 estimates reported by the UKBB[14].

59

60   ***Genome-wide associations***

61   We undertook a GWA analysis on the MRI-derived volumes of subcortical structures

62   using the 1000 Genomes Project[16] reference panel (phase 1 v.3) for imputation of missing

63   variants in CHARGE and ENIGMA. The UKBB performed imputation of variants using the

64   HRC reference panel[17] (see details on image acquisition and genotyping in Supplementary

65   Table 5 and Supplementary Table 6, respectively). Our sample comprised up to n = 37,741

66   individuals of European ancestry from 48 study samples across CHARGE, ENIGMA and the

67   UKBB. Additionally, we included three samples for generalization in African-Americans (up

68   to n = 769), and two for generalization in Asians (n = 341). Details on the population

69   characteristics, definition of the outcome and genotyping can be found in the supplement

70   (Supplementary Tables 2-5). Each study examined the association of genetic variants with

71   minor allele frequency (MAF) ≥1% to the volumes of subcortical structures (average

72   volume for bilateral structures) using additive genetic models adjusted for sex, age, total

73   intracranial volume (or total brain volume in the UKBB); as well as age^2, population

74   structure, psychiatric diagnosis (ENIGMA cohorts), and study site when applicable. After

75   quality control, we conducted meta-analyses per ethnicity combining all samples using

76   sample-size-weighted fixed effects methods in METAL[18]. An analysis of genetic correlations

77   showed consistency of associations across the CHARGE-ENIGMA and the UKBB (r_g > 0.94; P

78   < 1.46 × 10^{-15}), demonstrating the similar genetic architecture of subcortical volumes in

79   these two datasets.

80   We identified 48 independent genome-wide significant single nucleotide

81   polymorphisms (SNPs) across all seven subcortical structures, 40 of which are novel at the

82   time of analysis (Table 1). Among these, 26 SNPs were located within genes (one missense,

83   25 intronic), and 22 in intergenic regions. Most of the inflation observed in the quantile

84   plots (Supplementary Figure 1) is due to polygenic effects. We carried forward these 48

85   SNPs for *in-silico* generalization in African-American and Asian samples, and performed a

86   combined meta-analysis of all samples (Supplementary Table 7). Of the 46 SNPs present in

87   the generalization samples, the direction of association was the same for 13 across all

88   ethnicities and for an additional 6 SNPs in either the African-American or the Asian

89   samples. In the combined meta-analysis, 43 of the 48 associations remained significant, and

90   for 21 SNPs, the strength of association increased when all samples were combined.

91   Although we did not find significant associations for most SNPs at the generalization

92   sample level, likely due to their limited sample size, the sign test for the direction of effect

93   suggested that a large proportion of the SNPs associated with subcortical volumes in the

94   European sample are also associated in the African-American and Asian samples at the

95   polygenic level ($P < 1 \times 10^{-4}$; Supplementary Table 8).

96   To functionally annotate the 48 SNPs identified in the European sample, we used Locus

97   Zoom[19], investigated expression quantitative trait loci (eQTL) and methylation QTL

98   (meQTL) in post-mortem brains from the Religious Order Study and the Rush Memory and

99   Aging Project (ROSMAP), and also queried *cis-* and *trans-*eQTL datasets in brain and non-

100   brain tissues for the top 48 SNPs or their proxies ($r^2 > 0.8$), using the European population

101   reference (Supplementary Tables 9-12). Lead variants and their proxies were annotated to

102   genes based on the combination of physical proximity, eQTL and meQTL, which in some

103   instances assigned more than one gene to a single SNP. Most of our index SNPs had genes

104   assigned based on more than one functional source. This strategy allowed us to identify

105   199 putatively associated genes (Supplementary Table 13). More details can be found in

106   the Supplementary note.

107

108   ***Associations with cognition and neuropathology***

109   Although individual SNPs were not related to neuro-pathological traits or cognitive

110   function in ROSMAP (Supplementary Table 14), we found that cortical mRNA expression of

111   12 of our putatively associated genes was associated with neuropathological alterations

112   typically observed in Alzheimer's Disease (Supplementary Table 15). These included β-

113   amyloid load / presence of neuritic plaques (*APOBR, FAM65C, KTN1, NUPR1, OPA1*) and tau

114   density / neurofibrillary tangles (*FAM65C, MEPCE, OPA1, STAT1*). Many of these genes,

115   together with *ANKRD42, BCL2L1, RAET1G, SGTB*, and *ZCCHC14*, were also related to

116   cognitive function.

117

118   ***Phenotypic and genetic correlations***

119   We explored both phenotypic (Supplementary Table 16) and genetic (Supplementary

120   Table 17) correlations among subcortical volumes. We also investigated genetic

121   correlations of subcortical volumes with traits previously examined in the CHARGE and

122   ENIGMA consortia, including MRI-defined brain volumes[20,21,22], stroke subtypes[23],

123   anthropometric traits[24], general cognitive function[25], Alzheimer's disease[26], Parkinson's

124   Disease[27], bipolar disorder and schizophrenia[28], and attention deficit/hyperactivity

125   disorder (ADHD)[29]. We observed strong phenotypic and genetic overlap among most

126   subcortical structures using LDSC methods, consistent with our finding that many of the

127   loci identified have pleiotropic effects on the volumes of several subcortical structures.

128   As expected, we found strong genetic correlations among the nuclei composing the

129   striatum, particularly for nucleus accumbens with caudate nucleus ($P = 9.83 \times 10^{-19}$), and

130   with putamen ($P = 1.02 \times 10^{-17}$). The genetic architecture of thalamic volume highly

131   overlapped with that of most subcortical volumes, except for the caudate nucleus. In

132   contrast, there were no significant genetic correlations for the volume of the brainstem

133   with that of most structures, with the exception of very strong correlations with volumes of

134   the thalamus ($P = 1.56 \times 10^{-22}$) and the globus pallidus ($P = 1.52 \times 10^{-21}$). Individual level

135   analyses using GCTA in the Rotterdam Study (n = 3,486) showed similar correlations

136   despite the smaller sample.

137   We also observed strong genetic correlations for hippocampal volumes with amygdalar

138   and thalamic volumes. Height correlated with thalamic volumes and volume of the

139   brainstem was inversely correlated with ADHD. Notably, caudate nucleus volumes

140   correlated with white matter hyperintensity burden.

141

142   ***Cross-species analysis***

143   To investigate for potential evolutionarily conserved requirements of our gene-set in

144   neurodevelopment, neuronal maintenance, or both, we examined available genetic and

145   phenotypic data from the fruit fly, *Drosophila melanogaster*. Importantly, compared to

146   mammalian models, the fly genome has been more comprehensively interrogated for roles

147   in the nervous system. We found that a large proportion of candidate genes for human

148 subcortical volumes are strongly conserved in the *Drosophila* genome (59%), and many of

149 these genes appear to have conserved nervous system requirements (Supplementary Table

150 18). To examine if this degree of conservation was greater than that expected by chance,

151 we leveraged systematic, standardized phenotype data based on FlyBase annotations using

152 controlled vocabulary terms. Indeed, 22% of the conserved fly homologs are documented

153 to cause "neuroanatomy defective" phenotypes in flies, representing a significant (P = 7.3 ×

154 10^{-4}), nearly two-fold enrichment compared to 12.9% representing all *Drosophila* genes

155 associated with such phenotypes (Supplementary Table 19).

156

157 ***Partitioning heritability***

158   We further investigated enrichment for functional categories of the genome using

159 stratified LDSC methods[30] (Figure 2). Super enhancers were significantly enriched in most

160 subcortical structures, with 17% of SNPs explaining 43% of SNP-heritability in the

161 brainstem, 39% in the caudate, 44% in the pallidum, 37% in the putamen, and 38% in the

162 thalamus. Similarly, strong enrichment was observed for regular enhancers (H3K27ac

163 annotations from Hnisz[31]) in several subcortical structures, explaining over 60% of their

164 SNP-heritability. Conserved regions were enriched in the nucleus accumbens and the

165 brainstem, with 2.6% of SNPs explaining 53% and 35% of their SNP heritability,

166 respectively. Finally, only the brainstem showed enrichment for transcription start sites

167 (TSS), with 1.8% of SNPs explaining 26% of this structure SNP-heritability. Full results are

168 presented in Supplementary Table 20.

169

Protein-protein interactions

171   To explore potential functional relationships between proteins encoded by our set of genes,

172   we conducted protein-protein interaction analyses in STRING[32]. Our results showed

173   enrichment of genes involved in brain-specific pathways (i.e. regulation of neuronal death

174   and neuronal apoptosis), as well as immune-related (i.e. antigen processing, Epstein-Barr

175   virus infection) and housekeeping processes (i.e. proteasome, cell differentiation,

176   signaling). Figure 3 shows these protein networks, and the detailed pathways are

177   presented in Supplementary Table 21.

170

178   **DISCUSSION**

179   We undertook the largest GWA meta-analysis of variants associated with MRI-derived

180   volumes of the nucleus accumbens, amygdala, brainstem, caudate nucleus, globus pallidus,

181   putamen, and thalamus; in almost 40,000 individuals from 53 study samples worldwide.

182   Our analyses identified a set of 199 candidate genes influencing the volume of these

183   subcortical brain structures, most of which have relevant roles in the nervous system.

184   Our results show wide overlap of genetic variants determining the volume of

185   subcortical structures as elucidated from genetic correlations and individual look-ups

186   among structures. We find that 26 candidate genes may influence more than one structure.

187   For instance, significant SNPs near *KTN1,* are also associated with the volume of the

188   nucleus accumbens, caudate nucleus, and globus pallidus, suggesting that this genomic

189   region may have an important role in determining multiple subcortical brain volumes

190   during development. Furthermore, 14 of the candidate genes were associated with the

191   caudate, globus pallidus and putamen, supporting the shared genetic architecture of the

192   functionally defined corpus striatum.

193   We identified genes implicated in ***neurodevelopment***. We confirm the 11q14.3 genomic

194   region near the *FAT3* gene, previously associated with the caudate nucleus[13], additionally

195   associated with the putamen in our analysis. This gene encodes a conserved cellular

196   adhesion molecule implicated in neuronal morphogenesis and cell migration based on

197   mouse genetic studies[33]. SNPs near *PBX3* were associated with caudate volume. *PBX3* is

198   robustly expressed in the developing caudate nucleus of the non-human primate, *Macaca*

199   *fuscata*, consistent with a role in striatal neurogenesis[34].

200   We found several genes involved in insulin/IGF1 signaling, including *IGF1*, *PAPPA,*

201   *GRB10, SH2B1* and *TXNDC5* across the amygdala, brainstem, caudate, and putamen. *PAPPA*

202   encodes a secreted metalloproteinase that cleaves IGFBPs, thereby releasing bound IGF.

203   Although IGF may be beneficial in early- and midlife, its effects may be detrimental during

204   aging. Studies of PAPPA similarly support antagonistic pleiotropy. Low circulating PAPPA

205   levels are a marker for adverse outcomes in human embryonic development[35], but in later

206   life, higher levels have been associated with acute coronary syndromes and heart

207   failure[36,37]. Further, Grb10 and SH2B1 act as regulators of insulin/IGF1 signaling through

208   their SH2 domains[38]. Finally, *TXNDC5* has been suggested to increase IGF1 activity by

209   inhibiting the expression IGFBP1 in the context of rheumatoid arthritis[39].

210   Additional genes related to neurodevelopment include *PTPN1* (brainstem), *ALPL* and

211   *NBPF3,* (both related to the globus pallidus), and *SLC20A2* (nucleus accumbens). In studies

212   of both human and mouse embryonic stem cells, *PTPN1* was implicated as a critical

213   regulator of neural differentiation[40]. In addition, *PTPN1* encodes a target for the

214   transcriptional regulator encoded by *MECP2*, which causes the neurodevelopmental

215   disorder Rett Syndrome, and inhibition of *PTPB1* is being explored as a therapeutic

216   strategy in mouse Rett models[41]. *ALPL* mediates neuronal differentiation early during

217   development and post-natal synaptogenesis in transgenic mouse models[42]. ALPL may also

218   help propagate the neurotoxicity induced by tau[43], and its activity increases in Alzheimer's

219   disease[44] and cognitive impairment[45]. *NBPF3* belongs to the neuroblastoma breakpoint

220   family, which encodes domains of the autism- and schizophrenia-related DUF1220

221   protein[46]. *SLC20A2*, related to the globus pallidus and the thalamus, encodes an inorganic

222   phosphate transporter for which more than 40 mutations have been described in

223   association with familial idiopathic basal ganglia calcification (Fahr's Syndrome)[47,48]. It is

224   interesting to note that other three solute carrier genes were identified in this GWA

225   (*SLC12A9, SLC25A29, SLC39A8*), suggesting that the molecular transport of metals, amino

226   acids, and other solutes across the cellular membrane could play an important role in the

227   development of subcortical brain structures.

228   Several genes were related to ***synaptic signaling pathways***. We found a SNP in *NP*TX1

229   related to the thalamus, a gene expressed in the nervous system which restricts synapse

230   plasticity[49], and induces β-amyloid neurodegeneration in human and mouse brain tissues[50].

231   Additionally, the identified an intronic SNP in *SGTB* for the brainstem, which was an eQTL

232   for the expression of SGTB in dorsolateral prefrontal cortex. Experimental rat models

233   showed that βSGT, highly expressed in brain, forms a complex with the cysteine string

234   protein and heat-shock protein cognate (CSP/Hsc70) complex to function as a chaperone

235   guiding the refolding of misfolded proteins near synaptic vesicles[51]. Other experimental

236   studies in *C. elegans*, showed that the genetic manipulation of the ortholog, *sgt-1*,

237   suppresses toxicity associated with expression of the human β-amyloid peptide[52]. Other

238   genes involved in synaptic signaling are *CHPT1* (brainstem), involved in

239   phosphatidylcholine metabolism in the brain; *KATNA1*(brainstem), a conserved regulator

240   of neuronal process formation, outgrowth, and synaptogenesis[53,54]; and *DLG2* (putamen),

241   encoding an evolutionarily conserved scaffolding protein involved in glutamatergic-

242   mediated synaptic signaling and cell polarity[55] that has been associated with

243   schizophrenia[56], cognitive impairment[57], and Parkinson's disease[58].

244　　　Another set of SNPs point to genes involved in ***autophagy and apoptotic processes***,

245　　such as *DRAM1* and *FOXO3*, both related to brainstem volumes. *DRAM1* encodes a

246　　lysosomal membrane protein involved in activating TP53-mediated autophagy and

247　　apoptosis,[59] and mouse models mimicking cerebral ischemia and reperfusion have found

248　　that inhibiting the expression of *DRAM1* worsens cell injury[60]. The top SNP was also

249　　associated with a CpG site proximate to active TSS upstream of *DRAM1* in several mature

250　　brain tissues (S3.6). *FOXO3* has been recently identified as pivotal in an astrocyte network

251　　conserved across humans and mice involved in stress, sleep, and Huntington's disease[61],

252　　and has been related to longevity[62]. In *Drosophila*, a *FOXO3* ortholog regulates dendrite

253　　number and length in the peripheral nervous system[63], and in the zebrafish, *Danio rario*,

254　　*Foxo3a* knockdown led to apoptosis and mispatterning of the embryonic CNS[64]. Additional

255　　genes involved in apoptotic processes are *BCL2L1* (globus pallidus and putamen), *BIRC6*

256　　(globus pallidus) and *OPA1* (brainstem).

257　　　　Other genes have been implicated in ***axonal transport***. We confirm the association

258　　between the 13q22 locus near *KTN1* with putamen volume[13]and expand by showing that

259　　this region is also associated with the nucleus accumbens, caudate and the globus pallidus .

260　　The most significant SNP (rs945270) is a robust eQTL for *KTN1* in peripheral blood cells.

261　　This gene encodes a kinesin-binding protein involved in the transport of cellular

262　　components along microtubules[65], and impairment of these molecular motors has been

263　　increasingly recognized in neurological diseases with a subcortical component[66]. The 5q12

264　　locus upstream from *MAST4* was associated with nucleus accumbens volume. *MAST4*

265　　encodes a member of the microtubule-associated serine/threonine kinases. This gene has

266   been associated with hippocampal volumes[20] and juvenile myoclonic epilepsy[67],and it

267   appears to be differentially expressed in the prefrontal cortex of atypical cases of

268   frontotemporal lobar degeneration[68]. In *Drosophila*, the knockdown of a conserved *MAST4*

269   homolog enhanced the neurotoxicity of human tau[69], which aggregates to form

270   neurofibrillary tangle pathology in Alzheimer's disease. Further, we identified SNPs near

271   *NEFL* and *NEFM* (globus pallidus), where the top SNP was an eQTL for these genes in

272   subcortical brain tissue and esophagus mucosa. *NEFL* encodes the light chain, and NEFM

273   the medium chain of the neurofilament. These proteins determine neuronal caliber and

274   conduction velocity[70]. Mutations in NEFL/M genes have been related to neuropsychiatric

275   disorders and both proteins are increasingly recognized as powerful biomarkers of

276   neurodegeneration[71].

277   Finally, several of our candidate genes are also involved in ***inflammation, immunity***

278   ***and infection*** (*ANKRD42, DEFB124, IL27, NLRC4, PILRA/B, TRIM23, TRIM4*), in line with the

279   PPI analysis highlighting the KEGG-Epstein-Barr virus infection pathway. This suggests that

280   immune-related processes may be an important determinant influencing subcortical

281   volumes, as has been shown by other GWAS of neurologic traits[72,73].

282

283   Overall, the loci identified by our study pinpoint candidate genes not only associated

284   with human subcortical brain volumes, but also reported to disrupt invertebrate

285   neuroanatomy when manipulated in *Drosophila* and many other animal models. Thus, our

286   results are in line with the knowledge that the genomic architecture of central nervous

287   system development has been strongly conserved during evolution. Partitioning

288   heritability results suggest the nucleus accumbens and the brainstem are particularly

289   enriched in conserved regions.

290

291   One of the main limitations of our study was the small size of our generalization

292   samples, which limits the generalizability of our results to non-European ethnicities.

293   However, our analyses suggest significant concordance for the direction of effect across all

294   ethnicities at the polygenic level. We hope diverse samples become increasingly available

295   to further confirm our findings and make new discoveries. Additionally, we have focused

296   on the discovery of common and less frequent variants. Further efforts to also reveal rare

297   variants and epigenetic signatures associated with subcortical structures will provide an

298   even more refined understanding of the underlying mechanisms involved.

299

300   In conclusion, we describe multiple genes associated with the volumes of MRI-derived

301   subcortical structures in a large sample, leveraging diverse bioinformatic resources to

302   validation and follow-up our findings. Our analyses indicate that the variability of

303   evolutionarily old subcortical volumes of humans is moderately to strongly heritable, and

304   that their genetic variation is also strongly conserved across different species. The majority

305   of the variants identified in this analysis point to genes involved in neurodevelopment,

306   regulation of neuronal apoptotic processes, synaptic signaling, axonal transport,

307   inflammation/immunity, and susceptibility to neurological disorders. We show that the

308   genetic architecture of subcortical volumes overlaps with that of anthropometric measures

309   and neuropsychiatric disorders. In summary, our findings greatly expand current

310   understanding of the genetic variation related to subcortical structures, which can help

311   identify novel biological pathways of relevance to human brain development and disease.

312 **ACKNOWLEDGEMENTS**

315   **AUTHOR CONTRIBUTIONS**

316   CLS drafted the manuscript with contributions from HHHA, DPH, CCW, TVL, AAV, SE,

317   AKH, MWV, DJ, TGMVE, CDW, MJW, SEF, KAM, PJH, BF, HJG, ADJ, OLL, SDe, SEM, JMS, PMT,

318   SS, and MAI.

319   MS, NJ, LRY, TVL, GC, LA, MER, ADB, IK, MA, SA, SE, RRS, AKH, HJJ, AS, JB, MWV, AVW,

320   KW, NA, SH, ALG, PHL, SG, SLH, DK, LS, SML, IA, EW, DTG, JCI, LNV, RB, FC, DJ, OC, UKH, BSA,

321   CYC, AAA, MPB, AFM, SKM, PA, AJS, DCML, TYW, LSh, PGS, EJCdG, MT, KRVE, NJAVd, AMM,

322   JSR, NR, WH, MCVH, JBJK, LMOL, AHo, GH, MBa, SR, JJHo, ASi, NH, PRS, TWM, PMa, OGru,

323   NAG, JES, HLe, BM, DVR, IJD, RMB, IM, RK, HV, MJW, DvtE, MMN, SEF, ASB, KAM, NRS, DJH,

324   HJG, CMvD, JMW, CDe, PLDJ, and VG contributed to the preparation of data; CLS, HHHA,

325   DPH, MJK, JLS, MS, MSa, NJ, GVR, AVS, JCB, XJ, ML, EH, AT, SJvdL, JY, LRY, SL, KJY, GC, MER,

326   NJA, HJJ, AVW, SH, NMS, SG, DTG, JS, CYC, LMOL, QY, ATh, IOF, DvtE, CDe, and PLDJ

327   performed statistical analyses; and CLS, HHHA, CCW, MJK, TVL, SL, YH, KJY, JDE, QY, and

328   ADJ carried out downstream analyses.

329   SEM, JMS, PMT, SS, and MAI jointly supervised this work.

330   All authors reviewed the manuscript for intellectual content.

331 **COMPETING INTERESTS**

REFERENCES

351

352   1.   Marsden, C.D. The mysterious motor function of the basal ganglia: the Robert Wartenberg

353   Lecture. *Neurology* **32**, 514-39 (1982).

354   2.   Yin, H.H. & Knowlton, B.J. The role of the basal ganglia in habit formation. *Nat Rev Neurosci*

355   **7**, 464-76 (2006).

356   3.   McDonald, A.J. & Mott, D.D. Functional neuroanatomy of amygdalohippocampal

357   interconnections and their role in learning and memory. *J Neurosci Res* (2016).

358   4.   Hikosaka, O., Kim, H.F., Yasuda, M. & Yamamoto, S. Basal ganglia circuits for reward value-

359   guided behavior. *Annu Rev Neurosci* **37**, 289-306 (2014).

360   5.   Salzman, C.D. & Fusi, S. Emotion, cognition, and mental state representation in amygdala

361   and prefrontal cortex. *Annu Rev Neurosci* **33**, 173-202 (2010).

362   6.   Floresco, S.B. The nucleus accumbens: an interface between cognition, emotion, and action.

363   *Annu Rev Psychol* **66**, 25-52 (2015).

364   7.   Fabbro, F., Aglioti, S.M., Bergamasco, M., Clarici, A. & Panksepp, J. Evolutionary aspects of

365   self- and world consciousness in vertebrates. *Front Hum Neurosci* **9**, 157 (2015).

366   8.   Alexander, G.E., DeLong, M.R. & Strick, P.L. Parallel organization of functionally segregated

367   circuits linking basal ganglia and cortex. *Annu Rev Neurosci* **9**, 357-81 (1986).

368   9.   Jahanshahi, M., Obeso, I., Rothwell, J.C. & Obeso, J.A. A fronto-striato-subthalamic-pallidal

369   network for goal-directed and habitual inhibition. *Nat Rev Neurosci* **16**, 719-32 (2015).

370   10.   Shepherd, G.M. Corticostriatal connectivity and its role in disease. *Nat Rev Neurosci* **14**, 278-

371   91 (2013).

372   11.   Stratmann, K. *et al.* Precortical Phase of Alzheimer's Disease (AD)-Related Tau Cytoskeletal

373   Pathology. *Brain Pathol* **26**, 371-86 (2016).

374   12.   Del Tredici, K., Rub, U., De Vos, R.A., Bohl, J.R. & Braak, H. Where does Parkinson disease

375   pathology begin in the brain? *J Neuropathol Exp Neurol* **61**, 413-26 (2002).

376   13.   Hibar, D.P. *et al.* Common genetic variants influence human subcortical brain structures.

377   *Nature* **520**, 224-9 (2015).

378   14.   Elliott, L.T. *et al.* Genome-wide association studies of brain imaging phenotypes in UK

379   Biobank. *Nature* **562**, 210-216 (2018).

380   15.   Renteria, M.E. *et al.* Genetic architecture of subcortical brain regions: common and region-

381   specific genetic contributions. *Genes Brain Behav* **13**, 821-30 (2014).

382   16.   Clarke, L. *et al.* The 1000 Genomes Project: data management and community access. *Nat*

383   *Methods* **9**, 459-62 (2012).

384   17.   McCarthy, S. *et al.* A reference panel of 64,976 haplotypes for genotype imputation. *Nat*

385   *Genet* **48**, 1279-83 (2016).

386   18.   Willer, C.J., Li, Y. & Abecasis, G.R. METAL: fast and efficient meta-analysis of genomewide

387   association scans. *Bioinformatics* **26**, 2190-1 (2010).

388   19.   Pruim, R.J. *et al.* LocusZoom: regional visualization of genome-wide association scan results.

389   *Bioinformatics* **26**, 2336-7 (2010).

390   20.   Hibar, D.P. *et al.* Novel genetic loci associated with hippocampal volume. *Nat Commun* **8**,

391   13624 (2017).

392   21.   Adams, H.H. *et al.* Novel genetic loci underlying human intracranial volume identified

393   through genome-wide association. *Nat Neurosci* **19**, 1569-1582 (2016).

394   22.   Verhaaren, B.F. *et al.* Multiethnic genome-wide association study of cerebral white matter

395   hyperintensities on MRI. *Circ Cardiovasc Genet* **8**, 398-409 (2015).

396   23.   Malik, R. *et al.* Multiancestry genome-wide association study of 520,000 subjects identifies

397   32 loci associated with stroke and stroke subtypes. *Nat Genet* **50**, 524-537 (2018).

398 24. Yengo, L. *et al.* Meta-analysis of genome-wide association studies for height and body mass

399   index in approximately 700000 individuals of European ancestry. *Hum Mol Genet* **27**, 3641-

400   3649 (2018).

401 25. Davies, G. *et al.* Study of 300,486 individuals identifies 148 independent genetic loci

402   influencing general cognitive function. *Nat Commun* **9**, 2098 (2018).

403 26. Kunkle, B.W. *et al.* Genetic meta-analysis of diagnosed Alzheimer's disease identifies new

404   risk loci and implicates Abeta, tau, immunity and lipid processing. *Nat Genet* **51**, 414-430

405   (2019).

406 27. Simon-Sanchez, J. *et al.* Genome-wide association study reveals genetic risk underlying

407   Parkinson's disease. *Nat Genet* **41**, 1308-12 (2009).

408 28. Bipolar, D., Schizophrenia Working Group of the Psychiatric Genomics Consortium.

409   Electronic address, d.r.v.e., Bipolar, D. & Schizophrenia Working Group of the Psychiatric

410   Genomics, C. Genomic Dissection of Bipolar Disorder and Schizophrenia, Including 28

411   Subphenotypes. *Cell* **173**, 1705-1715 e16 (2018).

412 29. Demontis, D. *et al.* Discovery of the first genome-wide significant risk loci for attention

413   deficit/hyperactivity disorder. *Nat Genet* **51**, 63-75 (2019).

414 30. Finucane, H.K. *et al.* Partitioning heritability by functional annotation using genome-wide

415   association summary statistics. *Nat Genet* **47**, 1228-35 (2015).

416 31. Hnisz, D. *et al.* Super-enhancers in the control of cell identity and disease. *Cell* **155**, 934-47

417   (2013).

418 32. Szklarczyk, D. *et al.* STRING v10: protein-protein interaction networks, integrated over the

419   tree of life. *Nucleic Acids Res* **43**, D447-52 (2015).

420 33. Deans, M.R. *et al.* Control of neuronal morphology by the atypical cadherin Fat3. *Neuron* **71**,

421   820-32 (2011).

422 34. Takahashi, K. *et al.* Expression of FOXP2 in the developing monkey forebrain: comparison

423   with the expression of the genes FOXP1, PBX3, and MEIS2. *J Comp Neurol* **509**, 180-9

424   (2008).

425 35. Kjaer-Sorensen, K. *et al.* Pregnancy-associated plasma protein A (PAPP-A) modulates the

426   early developmental rate in zebrafish independently of its proteolytic activity. *J Biol Chem*

427   **288**, 9982-92 (2013).

428 36. Bayes-Genis, A. *et al.* Pregnancy-associated plasma protein A as a marker of acute coronary

429   syndromes. *N Engl J Med* **345**, 1022-9 (2001).

430 37. Funayama, A. *et al.* Serum pregnancy-associated plasma protein a in patients with heart

431   failure. *J Card Fail* **17**, 819-26 (2011).

432 38. Desbuquois, B., Carre, N. & Burnol, A.F. Regulation of insulin and type 1 insulin-like growth

433   factor signaling and action by the Grb10/14 and SH2B1/B2 adaptor proteins. *FEBS J* **280**,

434   794-816 (2013).

435 39. Li, J. *et al.* TXNDC5 contributes to rheumatoid arthritis by down-regulating IGFBP1

436   expression. *Clin Exp Immunol* **192**, 82-94 (2018).

437 40. Matulka, K. *et al.* PTP1B is an effector of activin signaling and regulates neural specification

438   of embryonic stem cells. *Cell Stem Cell* **13**, 706-19 (2013).

439 41. Krishnan, N. *et al.* PTP1B inhibition suggests a therapeutic strategy for Rett syndrome. *J Clin*

440   *Invest* **125**, 3163-77 (2015).

441 42. Sebastian-Serrano, A. *et al.* Tissue-nonspecific Alkaline Phosphatase Regulates Purinergic

442   Transmission in the Central Nervous System During Development and Disease. *Comput*

443   *Struct Biotechnol J* **13**, 95-100 (2015).

444 43. Diaz-Hernandez, M. *et al.* Tissue-nonspecific alkaline phosphatase promotes the

445   neurotoxicity effect of extracellular tau. *J Biol Chem* **285**, 32539-48 (2010).

446 44.   Vardy, E.R., Kellett, K.A., Cocklin, S.L. & Hooper, N.M. Alkaline phosphatase is increased in

447   both brain and plasma in Alzheimer's disease. *Neurodegener Dis* **9**, 31-7 (2012).

448 45.   Kellett, K.A., Williams, J., Vardy, E.R., Smith, A.D. & Hooper, N.M. Plasma alkaline

449   phosphatase is elevated in Alzheimer's disease and inversely correlates with cognitive

450   function. *Int J Mol Epidemiol Genet* **2**, 114-21 (2011).

451 46.   Searles Quick, V.B., Davis, J.M., Olincy, A. & Sikela, J.M. DUF1220 copy number is associated

452   with schizophrenia risk and severity: implications for understanding autism and

453   schizophrenia as related diseases. *Transl Psychiatry* **5**, e697 (2015).

454 47.   Hsu, S.C. *et al.* Mutations in SLC20A2 are a major cause of familial idiopathic basal ganglia

455   calcification. *Neurogenetics* **14**, 11-22 (2013).

456 48.   Taglia, I., Bonifati, V., Mignarri, A., Dotti, M.T. & Federico, A. Primary familial brain

457   calcification: update on molecular genetics. *Neurol Sci* **36**, 787-94 (2015).

458 49.   Figueiro-Silva, J. *et al.* Neuronal pentraxin 1 negatively regulates excitatory synapse density

459   and synaptic plasticity. *J Neurosci* **35**, 5504-21 (2015).

460 50.   Abad, M.A., Enguita, M., DeGregorio-Rocasolano, N., Ferrer, I. & Trullas, R. Neuronal

461   pentraxin 1 contributes to the neuronal damage evoked by amyloid-beta and is

462   overexpressed in dystrophic neurites in Alzheimer's brain. *J Neurosci* **26**, 12735-47 (2006).

463 51.   Tobaben, S., Varoqueaux, F., Brose, N., Stahl, B. & Meyer, G. A brain-specific isoform of small

464   glutamine-rich tetratricopeptide repeat-containing protein binds to Hsc70 and the cysteine

465   string protein. *J Biol Chem* **278**, 38376-83 (2003).

466 52.   Fonte, V. *et al.* Interaction of intracellular beta amyloid peptide with chaperone proteins.

467   *Proc Natl Acad Sci U S A* **99**, 9439-44 (2002).

468 53.   Mao, C.X. *et al.* Microtubule-severing protein Katanin regulates neuromuscular junction

469   development and dendritic elaboration in Drosophila. *Development* **141**, 1064-74 (2014).

470   54.   Yu, W. *et al.* The microtubule-severing proteins spastin and katanin participate differently in

471   the formation of axonal branches. *Mol Biol Cell* **19**, 1485-98 (2008).

472   55.   Zhu, J., Shang, Y. & Zhang, M. Mechanistic basis of MAGUK-organized complexes in synaptic

473   development and signalling. *Nat Rev Neurosci* **17**, 209-23 (2016).

474   56.   Ingason, A. *et al.* Expression analysis in a rat psychosis model identifies novel candidate

475   genes validated in a large case-control sample of schizophrenia. *Transl Psychiatry* **5**, e656

476   (2015).

477   57.   Nithianantharajah, J. *et al.* Synaptic scaffold evolution generated components of vertebrate

478   cognitive complexity. *Nat Neurosci* **16**, 16-24 (2013).

479   58.   Nalls, M.A. *et al.* Large-scale meta-analysis of genome-wide association data identifies six

480   new risk loci for Parkinson's disease. *Nat Genet* **46**, 989-93 (2014).

481   59.   Guan, J.J. *et al.* DRAM1 regulates apoptosis through increasing protein levels and lysosomal

482   localization of BAX. *Cell Death Dis* **6**, e1624 (2015).

483   60.   Yu, M., Jiang, Y., Feng, Q., Ouyang, Y. & Gan, J. DRAM1 protects neuroblastoma cells from

484   oxygen-glucose deprivation/reperfusion-induced injury via autophagy. *Int J Mol Sci* **15**,

485   19253-64 (2014).

486   61.   Scarpa, J.R. *et al.* Systems Genetic Analyses Highlight a TGFbeta-FOXO3 Dependent Striatal

487   Astrocyte Network Conserved across Species and Associated with Stress, Sleep, and

488   Huntington's Disease. *PLoS Genet* **12**, e1006137 (2016).

489   62.   Donlon, T.A. *et al.* FOXO3 longevity interactome on chromosome 6. *Aging Cell* (2017).

490   63.   Sears, J.C. & Broihier, H.T. FoxO regulates microtubule dynamics and polarity to promote

491   dendrite branching in Drosophila sensory neurons. *Dev Biol* **418**, 40-54 (2016).

492   64.   Peng, K. *et al.* Knockdown of FoxO3a induces increased neuronal apoptosis during

493   embryonic development in zebrafish. *Neurosci Lett* **484**, 98-103 (2010).

47

494 65. Santama, N., Er, C.P., Ong, L.L. & Yu, H. Distribution and functions of kinectin isoforms. *J Cell*

495   *Sci* **117**, 4537-49 (2004).

496 66. Liu, X.A., Rizzo, V. & Puthanveettil, S.V. Pathologies of Axonal Transport in

497   Neurodegenerative Diseases. *Transl Neurosci* **3**, 355-372 (2012).

498 67. Consortium, E. *et al.* Genome-wide association analysis of genetic generalized epilepsies

499   implicates susceptibility loci at 1q43, 2p16.1, 2q22.3 and 17q21.32. *Hum Mol Genet* **21**,

500   5359-72 (2012).

501 68. Martins-de-Souza, D. *et al.* Proteomic analysis identifies dysfunction in cellular transport,

502   energy, and protein metabolism in different brain regions of atypical frontotemporal lobar

503   degeneration. *J Proteome Res* **11**, 2533-43 (2012).

504 69. Shulman, J.M. *et al.* Functional screening in Drosophila identifies Alzheimer's disease

505   susceptibility genes and implicates Tau-mediated mechanisms. *Hum Mol Genet* **23**, 870-7

506   (2014).

507 70. Friede, R.L. & Samorajski, T. Axon caliber related to neurofilaments and microtubules in

508   sciatic nerve fibers of rats and mice. *Anat Rec* **167**, 379-87 (1970).

509 71. Yuan, A., Rao, M.V., Veeranna & Nixon, R.A. Neurofilaments and Neurofilament Proteins in

510   Health and Disease. *Cold Spring Harb Perspect Biol* **9**(2017).

511 72. Bis, J.C. *et al.* Whole exome sequencing study identifies novel rare and common Alzheimer's-

512   Associated variants involved in immune response and transcriptional regulation. *Mol*

513   *Psychiatry* (2018).

514 73. Marioni, R.E. *et al.* GWAS on family history of Alzheimer's disease. *Transl Psychiatry* **8**, 99

515   (2018).

516

517

518   **Figure 1. Heritability and Manhattan plot of genetic variants associated with subcortical brain volumes in the**

519   **European sample.**

520

521   **a.** Family-based heritability (h^2) estimates were performed with SOLAR in the Framingham Heart Study (n = 895) and the

522   Austrian Stroke Prevention-Family Study (n = 370). **b.** Combined Manhattan plot highlighting the most significant SNPs across

523   all subcortical structures (nucleus accumbens = 32,562; amygdala = 34,431; brainstem = 28,809; caudate = 37,741; pallidum =

524   34,413; putamen = 37,571; thalamus = 34,464). Variants are colored differently for each structure (see legend in a). Linear

525   regression models were adjusted for sex, age, age^2, total intracranial volume (CHARGE) or total brain volume (UKBB), and

526   population stratification. The solid horizontal line denotes genome-wide significance as set in this study after additional

527   Bonferroni correction for six independent traits ($P < 5 \times 10^{-8}/6 = 8.3 \times 10^{-9}$ for two-sided tests), the dashed horizontal line

528   denotes the classic genome-wide threshold of $P < 5 \times 10^{-8}$. Individual Manhattan plots can be found in the Supplementary note.

529   **Figure 2**. **Partitioning heritability by functional annotation categories.**

530

531   Analyses performed in the European sample (nucleus accumbens = 32,562; amygdala = 34,431; brainstem =28,809; caudate =

532   37,741; pallidum = 34,413; putamen = 37,571; thalamus = 34,464). Plotted ellipses represent enrichment (proportion of h^2_g

533   explained / proportion of SNPs in a given functional category) for subcortical structures (y-axis) across 28 functional

534   categories (x-axis). The color bar indicates the magnitude and direction of enrichment. Starred pairs denote significant over-

535   representation after Bonferroni correction for 168 tests (28 annotation categories and 6 independent traits, $P < 3 \times 10^{-4}$).

536   DHS, DNase I hypersensitivity site; TSS, transcription start site.

537 **Figure 3**. **Protein-protein interaction network of 158 genes enriched for common variants influencing the volume of**

538 **subcortical structures.**

539

540 The edges represent protein-protein associations, where the edge color indicates the predicted mode of action (bright green,

541 activation; pink, posttranslational modification; red, inhibition; dark blue, binding, purple, catalysis; light blue, phenotype;

542 black, reaction; yellow, transcriptional regulation) and the edge shape the predicted action effects (arrow, positive, flat arrow,

543 negative; oval arrow, unspecified). Colored nodes represent the queried proteins and first shell of interactors (5 maximum),

544 whereas white nodes represent the second shell of interactors (5 maximum).

545 **Table 1**. Genome-wide association[a] results for subcortical brain volumes in Europeans from CHARGE, ENIGMA, and the UKBB

SNP	Chr	Position	Function	A1/A2	A1 Freq.	Weight	Z-score	P[b]	Direction	I²
Nucleus accumbens (n=32,562)										
rs9818981[c]	3	190602087	intergenic	A/G	0.09	32,282	-6.23	4.70E-10	---	63.2
rs13107325	4	103188709	missense	T/C	0.06	32,283	6.15	7.74E-10	+++	76.2
rs11747514[c]	5	65839259	intronic	T/G	0.22	32,562	-5.99	2.11E-09	---	0.0
rs868202[c]	14	56195762	intergenic	T/C	0.56	32,562	5.90	3.55E-09	+++	0.0
Amygdala (n=34,431)										
rs11111293[c]	12	102921296	intergenic	T/C	0.78	34,313	6.25	4.16E-10	+++	0.0
Brainstem (n=28,809)										
rs11111090	12	102326461	intergenic	A/C	0.52	28,809	10.79	3.70E-27	+++	0.0
rs10217651[c]	9	118923652	intronic	A/G	0.39	28,809	9.78	1.40E-22	+++	0.0
rs869640[c]	5	65015128	intronic	A/C	0.72	28,809	-8.40	4.36E-17	---	9.5
rs9398173[c]	6	109000316	intronic	T/C	0.33	28,809	-7.95	1.80E-15	---	19.0
rs10792032[c]	11	68984602	intergenic	A/G	0.49	28,648	7.75	9.08E-15	+++	39.4
rs4396983[c]	4	15132604	intergenic	A/G	0.44	28,809	-7.02	2.27E-12	---	73.6
rs9322194[c]	6	149920249	intronic	T/C	0.34	28,156	6.91	4.94E-12	+++	0.0
rs7972561[c]	12	107139983	intronic	A/T	0.33	28,809	6.90	5.05E-12	+++	0.0

rs2206656[c]	20	49130119	intronic	C/G	0.61	28,809	6.83	8.26E-12	+++	0.0
rs12479469[c]	20	61145196	intergenic	A/G	0.33	25,822	-6.80	1.08E-11	---	65.6
rs4784256[c]	16	52814559	intergenic	A/G	0.40	28,809	6.76	1.41E-11	+++	0.0
rs555925[c]	3	193544359	intergenic	T/G	0.41	27,934	6.37	1.88E-10	+++	62.9
rs12313279[c]	12	102846504	intronic	A/G	0.29	28,809	6.21	5.39E-10	+++	24.9
rs9505301[c]	6	7887131	intronic	A/G	0.89	28,691	-6.05	1.41E-09	---	43.2
rs11684404[c]	2	88924622	intronic	T/C	0.66	28,809	-5.95	2.73E-09	---	0.0
rs112178027[c]	17	27564013	intergenic	T/C	0.17	28,809	-5.90	3.67E-09	---	0.0

Caudate nucleus (n=37,741)

rs3133370	11	92026446	intergenic	T/C	0.67	37,741	7.52	5.59E-14	+++	44.9
rs6060983[c]	20	30420924	intronic	T/C	0.70	37,741	7.04	1.95E-12	+++	0.0
rs7040561[c]	9	128528978	intronic	A/T	0.85	34,049	-6.26	3.84E-10	---	0.0
rs2817145[c]	1	3133422	intronic	A/T	0.19	35,598	6.20	5.71E-10	+++	65.3
rs148470213[c]	14	56193700	intergenic	T/C	0.54	29,429	6.18	6.48E-10	++?	0.0
rs1987471[c]	16	28825866	intergenic	T/G	0.63	37,741	5.87	4.40E-09	+++	0.0
rs12445022[c]	16	87575332	intergenic	A/G	0.33	37,741	5.87	4.45E-09	+++	0.0
rs55989340[c]	14	100635222	intergenic	A/G	0.74	37,741	-5.86	4.62E-09	---	52.0
rs4888010[c]	16	73895046	intergenic	A/G	0.47	37,741	5.86	4.67E-09	+++	74.9

rs35305377[c]	7	99938955	intronic	A/G	0.55	33,429	-5.84	5.36E-09	---	47.8
Globus pallidus (n=34,413)										
rs2923447	8	42439848	intergenic	T/G	0.59	34,413	8.11	4.88E-16	+++	34.0
rs10129414[c]	14	56193272	intergenic	A/G	0.44	34,413	-7.53	5.11E-14	---	0.0
rs196807[c]	8	24682649	intergenic	A/G	0.18	34,295	6.44	1.17E-10	+++	21.1
rs10439607[c]	20	30258541	intronic	A/G	0.30	34,413	-6.28	3.35E-10	---	0.0
rs4952211[c]	2	32611512	intronic	T/C	0.43	34,252	-5.86	4.72E-09	---	61.9
rs12567402[c]	1	21870213	intronic	T/C	0.33	34,214	5.81	6.17E-09	+++	0.0
Putamen (n=37,571)										
rs945270	14	56200473	intergenic	C/G	0.58	37,571	15.03	5.02E-51	+++	57.3
rs62098013	18	50863861	intronic	A/G	0.38	37,571	8.92	4.59E-19	+++	33.9
rs6087771	20	30306724	intronic	T/C	0.71	36,291	8.69	3.75E-18	+++	7.5
rs35200015[c]	11	117383215	intronic	A/G	0.19	37,571	-8.19	2.51E-16	---	0.0
rs1432054	11	83260225	intronic	A/G	0.64	37,571	-7.94	2.10E-15	---	0.0
rs7902527[c]	10	118715399	intronic	A/G	0.24	37,108	6.29	3.13E-10	+++	0.0
rs2244479[c]	7	50738987	intronic	T/C	0.65	36,291	-5.92	3.17E-09	---	32.1
rs2410767[c]	5	87705268	intronic	C/G	0.78	37,571	5.88	3.99E-09	+++	0.0
rs1187162[c]	11	92011126	intergenic	T/C	0.42	37,571	5.84	5.14E-09	+++	0.0

Thalamus (n=34,464)									
rs12600720[c]	17	78448640	intronic	C/G	0.69	33,023	6.25	4.06E-10	+++	0.0
rs142461330[c]	7	55012097	intergenic	T/C	0.92	34,185	-5.90	3.69E-09	---	0.0

546 [a] Linear regression models are adjusted for sex, age, age^2, total intracranial volume (CHARGE) or total brain volume (UKBB),

547 and population stratification.

548 [b] P-values are two-tailed. Significance was set at $P < 8.3 \times 10^{-9}$ after additional Bonferroni correction for six independent traits

549 $(5 \times 10^{-8}/6)$.

550 [c] Novel SNPs

551 Chr = chromosome; Freq. = frequency of the coded allele; A1 = coded allele; A2 = non-coded allele

552   **ONLINE METHODS**

553

554   ***Study population***

555   The present effort included 53 study samples from the Cohorts of Heart and Aging

556   Research in Genomic Epidemiology (CHARGE) consortium [74], the Enhancing Neuro Imaging

557   Genetics through Meta-Analysis (ENIGMA) consortium[75], and the United Kingdom Biobank

558   (UKBB)[76]. Briefly, the CHARGE consortium is a collaboration of predominantly population-

559   based cohort studies investigating the genomics of age-related complex diseases, including

560   those of the brain (depts.washington.edu/chargeco/wiki/). The ENIGMA consortium brings

561   together various studies, approximately 75% of which are population-based, with the

562   remainder using case-control designs for various neuropsychiatric or neurodegenerative

563   diseases (enigma.ini.usc.edu/). The UKBB is a large-scale prospective epidemiological

564   study of over 500,000 individuals aged 40-69 years from the United Kingdom, established

565   to investigate the genetic and non-genetic determinants of middle and old age diseases

566   (www.ukbiobank.ac.uk/).

567   Our sample consisted of up to n=37,741 individuals of European ancestry. We

568   additionally included three generalization samples of African-Americans (up to n=769),

569   and two generalization samples of Asians (n=341). All participants have provided written

570   informed consent and participating studies obtained approval from their institutional

571   review board or equivalent organization. The institutional review boards of Boston

572   University and the University of Southern California, as well as the local ethics board of

573   Erasmus University Medical Center approved this study.

574   Exclusion criteria comprised prevalent dementia or stroke at the time of the MRI scan,

575   and when available, presence of large brain infarcts or other neurological pathologies seen

576   at the MRI that could substantially influence the measurement of brain volumes (e.g. brain

577   tumor, trauma). Individual studies applied the exclusion criteria prior to analyses.

578

579   ***Definition of phenotypes***

580   Our study investigated the volumes of seven subcortical structures: nucleus accumbens,

581   amygdala, brainstem, caudate nucleus, globus pallidus, putamen, and thalamus. These

582   phenotypes were defined as the mean volume (in cm^3) of the left and right hemispheres,

583   with the exception if the brainstem that was simply defined as total volume (in cm^3). Each

584   study contributed magnetic resonance imaging (MRI) data obtained using diverse

585   scanners, field strengths, and acquisition protocols. The estimation of volumes for the

586   seven subcortical brain structures and total intracranial volume was generated by freely

587   available and in-house segmentation methods previously described and validated.

588   Summary statistics for subcortical brain volumes in CHARGE study samples are presented

589   in Supplementary Table 3, and the study-specific MRI protocols and software are described

590   in Supplementary Table 5. We have recently published results describing the genetic

591   variation associated with hippocampal volumes[20], and therefore, we have not included that

592   brain structure in this report.

593

594   ***Genotyping***

595   Genotyping was performed using a variety of commercial arrays across the

596   participating studies. Study samples and genetic variants underwent similar quality control

597   procedures based on genetic homogeneity, call rate, minor allele frequency (MAF), and

598   Hardy-Weinberg Equilibrium. Good quality variants were used as input for imputation to

599   the 1000 Genomes Project (phase 1, version 3) reference panel[16], or the Haplotype

600   Reference Consortium (HRC, version 1.1)[17] in the UKBB, using validated software packages.

601   A detailed description of the genotyping and quality control carried by each study is

602   described in Supplementary Table 6.

603

604   ***Heritability***

605   Heritability of subcortical brain volumes was estimated in the Framingham Heart Study

606   (FHS)[77] and the Austrian Stroke Prevention Study Family Study (ASPS Fam)[78], two

607   population-based cohorts with family structure. We used SOLAR[79] to determine the ratio of

608   the genetic variance to the phenotypic variance, including variance component models that

609   were adjusted for age, sex, total intracranial volume, as well as age squared and principal

610   components if required, in the same way it is described for the genome-wide association

611   (GWA) analysis. We also estimated the variance of subcortical structures explained by SNPs

612   in a sample of n=3,486 unrelated participants from the Rotterdam Study using GCTA[80], and

613   additionally in the full European sample using LDSC regression methods[81]. Supplementary

614   Table 4 provides family- and SNP-based heritabilities for subcortical structures.

615

616   ***Genome-wide associations and meta-analysis***

617   In CHARGE and ENIGMA, each study undertook a GWA analysis on the volumes of seven

618   MRI subcortical brain structures (or those that were available to each study) according to a

619   common predefined analysis plan. Studies including unrelated participants performed

620   linear regression analyses, whereas those including related participants conducted linear

621   mixed models to account for familial relationships. Models assumed additive genetic effects

622   and were adjusted for age, sex, total intracranial volume and, if applicable, they were

623   additionally adjusted for age^2, principal components to account for population

624   stratification, psychiatric diagnosis (ENIGMA cohorts), and study site. Individual studies

625   shared summary statistics to a centralized, secured computing space. Analysis in the UKBB

626   sample followed a similar approach in n=8,312 unrelated participants although the genetic

627   data used for these analyses uses only those variants imputed using the HRC[17] reference

628   panel. As the data released by the UKBB did not include total intracranial volume, linear

629   regression models in this sample are adjusted for age, age^2, sex, *total brain volume*, and

630   principal components. We used LDSC methods[81] to investigate the genetic correlations for

631   all subcortical structures between the CHARGE-ENIGMA and the UKBB. There was no

632   evidence suggesting differences in the genetic architecture of both samples.

633   Prior to meta-analysis, we performed quality control at the study-level summary

634   statistics using a series of quality checks implemented in EasyQC[82]. Filters were set to

635   remove SNPs with poor imputation ($R^2 < 0.5$), rare (MAF < 0.1%), or with an effective allele

636   count (2 x MAF x study sample size x imputation quality) < 20. Finally, we only considered

637   variants present in at least 70% of the total European sample for each structure.

638   Fixed-effects meta-analyses weighting for sample size were performed using METAL[18],

639   given that not all samples used the same methods for acquisition and post-processing of

640   brain images. We used the LD score regression intercept to correct for population

641   stratification and cryptic relatedness[81]. Quantile and Manhattan plots are presented for

642   each subcortical structure in Supplementary Figure 1. To correct for multiple comparisons

643   across our seven traits, we calculated the Pearson's correlation among subcortical

644   structures adjusting for age, sex and intracranial volume in n=4,459 participants from the

645   Rotterdam Study. After 1,000 permutations, the resulting number of independent traits

646   was of six, leading to the definition of a significant threshold as $P < (5 \times 10^{-8}/6) = 8.3 \times 10^{-9}$.

647   To select our top independent SNPs in the European meta-analysis, we ran a multi-SNP-

648   based conditional & joint association analysis (GCTA-COJO)[80] using n=6,921 participants

649   from the Rotterdam Study as the reference sample. In secondary analyses, we looked for

650   the association of our index SNPs (the most significant variant in each locus) with the other

651   six subcortical structures.

652   We conducted separate meta-analyses by ancestry, and further performed a combined

653   meta-analysis including all samples. Forest plots were created to explore the contribution

654   of participating studies to each of the significant SNPs (Supplementary Figure 4). To assess

655   signal overlap with African-American and Asian samples, we first clumped variants with P

656    $< 1 \times 10^{-4}$ in the European sample, and then ran binomial sign tests for the correlation of

657   the direction of association across ethnic groups.

658

659   ***Functional annotations***

660   We used Locus Zoom[19] based on the hg 19 UCSC Genome Browser assembly for the

661   visualization of the nearest genes within a ±500 Kb genomic region. We also investigated

662   *cis* (1 Mb) expression quantitative trait loci (eQTL) and methylation QTL (meQTL) for our

663   index SNPs in post-mortem brains from the Religious Order Study and the Rush Memory

664   and Aging Project (ROSMAP). In ROSMAP, the dorsolateral prefrontal cortex (DLPFC) was

665   selected for initial multi-omics data generation, as it is relevant to multiple common

666   neuropathologies and cognitive phenotypes in the aging population[83]. RNA was extracted

667   from the gray matter of DLPFC, and next-generation RNA sequencing (RNA-Seq) was done

668   on the Illumina HiSeq for samples with an RNA integrity score > 5 and a quantity threshold

669   > 5 ug, as previously described[83,84]. We quantile-normalized the fragments per kilobase of

670   transcript per million fragments mapped (FPKM), correcting for batch effect with

671   Combat[84,85]. These adjusted FPKM values were used for analysis. A subset of 407

672   participants had quality-controlled RNA-Seq data and were included in the eQTL analysis.

673   DNA methylation levels from the gray matter of DLPFC were measured using the

674   Illumina HumanMethylation450 BeadChip, and the measurements underwent QC

675   processing as previously described (i.e. detection $p < 0.01$ for all samples)[83], yielding 708

676   participants with 415,848 discrete CpG dinucleotide sites with methylation measurement.

677   Any missing methylation levels from any of quality-controlled CpG dinucleotide sites were

678   imputed using a k-nearest neighbor algorithm for k = 100[83]. A subset of 488 participants in

679   our study had quality-controlled genome-wide methylation data and were included in the

680   *cis*-methylation QTL analysis. Finally, the associations between our index SNPs and CpG

681   sites were plotted along Roadmap Epigenomic chromatin states for ten brain tissues[86].

682   We further queried *cis* and *trans* eQTLs in non-brain and brain tissues from additional

683   eQTL repositories[87]. We searched for proxies to our index SNPs with a $r^2 > 0.8$ using the

684   European population reference in rAggr (1000G, phase 1, Mar 2012), and then queried

685   index and proxy SNPs against eQTLs from diverse databases.[88] Blood cell related eQTL

686   studies included fresh lymphocytes and leukocytes, leukocyte samples in individuals with

687   Celiac disease, whole blood samples, lymphoblastoid cell lines (LCL) derived from

688   asthmatic children, HapMap LCL from 3 populations, a separate study on HapMap CEU LCL,

689   LCL population samples, neutrophils, CD19+ B cells, primary PHA-stimulated T cells, CD4+

690   T cells, peripheral blood monocytes, long non-coding RNAs in monocytes and CD14+

691   monocytes before and after stimulation with LPS or interferon-gamma, CD11+ dendritic

692   cells before and after *Mycobacterium tuberculosis* infection and a separate study of

693   dendritic cells before or after stimulation with lipopolysaccharide (LPS), influenza or

694   interferon-beta; micro-RNA QTLs, DNase-I QTLs, histone acetylation QTLs, and ribosomal

695   occupancy QTLs were also queried for LCL; splicing QTLs and micro-RNA QTLs were

696   queried in whole blood. Non-blood cell tissue eQTL searches included omental and

697   subcutaneous adipose, visceral fat stomach, endometrial carcinomas, ER+ and ER- breast

698   cancer tumor cells, liver, osteoblasts, intestine and normal and cancerous colon, skeletal

699   muscle, breast tissue (normal and cancer), lung, skin, primary fibroblasts, sputum,

700   pancreatic islet cells, prostate, rectal mucosa, arterial wall and heart tissue from left

701   ventricles and left and right atria. Micro-RNA QTLs were also queried for gluteal and

702   abdominal adipose and liver. Methylation QTLs were queried in pancreatic islet cells.

703   Further mRNA and micro-RNA QTLs were queried from ER+ invasive breast cancer

704   samples, colon-, kidney renal clear-, lung- and prostate-adenocarcinoma samples. Brain

705   eQTL studies included brain cortex, cerebellar cortex, cerebellum, frontal cortex, gliomas,

706   hippocampus, inferior olivary nucleus (from medulla), intralobular white matter, occiptal

707   cortex, parietal lobe, pons, pre-frontal cortex, putamen (at the level of anterior

708   commussure), substantia nigra, temporal cortex, thalamus and visual cortex. eQTL data

709   was integrated from online sources including ScanDB[89], the GTEx Portal[90], and the

710   Pritchard Lab[91]. Cerebellum, parietal lobe and liver eQTL data was downloaded from

711   ScanDB and cis-eQTL were limited to those with $P<1.0 \times 10^{-6}$ and trans-eQTLs with $P < 5.0$

712   × 10[-8]. Results for GTEx Analysis V6 for 48 tissues were downloaded from the GTEx Portal

713   (www.gtexportal.org). For all gene-level eQTL, if at least 1 SNP passed the tissue-specific

714   empirical threshold in GTEx, the best SNP for that eQTL was always retained.

715

716   ***Associations of cognition and neuropathology phenotypes with gene expression in***

717   ***brain***

718   We further related cognitive function and neuropathological findings to the expression

719   of the 199 gene set influencing subcortical volumes in 508 brains from the ROSMAP

720   samples.

721   Briefly, brain autopsies were performed as previously described and each brain was

722   inspected for common pathologies relating to loss of cognition in aging populations[92,93]. In

723   this report, we included: neurofibrillary tangles, neuritic plaques, β-amyloid load, tau

724   density, hippocampal sclerosis, Lewy bodies and neuronal loss in substantia nigra.

725   Neurofibrillary tangles and neuritic plaques were visualized by modified Bielschowsky

726   silver stain, then counted and scaled in five brain regions: mid-frontal, temporal, inferior

727   parietal, entorhinal cortex, and hippocampus CA1. Composite scores for each of these three

728   pathology types were derived by scaling the counts within each of the five regions, and

729   taking the square root of the average of the regional scaled values to account for their

730   positively skewed distribution[92-94]. β-amyloid load and tau tangle density were measured

731   by immunohistochemistry and square root transformed as previously described[95]. Lewy

732   bodies were identified using immunohistochemistry and were further dichotomized as

733   present or absent based on the recommendations of the Report of the Consortium on DLB

734   International Workshop[96]. Hippocampal sclerosis was recorded as either present or absent

735   as evaluated with H&E stain. Nigral neuronal loss was assessed in the substantia nigra in

736   the mid to rostral midbrain near or at the exit of the 3rd nerve using H&E stain and 6

737   micron sections using a semi-quantitative scale (0–3)[97].

738   Global cognition was computed as a composite score of 19 (ROS) and 17 (MAP)

739   cognitive tests performed at annual evaluations including five cognitive domains: episodic

740   memory, semantic memory, working memory, perceptual speed, and visuospatial

741   ability[92,93]. From these scores, we created normalized summary measures to limit the

742   influence of outliers. We used global cognition proximate to death to derive cognitive

743   reserve. Separately, the residual slope of global cognitive change and the residual slopes of

744   cognitive change in the five cognitive domains were derived through general linear mixed

745   models, controlling for age at enrollment, sex, and education.

746

747   ***Phenotypic and genetic correlations***

748   We estimated the Pearson's partial phenotypic correlations among the volumes of

749   subcortical structures in 894 participants from the Framingham Heart Study. Similarly, to

750   the GWA, these analyses were corrected for the effects of sex, age, age^2, total intracranial

751   volume and PC1.

752   Genetic correlation analyses were performed using LDSC regression methods[81]. The

753   GWA meta-analysis results for the seven subcortical brain structures were correlated with

754   each other's, as well as with published GWA studies on the following traits: hippocampal

755   volume[20], intracranial volume[21], white matter hyperintensities[22], stroke subtypes[23], adult

756   height and body mass index[24], fat-free mass and whole-body water mass[98], Alzheimer's

757   disease[26], Parkinson's Disease[27], general cognitive function[25], bipolar disorder and

758   schizophrenia[28], and ADHD[29].

759

760   ***Look-up of functional orthologs in Drosophila melanogaster***

761   For the cross-species assessment of gene-phenotype relationships in *Drosophila*, we

762   relied on a similar analytic approach as in prior work[99]. Human genes were mapped to

763   corresponding *Drosophila* orthologs using DIOPT: Drosophila Integrated Ortholog

764   Prediction Tool (www.flyrnai.org/diopt)[100], which incorporates 14 distinct algorithms to

765   define orthology. Fly gene orthologs were defined based on a DIOPT score of 2 or greater,

766   indicating at least 2 algorithms were in agreement on the pairing. When more than one of

767   the fly ortholog was predicted, all such genes meeting this threshold were included in our

768   analyses. This resulted in a gene set consisting of 168 *Drosophila* homologs of human

769   candidate genes at subcortical volume susceptibility loci. The resulting 37 genes associated

770   with "neuroanatomy defective" phenotypes in *Drosophila* (22%) were annotated based on

771   the controlled vocabulary terms implemented in FlyBase (flybase.org/)[101]. Genes causing

772   "neuroanatomy defective" phenotypes in *Drosophila* include both loss- or gain-of-function

773   genetic manipulations of fly gene homologs. Loss-of-function studies included both

774   classical mutant alleles (e.g. point mutations, gene deletions, or transposon insertions) or

775   gene knockdown using RNA interference transgenic strains. Gain-of-function experiments

776   were based on tissue specific overexpression of the fly gene orthologs. The hypergeometric

777   overlap test was used to assess for enrichment of "neuroanatomy defective" phenotypes

778   among the conserved gene set.

779

780 ***Protein-protein interactions and network analysis***

781   We used the human STRING database resource (string-db.org)[32] for the exploration of

782 direct (physical) and indirect (functional) protein-protein interactions based on the gene

783 set derived from the GWA results and functional annotations (Supplementary Table 13).

784 The input parameters included a medium-confidence interaction scores (0.4) with first and

785 second shells of maximum 5 interactors. Finally, we generated a protein-protein

786 interaction network based on known and predicted interactions.

787

788 ***Partitioning heritability***

789   Partitioned heritability was estimated with stratified LDSC methods[30]. This method

790 partitions SNP heritability using GWAS summary results and accounting by LD. We used

791 the meta-analysis results from the European sample to partitioning SNPs by 28 functional

792 categories, including: coding, intron, promoter, 3'/5' UTRs, digital genomic footprint (DGF),

793 transcription factor binding sites, chromHMM and Segway annotations for six cell lines,

794 DNase I hypersensitivity sites (DHS), H3K4me1, H3K4me3 and H3K9ac marks, two sets of

795 H3K27ac marks, super-enhancers, conserved regions in mammals, and FANTOM5

796 enhancers. Significance was set at $P < (0.05/(28 \times 6)) = 3 \times 10^{-4}$.

797

798 ***Data availability***

799 The genome-wide summary statistics that support the findings of this study will be made

800 available through the CHARGE dbGaP (accession number phs000930) and ENIGMA

801 (http://enigma.ini.usc.edu/research/download-enigma-gwas-results) websites.

802

803 **METHODS REFERENCES**

804 74. Psaty, B.M. *et al.* Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE)

805   Consortium: Design of prospective meta-analyses of genome-wide association studies from

806   5 cohorts. *Circ Cardiovasc Genet* **2**, 73-80 (2009).

807 75. Thompson, P.M. *et al.* The ENIGMA Consortium: large-scale collaborative analyses of

808   neuroimaging and genetic data. *Brain Imaging Behav* **8**, 153-82 (2014).

809 76. Sudlow, C. *et al.* UK biobank: an open access resource for identifying the causes of a wide

810   range of complex diseases of middle and old age. *PLoS Med* **12**, e1001779 (2015).

811 77. Tsao, C.W. & Vasan, R.S. Cohort Profile: The Framingham Heart Study (FHS): overview of

812   milestones in cardiovascular epidemiology. *Int J Epidemiol* **44**, 1800-13 (2015).

813 78. Schmidt, R. *et al.* Assessment of cerebrovascular risk profiles in healthy persons: definition

814   of research goals and the Austrian Stroke Prevention Study (ASPS). *Neuroepidemiology* **13**,

815   308-13 (1994).

816 79. Almasy, L. & Blangero, J. Multipoint quantitative-trait linkage analysis in general pedigrees.

817   *Am J Hum Genet* **62**, 1198-211 (1998).

818 80. Yang, J., Lee, S.H., Goddard, M.E. & Visscher, P.M. GCTA: a tool for genome-wide complex trait

819   analysis. *Am J Hum Genet* **88**, 76-82 (2011).

820 81. Bulik-Sullivan, B.K. *et al.* LD Score regression distinguishes confounding from polygenicity

821   in genome-wide association studies. *Nat Genet* **47**, 291-5 (2015).

822 82. Winkler, T.W. *et al.* Quality control and conduct of genome-wide association meta-analyses.

823   *Nat Protoc* **9**, 1192-212 (2014).

824 83. Bennett, D.A., Yu, L. & De Jager, P.L. Building a pipeline to discover and validate novel

825   therapeutic targets and lead compounds for Alzheimer's disease. *Biochem Pharmacol* **88**,

826   617-30 (2014).

827 84. Chan, G. *et al.* CD33 modulates TREM2: convergence of Alzheimer loci. *Nat Neurosci* **18**,

828   1556-8 (2015).

829 85. Johnson, W.E., Li, C. & Rabinovic, A. Adjusting batch effects in microarray expression data

830   using empirical Bayes methods. *Biostatistics* **8**, 118-27 (2007).

831 86. Roadmap Epigenomics, C. *et al.* Integrative analysis of 111 reference human epigenomes.

832   *Nature* **518**, 317-30 (2015).

833 87. Eicher, J.D. *et al.* GRASP v2.0: an update on the Genome-Wide Repository of Associations

834   between SNPs and phenotypes. *Nucleic Acids Res* **43**, D799-804 (2015).

835 88. Zhang, X. *et al.* Synthesis of 53 tissue and cell line expression QTL datasets reveals master

836   eQTLs. *BMC Genomics* **15**, 532 (2014).

837 89. Zhang, W. *et al.* SCAN database: facilitating integrative analyses of cytosine modification and

838   expression QTL. *Database (Oxford)* **2015**(2015).

839 90. Consortium, G.T. The Genotype-Tissue Expression (GTEx) project. *Nat Genet* **45**, 580-5

840   (2013).

841 91. Veyrieras, J.B. *et al.* High-resolution mapping of expression-QTLs yields insight into human

842   gene regulation. *PLoS Genet* **4**, e1000214 (2008).

843 92. Bennett, D.A. *et al.* Overview and findings from the rush Memory and Aging Project. *Curr*

844   *Alzheimer Res* **9**, 646-63 (2012).

845 93. Bennett, D.A., Schneider, J.A., Arvanitakis, Z. & Wilson, R.S. Overview and findings from the

846   religious orders study. *Curr Alzheimer Res* **9**, 628-45 (2012).

847 94. Replogle, J.M. *et al.* A TREM1 variant alters the accumulation of Alzheimer-related amyloid

848   pathology. *Ann Neurol* **77**, 469-77 (2015).

849 95. Barnes, L.L., Schneider, J.A., Boyle, P.A., Bienias, J.L. & Bennett, D.A. Memory complaints are

850   related to Alzheimer disease pathology in older persons. *Neurology* **67**, 1581-5 (2006).

851   96.   McKeith, I.G. *et al.* Consensus guidelines for the clinical and pathologic diagnosis of

852   dementia with Lewy bodies (DLB): report of the consortium on DLB international

853   workshop. *Neurology* **47**, 1113-24 (1996).

854   97.   Schneider, J.A. *et al.* Substantia nigra tangles are related to gait impairment in older persons.

855   *Ann Neurol* **59**, 166-73 (2006).

856   98.   Zheng, J. *et al.* LD Hub: a centralized database and web interface to perform LD score

857   regression that maximizes the potential of summary level GWAS data for SNP heritability

858   and genetic correlation analysis. *Bioinformatics* **33**, 272-279 (2017).

859   99.   Wangler, M.F., Hu, Y. & Shulman, J.M. Drosophila and genome-wide association studies: a

860   review and resource for the functional dissection of human complex traits. *Dis Model Mech*

861   **10**, 77-88 (2017).

862   100.   Hu, Y. *et al.* An integrative approach to ortholog prediction for disease-focused and other

863   functional studies. *BMC Bioinformatics* **12**, 357 (2011).

864   101.   Marygold, S.J., Crosby, M.A., Goodman, J.L. & FlyBase, C. Using FlyBase, a Database of

865   Drosophila Genes and Genomes. *Methods Mol Biol* **1478**, 1-31 (2016).

866