Genome-wide analyses identify 30 loci associated with obsessive-compulsive disorder

Nora I. Strom[1,2,3,4,#], Zachary F. Gerring[5,6,#], Marco Galimberti[7,8,#], Dongmei Yu[9,10,#], Matthew W. Halvorsen[11], Abdel Abdellaoui[12], Cristina Rodriguez-Fontenla[13,14], Julia M. Sealock[15], Tim Bigdeli[16,17], Jonathan R. Coleman[18,19], Behrang Mahjani[20,21], Jackson G. Thorp[22,23], Katharina Bey[24], Christie L. Burton[25], Jurjen J. Luykx[26,27], Gwyneth Zai[28,29], Silvia Alemany[30,31,32], Christine Andre[33], Kathleen D. Askland[34], Julia Bäckmann[3], Nerisa Banaj[35], Cristina Barlassina[36], Judith Becker Nissen[37,38], O. Joseph Bienvenu[39], Donald Black[40], Michael H. Bloch[41], Sigrid Børte[42,43,44], Rosa Bosch[45,46], Michael Breen[47,48,49], Brian P. Brennan[33,50], Helena Brentani[51], Joseph D. Buxbaum[20], Jonas Bybjerg-Grauholm[52], Enda M. Byrne[53], Judit Cabana-Dominguez[30,31,32], Beatriz Camarena[54], Adrian Camarena[55], Carolina Cappi[56,57], Angel Carracedo[58,59,60], Miguel Casas[61,62], Maria Cristina Cavallini[63], Valentina Ciullo[35], Edwin H. Cook[64], Jesse Crosby[33,50], Bernadette A. Cullen[65,66], Elles J. De Schipper[3], Richard Delorme[67], Srdjan Djurovic[44,68], Jason A. Elias[69,70], Xavier Estivill[71], Martha J. Falkenstein[33,50], Bengt T. Fundin[72], Lauryn Garner[33], Christina Gironda[33], Fernando S. Goes[73], Marco A. Grados[74], Jakob Grove[4,75,76,77], Wei Guo[78], Jan Haavik[79,80], Kristen Hagen[81,82,83], Kelly Harrington[84,85], Alexandra Havdahl[86,87], Kira D. Höffler[68,88,89], Ana G. Hounie[90], Donald Hucks[91], Christina Hultman[21], Magdalena Janecka[92,93], Eric Jenike[33], Elinor K. Karlsson[94,95], Kara Kelley[33], Julia Klawohn[1,96], Janice E. Krasnow[97], Kristi Krebs[98], Christoph Lange[99,100], Nuria Lanzagorta[101], Daniel Levey[102,103], Kerstin Lindblad-Toh[95,104,105], Fabio Macciardi[106], Brion Maher[107], Brittany Mathes[33], Evonne McArthur[108], Nathaniel McGregor[109], Nicole C. McLaughlin[110,111], Sandra Meier[112], Euripedes C. Miguel[113], Maureen Mulhern[24], Paul S. Nestadt[114], Erika L. Nurmi[115], Kevin S. O'Connell[116,117], Lisa Osiecki[118,119], Olga Therese Ousdal[79,120], Teemu Palviainen[121], Nancy L. Pedersen[21], Fabrizio Piras[35], Federica Piras[122], Sriramya Potluri[33], Raquel Rabionet[123,124,125], Alfredo Ramirez[126,127,128,129,130], Scott Rauch[131], Abraham Reichenberg[132], Mark A. Riddle[133], Stephan Ripke[134,135,136], Maria C. Rosário[137], Aline S. Sampaio[138], Miriam A. Schiele[139], Anne Heidi Skogholt[140], Laura G. Sloofman[20], Jan Smit[141], María Soler Artigas[30,31,32,142], Laurent F. Thomas[143,144,145,146], Eric Tifft[33], Homero Vallada[147,148], Nathanial van Kirk[149,150], Jeremy Veenstra-VanderWeele[151,152], Nienke N. Vulink[153], Christopher P. Walker[154], Ying Wang[155], Jens R. Wendland[156], Bendik S. Winsvold[42,157,158], Yin Yao[159], Hang Zhou[160,161,162], Estonian Biobank*, 23andMe Inc.*, Arpana Agrawal[163], Pino Alonso[164,165,166,167], Götz Berberich[168], Kathleen K. Bucholz[169], Cynthia M. Bulik[21,170,171], Danielle Cath[172,173], Damiaan Denys[174], Valsamma Eapen[175,176], Howard Edenberg[177], Peter Falkai[178,179], Thomas V. Fernandez[180], Abby J. Fyer[181,182], J. M. Gaziano[183,184], Dan A. Geller[50,185], Hans J. Grabe[186], Benjamin D. Greenberg[187,188,189], Gregory L. Hanna[190], Ian B. Hickie[191], David M. Hougaard[52,76], Norbert Kathmann[1], James Kennedy[29], Dongbing Lai[192], Mikael Landén[21,193], Stéphanie Le Hellard[68,194], Marion Leboyer[195], Christine Lochner[196], James T. McCracken[115], Sarah E. Medland[197], Preben B. Mortensen[76,198,199], Benjamin M. Neale[200,201], Humberto Nicolini[202,203], Merete Nordentoft[204,205], Michele Pato[206], Carlos Pato[206], David L. Pauls[50], John Piacentini[207], Christopher Pittenger[208], Danielle Posthuma[209,210], Josep Antoni Ramos-Quiroga[211,212,213,214], Steven A. Rasmussen[110], Margaret A. Richter[29,215], David R. Rosenberg[216], Stephan Ruhrmann[217], Jack F. Samuels[218], Sven Sandin[20,21], Paul Sandor[29], Gianfranco Spalletta[35,219], Dan J. Stein[220], S. Evelyn Stewart[221,222,223], Eric A. Storch[224], Barbara E. Stranger[225,226], Maurizio Turiel[227], Thomas Werge[228,229], Ole A. Andreassen[230,231], Anders D. Børglum[4,232,233], Susanne Walitza[234,235,236], Kristian Hveem[158,237,238], Bjarne K. Hansen[239,240], Christian Rück[3], Nicholas G. Martin[241], Lili Milani[98], Ole Mors[242], Ted Reichborn-Kjennerud[243,244], Marta Ribasés[30,32,142,245], Gerd Kvale[83,246], David Mataix-Cols[3], Katharina Domschke[139,247], Edna Grünblatt[235,236,248], Michael Wagner[24,249,250], John-Anker Zwart[158,251], Gerome Breen[252], Gerald Nestadt[114], Jaakko Kaprio[253], Paul D. Arnold[254,255], Dorothy E. Grice[20], James A. Knowles[256], Helga Ask[257,258], Karin J. Verweij[153], Lea K. Davis[259], Dirk J. Smit[260], James J. Crowley[3,11,170], Jeremiah M. Scharf[118,135,##], Murray B. Stein[261,262,##], Joel Gelernter[263,264,##], Carol A. Mathews[265,##], Eske M. Derks[266,##] & Manuel Mattheisen[2,4,267,##]

*A list of cohort authors and their affiliations appears at the end of the paper.

[#]These authors contributed equally to this work.

[##]These authors jointly supervised this work.

Corresponding authors: Nora I. Strom (nora.strom@hu-berlin.de), Manuel Mattheisen

(manuel.mattheisen@dal.ca)

Affiliations

[1]Department of Psychology, Humboldt-Universität zu Berlin, Berlin, Germany. [2]Department of Psychiatric Phenomics and Genomics (IPPG), Ludwig-Maximilians University Munich, Munich, Germany. [3]Centre for Psychiatry Research, Department of Clinical Neuroscience, Karolinska Institutet & Stockholm Health Services, Region Stockholm, Stockholm, Sweden. [4]Department of Biomedicine, Aarhus University, Aarhus, Denmark. [5]Department of Mental Health and Neuroscience, Translational Neurogenomics, QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia. [6]Department of Population Health and Immunity, Healthy Development and Ageing, Walter and Eliza Hall Institute of Medical Research, Melbourne, Victoria, Australia. [7]Department of Psychiatry, Human Genetics, Yale University, New Haven, CT, USA. [8]VA Connecticut Healthcare System, West Haven, CT, USA. [9]Department of Center for Genomic Medicine, Psychiatric and Neurodevelopmental Genetics Unit, Massachusetts General Hospital, Boston, MA, USA. [10]Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge, MA, USA. [11]Department of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA. [12]Department of Psychiatry, Amsterdam University Medical Centers, University of Amsterdam, Amsterdam, The Netherlands. [13]CIMUS (Center for Research in Molecular Medicine and Chronic Diseases), Genomics and Bioinformatics, University of Santiago de Compostela, Santiago de Compostela, A Coruña, Spain. [14]Grupo de Medicina Xenómica, Genetics, FIDIS (Instituto de Investigación Sanitaria de Santiago de Compostela), Santiago de Compostela, A Coruña, Spain. [15]Vanderbilt Genetics Institute, Vanderbilt University, Nashville, TN, USA. [16]Department of Psychiatry and Behavioral Sciences, SUNY Downstate Health Sciences University, Brooklyn, NY, USA. [17]VA NY Harbor Healthcare System, Brooklyn, NY, USA. [18]Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, Psychology, and Neuroscience, King's College London, London, UK. [19]National Institute for Health and Care Research Maudsley Biomedical Research Centre, South London and Maudsley NHS Trust, London, UK. [20]Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, NY, USA. [21]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden. [22]Mental Health and Neuroscience Program, QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia. [23]Faculty of Medicine, School of Biomedical Sciences, University of Queensland, Brisbane, Queensland, Australia. [24]Department of Psychiatry and Psychotherapy, University Hospital Bonn, Bonn, Germany. [25]Department of Neurosciences and Mental Health, Hospital for Sick Children, Toronto, Ontario, Canada. [26]Department of Psychiatry, Brain, University Medical Center Utrecht, Utrecht, The Netherlands. [27]Second opinion outpatient clinic, GGNet, Warnsveld, The Netherlands. [28]Molecular Brain Science Department, Campbell Family Mental Health Research Institute, Centre for Addiction and Mental Health, Toronto, Ontario, Canada. [29]Department of Psychiatry, University of Toronto, Toronto, Ontario, Canada. [30]Psychiatric Genetics Unit, Group of Psychiatry, Mental Health and Addiction, Vall d'Hebron Research Institute (VHIR), Universitat Autònoma de Barcelona, Barcelona, Spain. [31]Department of Mental Health, Hospital Universitari Vall d'Hebron, Barcelona, Spain. [32]Biomedical Network Research Centre on Mental Health (CIBERSAM), Madrid, Spain. [33]Obsessive-Compulsive Disorder Institute, McLean Hospital, Belmont, MA, USA. [34]Department of Psychiatry & Behavioural Neurosciences, McMaster University, Hamilton, Ontario, Canada. [35]Laboratory of Neuropsychiatry, IRCCS Santa Lucia Foundation, Rome, Italy. [36]Department of Health Sciences, University of Milano, Milano, Italy. [37]Department of Child and Adolescent Psychiatry, Aarhus University Hospital, Psychiatry, Aarhus, Denmark. [38]Institute of Clinical Medicine, Health, Aarhus University, Aarhus, Denmark. [39]Department of Psychiatry and Behavioral Sciences, General Hospital Psychiatry, Johns Hopkins University School of Medicine, Baltimore, MD, USA. [40]Departments of Roy J. and Lucille A. Carver College of Medicine, University of Iowa, Iowa City, IA, USA. [41]Department of Child Study Center and Psychiatry, Yale University, New Haven, CT, USA. [42]Department of Research and Innovation, Division of Clinical Neuroscience, Oslo University Hospital, Oslo, Norway. [43]Department of Public Health and Nursing, Faculty of Medicine and Health Sciences, HUNT Center for Molecular and Clinical Epidemiology, Norwegian University of Science and Technology (NTNU), Trondheim, Norway. [44]Department of Medical Genetics, Oslo University Hospital, Oslo, Norway. [45]Department of Child and Adolescent Mental Health, Hospital Sant Joan de Déu, Esplugues de Llobregat, Barcelona, Spain. [46]Instituto de Salut Carlos III, Centro de Investigación Biomédica en Red de Salut Mental (CIBERSAM), Madrid, Spain. [47]Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, NY, USA. [48]Seaver Autism Center for Research and Treatment, Icahn School of Medicine at Mount Sinai, New York, NY, USA. [49]The Mindich Child Health and Development Institute, Icahn School of Medicine at Mount Sinai, New York, NY, USA. [50]Department of Psychiatry, Harvard Medical School, Boston, MA, USA. [51]Department of Psychiatry, Universidade de São Paulo, São Paulo, Brazil. [52]Department for Congenital Disorders, Statens Serum Institut, Copenhagen, Denmark. [53]Child Health Research Centre, University of Queensland, Brisbane, Queensland, Australia. [54]Pharmacogenetics Department, Investigaciones Clínicas, Instituto Nacional de Psiquiatría Ramon de la Fuente Muñiz, Mexico City, México. [55]Department of Surgery, Duke University, Durham, NC, USA. [56]Department of Psychiatry, Icahn School of Medicine at Mount Sinai Hospital, New York, NY, USA. [57]Department of Psychiatry, University of São Paulo, São Paulo, Brazil. [58]Center for Research in Molecular Medicine and Chronic Diseases (CiMUS), Genomics and Bioinformatics Group, University of Santiago de Compostela, Santiago de Compostela, Spain. [59]Galiician Foundation of Genomic Medicine, Grupo de Medicina Xenómica, Instituto de Investigación Sanitaria de Santiago -IDIS-, Santiago de Compostela, Spain. [60]Medicina Genómica, Centro de Investigación Biomédica en Red, Enfermedades Raras (CIBERER), Santiago de Compostela, Spain. [61]Programa MIND

Escoles, Hospital Sant Joan de Déu, Esplugues de Llobregat, Barcelona, Spain. [62]Departamento de Psiquiatría y Medicina Legal, Universitat Autònoma de Barcelona, Bellaterra, Spain. [63]Department of Psychiatry, Ospedale San Raffaele, Milano, Italy. [64]Department of Psychiatry, University of Illinois Chicago, Chicago, IL, USA. [65]Department of Psychiatry and Behavioral Sciences, Johns Hopkins Medical Institutions, Baltimore, MD, USA. [66]Department of Mental Health, Bloomberg School of Public Health, Baltimore, MD, USA. [67]Child and Adolesccent Psychiatry Department, APHP, Paris, France. [68]Department of Clinical Science, University of Bergen, Bergen, Norway. [69]Psychiatry, McLean Hospital OCDI, Harvard Medical School, Belmont, MA, USA. [70]Adult Psychological Services, CBTeam LLC, Lexington, MA, USA. [71]qGenomics (Quantitative Genomics Laboratories), Esplugues de Llobregat, Barcelona, Spain. [72]Department of Medical Epidemiology and Biostatistics, Center for Eating Disorders Innovation, Karolinska Institutet, Stockholm, Sweden. [73]Department of Psychiatry, Johns Hopkins University, Baltimore, MD, USA. [74]Department of Psychiatry and Behavioral Sciences, Child & Adolescent Psychiatry, Johns Hopkins University, Baltimore, MD, USA. [75]The Lundbeck Foundation Initiative for Integrative Psychiatric Research, iPSYCH, Aarhus, Denmark. [76]Center for Genomics and Personalized Medicine, Aarhus, Denmark. [77]Bioinformatics Research Centre, Aarhus, Denmark. [78]Genetic Epidemiology Research Branch, National Institute of Mental Health, Bethesda, MD, USA. [79]Department of Biomedicine, University of Bergen, Bergen, Norway. [80]Bergen Center for Brain Plasticity, Division of Psychiatry, Haukeland University Hospital, Bergen, Norway. [81]Department of Psychiatry, Møre og Romsdal Hospital Trust, Molde, Norway. [82]Bergen Center for Brain Plasticity, Psychiatry, Haukeland University Hospital, Bergen, Norway. [83]Department of Mental Health, Norwegian University for Science and Technology, Trondheim, Norway. [84]Million Veteran Program (MVP) Coordinating Center, VA Boston Healthcare System, Boston, MA, USA. [85]Department of Psychiatry, Boston University Chobanian & Avedisian School of Medicine, Boston, MA, USA. [86]PsychGen Centre for Genetic Epidemiology and Mental Health, Norwegian Institute of Public Health, Oslo, Norway. [87]Nic Waals Institute, Lovisenberg Diaconal Hospital, Oslo, Norway. [88]Bergen Center for Brain Plasticity, Haukeland University Hospital, Bergen, Norway. [89]Department of Medical Genetics, Dr. Einar Martens Research Group for Biological Psychiatry, Haukeland University Hospital, Bergen, Norway. [90]Department of Psychiatry, University of São Paulo, São Paulo, Brazil. [91]Department of Medicine, Genetic Medicine, Vanderbilt University Medical Center, Nashville, TN, USA. [92]Department of Child and Adolescent Psychiatry, NYU Grossman School of Medicine, New York, NY, USA. [93]Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, NY, USA. [94]Department of Bioinformatics and Integrative Biology, University of Massachusetts Medical School, Worcester, MA, USA. [95]Department of Vertebrate Genomics, Broad Institute of MIT and Harvard, Cambridge, MA, USA. [96]Department of Medicine, MSB Medical School Berlin, Berlin, Germany. [97]Department of Psychiatry and Behavioral Sciences, Johns Hopkins University, Baltimore, MD, USA. [98]Estonian Genome Centre, Institute of Genomics, University of Tartu, Tartu, Estonia. [99]Department of Biostatistics, T.H. Chan School of Public Health, Boston, MA, USA. [100]Department of Medicine, Channing Division of Network Medicine, Brigham and Women's Hospital, Boston, MA, USA. [101]Research Department, Grupo Medico Carracci, Mexico City, Mexico. [102]Department of Psychiatry, Yale University, West Haven, CT, USA. [103]Office of Research & Development, United States Department of Veterans Affairs, West Haven, CT, USA. [104]Department of Medical Biochemistry and Microbiology, Uppsala University, Uppsala, Sweden. [105]SciLifeLab, Uppsala University, Uppsala, Sweden. [106]Department of Psychiatry, University of California, Irvine (UCI), Irvine, CA, USA. [107]Department of Mental Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA. [108]Vanderbilt Genetics Institute, Vanderbilt University Medical Center, Nashville, TN, USA. [109]COBRE Center for Neuromodulation, Butler Hospital, Providence, RI, USA. [110]Department of Psychiatry & Human Behavior, Alpert Medical School, Brown University, Providence, RI, USA. [111]Butler Hospital, Providence, RI, USA. [112]Department of Psychiatry, Dalhousie University, Halifax, Nova Scotia, Canada. [113]Department of Psychiatry, Faculdade de Medicina da Universidade de São Paulo, São Paulo, Brazil. [114]Department of Psychiatry and Behavioral Science, Johns Hopkins University, Baltimore, MD, USA. [115]Department of Psychiatry and Biobehavioral Sciences, Division of Child and Adolescent Psychiatry, University of California, Los Angeles, Los Angeles, CA, USA. [116]Department of Clinical Medicine, Division of Mental Health and Addiction, Oslo University Hospital, Oslo, Norway. [117]NORMENT, University of Oslo, Oslo, Norway. [118]Department of Psychiatry, Massachusetts General Hospital, Boston, MA, USA. [119]Psychiatric and Neurodevelopmental Genetics Unit, Harvard Medical School, Boston, MA, USA. [120]Department of Biomedicine, Haukeland University Hospital, Bergen, Norway. [121]Institute for Molecular Medicine Finland - FIMM, University of Helsinki, Helsinki, Finland. [122]Department of Clinical Neuroscience and Neurorehabilitation, Neuropsychiatry Laboratory, IRCCS Santa Lucia Foundation, Rome, Italy. [123]Department of Genetics, Microbiology and Statistics, IBUB, Universitat de Barcelona, Barcelona, Spain. [124]CIBERER, Centro de investigación biomédica en red, Madrid, Spain. [125]Department of Human Molecular Genetics, Institut de Recerca Sant Joan de Déu, Esplugues de Llobregat, Barcelona, Spain. [126]Department of Psychiatry and Psychotherapy, Division of Neurogenetics and Molecular Psychiatry, Faculty of Medicine and University Hospital Cologne, University of Cologne, Cologne, Germany. [127]Department of Neurodegenerative Diseases and Geriatric Psychiatry, University Hospital Bonn, Medical Faculty, Bonn, Germany. [128]DZNE Bonn, German Center for Neurodegenerative Diseases (DZNE), Bonn, Germany. [129]Department of Psychiatry and Glenn Biggs Institute for Alzheimer's and Neurodegenerative Diseases, UT Health San Antonio, San Antonio, TX, USA. [130]Cologne Excellence Cluster for Stress Responses in Ageing-associated diseases (CECAD), University of Cologne, Cologne, Germany. [131]Department of Psychiatry, McLean Hospital, Harvard Medical School, Belmont, MA, USA. [132]Department of Mental Disorders, Norwegian Institute of Public Health, New York, NY, USA. [133]Department of Psychiatry and Behavioral Sciences, Child and Adolescent, Johns Hopkins University School of Medicine, Baltimore, MD, USA. [134]Department of

Psychiatry and Psychotherapy, Charité - Universitätsmedizin, Berlin, Germany. [135]Stanley Center for Psychiatric Research, Broad Institute of Harvard and MIT, Cambridge, MA, USA. [136]site Berlin-Potsdam, German Center for Mental Health (DZPG), Berlin, Germany. [137]Department of Psychiatry, Child and Adolescent Psychiatry Unit (UPIA), Federal University of São Paulo (UNIFESP), São Paulo, Brazil. [138]Department of Neurosciences and Mental Health, Medical School, Federal University of Bahia, Salvador, Brazil. [139]Department of Psychiatry and Psychotherapy, Faculty of Medicine, University of Freiburg, Medical Center - University of Freiburg, Freiburg, Germany. [140]Department of Public Health and Nursing, HUNT Center for Molecular and Clinical Epidemiology, Trondheim, Norway. [141]Department of Psychiatry, Faculty of Medicine, Locaion Vumc, Amsterdam UMC, University of Amsterdam, Amsterdam, The Netherlands. [142]Department of Genetics, Microbiology, and Statistics, Faculty of Biology, Universitat de Barcelona (UB), Barcelona, Spain. [143]Department of Clinical and Molecular Medicine, Norwegian University of Science and Technology, Trondheim, Norway. [144]Department of Public Health and Nursing, K. G. Jebsen Center for Genetic Epidemiology, Norwegian University of Science and Technology, Trondheim, Norway. [145]BioCore - Bioinformatics Core Facility, Norwegian University of Science and Technology, Trondheim, Norway. [146]Clinic of Laboratory Medicine, St.Olavs Hospital, Trondheim University Hospital, Trondheim, Norway. [147]Department of Psychiatry, Universidade de Sao Paulo, São Paulo, Brazil. [148]Department of Molecular Medicine and Surgery, CMM, Karolinska Institutet, Stockholm, Sweden. [149]OCD Institute, Division of Depression and Anxiety, McLean Hospital, Belmont, MA, USA. [150]Department of Psychiatry, Harvard Medical School, Belmont, MA, USA. [151]Department of Psychiatry, Division of Child and Adolescent Psychiatry, Columbia University, New York, NY, USA. [152]Department of Child and Adolescent Psychiatry, New York State Psychiatric Institute, New York, NY, USA. [153]Department of Psychiatry, Amsterdam UMC, University of Amsterdam, Amsterdam, The Netherlands. [154]Human Genetics Institute of New Jersey, Rutgers University, Piscataway, NJ, USA. [155]Department of Neurology, The Johns Hopkins University School of Medicine, Baltimore, MD, USA. [156]Laboratory of Clinical Science, NIMH Intramural Research Program, Bethesda, MD, USA. [157]Department of Neurology, Oslo University Hospital, Oslo, Norway. [158]HUNT Center for Molecular and Clinical Epidemiology, Department of Public Health and Nursing, Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology, Trondheim, Norway. [159]Department of Computional Biology, Institute of Life Science, Fudan University, Fudan, China. [160]Department of Psychiatry, Yale School of Medicine, New Haven, CT, USA. [161]Department of Psychiatry, Veterans Affairs Connecticut Healthcare System, West Haven, CT, USA. [162]Section of Biomedical Informatics and Data Science, Yale School of Medicine, New Haven, CT, USA. [163]Department of Psychiatry, Washington University School of Medicine, Saint Louis, MO, USA. [164]Department of Psychiatry, OCD Clinical and Research Unit, Bellvitge Hospital, Barcelona, Spain. [165]Department of Clinical Sciences, University of Barcelona, Barcelona, Spain. [166]Department of Psychiatry and Mental Health, Bellvitge Biomedical Research Institute IDIBELLL, Barcelona, Spain. [167]CIBERSAM, Mental Health Network Biomedical Research Center, Madrid, Spain. [168]Psychosomatic Department, Windach Hospital of Neurobehavioural Research and Therapy, Windach, Germany. [169]Department of Psychiatry, Washington University School of Medicine, St Louis, MO, USA. [170]Department of Psychiatry, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA. [171]Department of Nutrition, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA. [172]Departments of Rijksuniversiteit Groningen and Psychiatry, University Medical Center Groningen, Groningen, The Netherlands. [173]Department of Specialized Training, Drenthe Mental Health Care Institute, Groningen, The Netherlands. [174]Department of Psychiatry, Institute of The Royal Netherlands Academy of Arts and Sciences (NIN-KNAW), Amsterdam, The Netherlands. [175]Discipline of Psychiatry and Mental Health, School of Clinical Medicine, UNSW, Sydney, New South Wales, Australia. [176]Academic Unit of Child Psychiatry South-West Sydney (AUCS), South-West Sydney Clinical School, SWSLHD & Ingham Institute, Sydney, New South Wales, Australia. [177]Department of Biochemistry and Molecular Biology, Indiana University School of Medicine, Indianapolis, IN, USA. [178]Department of Psychiatry and Psychotherapy, University Hospital LMU, Munich, Germany. [179]Department of Psychiatry, Max Planck Institute, Munich, Germany. [180]Child Study Center and Psychiatry, Yale School of Medicine, New Haven, CT, USA. [181]Department of Psychiatry, New York State Psychiatric Institute, New York, NY, USA. [182]Department of Psychiatry, Columbia University Medical Center, New York, NY, USA. [183]Department of Medicine, VA Boston Healthcare System, Boston, MA, USA. [184]Department of Medicine, Mass General Brigham, Boston, MA, USA. [185]Department of Psychiatry, Child Psychiatry, Massachusetts General Hospital, Boston, MA, USA. [186]Department of Psychiatry & Psychotherapy, University Medicine Greifswald, Greifswald, Germany. [187]COBRE Center on Neuromodulation, Butler Hospital, Providence, RI, USA. [188]Center for Neurorestoration and Neurotechnology, VA Providence Healthcare System, Providence, RI, USA. [189]Department of Psychiatry and Human Behavior, Alpert Medical School, Brown University, Providence, RI, USA. [190]Department of Psychiatry, Child and Adolescent Psychiatry, University of Michigan, Ann Arbor, MI, USA. [191]Brain and Mind Centre, The University of Sydney, Sydney, New South Wales, Australia. [192]Department of Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, IN, USA. [193]Institute of Neuroscience and Physiology, Department of Psychiatry and Neurochemistry, University of Gothenburg, Gothenburg, Sweden. [194]Bergen Center for Brain Plasticity, Division of Psychiatry, Haukeland University Hospital, Bergen, Norway. [195]Department of Addictology and Psychiatry, Univ Paris Est Créteil, AP-HP, Inserm, Paris, France. [196]Department of Psychiatry, SA MRC Unit on Risk and Resilience in Mental Disorders, Stellenbosch University, Stellenbosch, South Africa. [197]Department of Mental Health, Psychiatric Genetics, QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia. [198]National Centre for Register-based Research, Aarhus University, Aarhus, Denmark. [199]Centre for Integrated Register-based Research, Aarhus University, Aarhus, Denmark. [200]Analytic and Translational Genetics Unit, Massachusetts General Hospital, Boston, MA, USA.

[201]Stanley Center for Psychiatric Research, Broad Institute of Harvard and MIT, Cambridge, MA, USA. [202]Department of Psychiatry, Psychiatry, Carracci Medical Group, Mexico City, México. [203]Psiquiatría, Instituto Nacional de Medicina Genómica, Mexico City, México. [204]Mental Health Center Copenhagen, Copenhagen Research Center for Mental Health, Mental Health Services in the Capital Region of Denmark, Copenhagen, Denmark. [205]Faculty of Health and Medical Sciences, Department of Clinical Medicine, University of Copenhagen, Copenhagen, Denmark. [206]Department of Psychiatry, Rutgers University, Piscataway, NJ, USA. [207]Department of Psychiatry and Biobehavioral Sciences, Child and Adolescent Psychiatry, UCLA Semel Institute for Neuroscience and Human Behavior, Los Angeles, CA, USA. [208]Department of Psychiatry, Yale University, New Haven, CT, USA. [209]Department of Complex Trait Genetics, Vrije Universiteit Amsterdam, Center for Neurogenomics and Cognitive Research, Amsterdam Neuroscience, Amsterdam, The Netherlands. [210]Department of Child and Adolescent Psychiatric, Section Complex Trait Genetics, VU Medical Center Amsterdam, Amsterdam, The Netherlands. [211]Department of Psychiatry, Hospital Universitari Vall d'Hebron, Barcelona, Spain. [212]Group of Psychiatry, Mental Health and Addictions, Psychiatric Genetics Unit, Vall d'Hebron Research Institute (VHIR), Barcelona, Spain. [213]Biomedical Network Research Centre on Mental Health (CIBERSAM), Barcelona, Spain. [214]Department of Psychiatry and Forensic Medicine, Universitat Autònoma de Barcelona, Barcelona, Spain. [215]Department of Psychiatry, Sunnybrook Health Sciences Centre, Toronto, Ontario, Canada. [216]Department of Psychiatry and Behavioral Neurosciences, Child and Adolescent Psychiatry, Wayne State University School of Medicine, Detroit, MI, USA. [217]Department of Psychiatry and Psychotherapy, University of Cologne, Cologne, Germany. [218]Department of Psychiatry and Behavioral Sciences, The Johns Hopkins University School of Medicine, Baltimore, MD, USA. [219]Department of Psychiatry and Behavioral Sciences, Division of Neuropsychiatry, Baylor College of Medicine, Houston, TX, USA. [220]Department of Psychiatry & Neuroscience Institute, SAMRC Unit on Risk & Resilience in Mental Disorders, University of Cape Town, Cape Town, Western Cape, South Africa. [221]Department of Psychiatry, University of British Columbia, Vancouver, British Columbia, Canada. [222]British Columbia Children's Hospital Research Institute, Vancouver, British Columbia, Canada. [223]British Columbia Mental Health and Substance Use Services Research Institute (BCMHSUS), Vancouver, British Columbia, Canada. [224]Department of Psychiatry and Behavioral Sciences, Baylor College of Medicine, Houston, TX, USA. [225]Department of Pharmacology, Northwestern University Feinberg School of Medicine, Chicago, IL, USA. [226]Center for Genetic Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL, USA. [227]Department of Cardiology, University of Milan, Milan, Italy. [228]Institute of Biological Psychiatry, Mental Health Center Sct. Hans, Copenhagen University Hospital, Mental Health Services (RHP), Copenhagen, Denmark. [229]Institute of Clinical Medicine, University of Copenhagen, Copenhagen, Denmark. [230]Institute of Clinical Medicine, NORMENT Centre, University of Oslo, Oslo, Norway. [231]Division of Mental Health and Addiction, Center for Precision Psychiatry, Oslo University Hospital, Oslo, Norway. [232]The Lundbeck Foundation Initiative for Integrative Psychiatric Research, iPSYCH, Aarhus University, Aarhus, Denmark. [233]Center for Genomics and Personalized Medicine, CGPM, Aarhus University, Aarhus, Denmark. [234]Department of Child and Adolescent Psychiatry and Psychotherapy, Psychiatric University Hospital Zurich (PUK), University of Zurich, Zürich, Switzerland. [235]Neuroscience Center Zurich, University of Zurich and the ETH Zuric, Zurich, Switzerland. [236]Zurich Center for Integrative Human Physiology, University of Zurich, Zurich, Switzerland. [237]HUNT Research Center, Department of Public Health and Nursing, Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology, Trondheim, Norway. [238]Department of Research, Innovation and Education, St. Olavs Hospital, Trondheim University Hospital, Trondheim, Norway. [239]Bergen Center for Brain Plasticity (BCBP), Psychiatry, Haukeland University Hospital, Bergen, Norway. [240]Centre for Crisis Psychology, Psychology, University of Bergen, Bergen, Norway. [241]Department of Genetic Epidemiology, QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia. [242]Psychosis Research Unit, Aarhus University Hospital - Psychiatry, Aarhus, Denmark. [243]Department of Mental Disorders, Norwegian Institute of Public Health, Oslo, Norway. [244]Institute of Clinical Medicine, University of Oslo, Oslo, Norway. [245]Department of Mental Health, Hospital Universitari Vall d'Hebron, Barcelona, Spain. [246]Department of Clinical Psychology, Faculty of Psychology, University of Bergen, Bergen, Norway. [247]German Center for Mental Health (DZPG), Partner Site Berlin, Berlin, Germany. [248]Department of Child and Adolescent Psychiatry and Psychotherapy, Psychiatric University Hospital Zurich (PUK), University of Zurich, Zürich, Switzerland. [249]Department of Neurodegenerative Diseases and Geriatric Psychiatry, University Hospital Bonn, Bonn, Germany. [250]German Center for Neurodegenerative Diseases, DZNE, Bonn, Germany. [251]Department of Research and Innovation, Clinical Neuroscience, Oslo University Hospital and University of Oslo, Oslo, Norway. [252]Social, Genetic, and Developmental Psychiatric Centre, Institute of Psychiatry, Psychology, and Neuroscience, King's College London, London, UK. [253]Institute for Molecular Medicine FIMM, University of Helsinki, Helsinki, Finland. [254]Department of Psychiatry, The Mathison Centre for Mental Health Research & Education, Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada. [255]Program in Genetics and Genome Biology, Hospital for Sick Children, Toronto, Ontario, Canada. [256]Department of Genetics, Human Genetics Institute of New Jersey, Rutgers University, Piscataway, NJ, USA. [257]PsychGen Center for Genetic Epidemiology, Norwegian Institute of Public Health, Oslo, Norway. [258]PROMENTA Research Center, Department of Psychology, University of Oslo, Oslo, Norway. [259]Department of Medicine, Division of Genetic Medicine, Vanderbilt University Medical Center, Nashville, TN, USA. [260]Department of Psychiatry, Amsterdam UMC location AMC, Amsterdam, The Netherlands. [261]Psychiatry Service, VA San Diego Healthcare System, San Diego, CA, USA. [262]Department of Psychiatry and School of Public Health, University of California San Diego, La Jolla, CA, USA. [263]Department of Psychiatry, Human Genetics (Psychiatry), Yale University School of Medicine, West Haven, CT, USA. [264]Department of Psychiatry, Veterans Affairs Connecticut Healthcare

Center, West Haven, CT, USA. [265]Psychiatry and Genetics Institute, Center for OCD, Anxiety and Related Disorders, University of Florida, Gainesville, FL, USA. [266]Department of Mental Health and Neuroscience, QIMR Berghofer, Brisbane, Queensland, Australia. [267]Department of Community Health and Epidemiology and Faculty of Computer Science, Dalhousie University, Halifax, Nova Scotia, Canada.

Abstract

Obsessive-compulsive disorder (OCD) affects ~1% of children and adults and is partly caused by genetic factors. We conducted a genome-wide association study (GWAS) meta-analysis combining 53,660 OCD cases and 2,044,417 controls and identified 30 independent genome-wide significant loci. Gene-based approaches identified 249 potential effector genes for OCD, with 25 of these classified as the most likely causal candidates, including *WDR6*, *DALRD3*, *CTNND1* and multiple genes in the MHC region. We estimated that ~11,500 genetic variants explained 90% of OCD genetic heritability. OCD genetic risk was associated with excitatory neurons in the hippocampus and cortex, along with D1- and D2-type dopamine receptor-containing medium spiny neurons. OCD genetic risk was shared with 65 of 112 additional phenotypes, including all of the psychiatric disorders we examined. In particular, OCD shared genetic risk with anxiety, depression, anorexia nervosa, and Tourette syndrome, and was negatively associated with inflammatory bowel diseases, educational attainment, and body mass index.

Introduction

Obsessive-compulsive disorder (OCD) is a chronic psychiatric illness that affects 1-3% of the population[1] and is characterized by obsessions and compulsions that vary in type and severity, and over time. OCD is responsible for profound personal and societal costs[2], including increased risk of suicide[3] and overall mortality[4]. OCD is moderately heritable; twin-based heritability estimates range between 27-47% in adults and between 45-65% in children[5–8], with SNP-based heritability estimates between 28-37%[9–11].

Two earlier OCD GWAS meta-analyses, both containing a subset of the data included in this analysis[12,13], showed SNP-based heritabilities of 8.5% (assuming a 3% population prevalence) and 16% (assuming a 2% population prevalence). The first GWAS (n_{cases} = 14,140, $n_{controls}$ = 562,117)[12] found one genome-wide significant locus associated with OCD, while the second (n_{cases} = 37,015, $n_{controls}$ = 948,616)[13] identified 15 independent genome-wide significant loci. As with other complex traits, increased sample sizes are needed for a more comprehensive understanding of the underlying genetic etiology of OCD and its genetic relationships with related disorders.

The current study combines data from the two unpublished OCD GWASs described above and includes additional cohorts (~9,000 cases). This results in the largest and most well-powered GWAS of OCD to date, with a ~20-fold increase of OCD cases compared to the previously published OCD GWASs[10]. Based on the results from the meta-analysis, we conducted secondary analyses, including positional and functional fine-mapping of SNPs and genes, structural equation modeling to examine possible genetic differences in sample ascertainment across cohorts, protein and transcriptome-wide association analyses, single-cell enrichment, and genetic correlations with other traits (**Supplementary Fig. 1**). Our results provide more detailed insight into the genetic underpinnings and biology of OCD.

Results

GWAS meta-analysis identifies 30 genome-wide significant loci

We conducted a GWAS meta-analysis of 28 European-ancestry OCD case-control cohorts, comprising 53,660 cases and 2,044,417 controls (effective sample size ~210,000 individuals). Ascertainment of cases varied across cohorts: OCD diagnosis was determined (i) by a healthcare professional in a clinical setting (18 cohorts, *n* = 9,089 cases), (ii) from health records or biobanks (7 cohorts, *n* = 9,138 cases), (iii) in a clinical setting or from health records with the additional characteristic that all OCD cases were primarily collected for another psychiatric disorder (3 cohorts, *n* = 5,266 cases), or (iv) by self-reported diagnosis in a consumer-based setting (23andMe, Inc., *n* = 30,167 cases). Cohort details, including phenotypic assessment, quality control, and individual cohort GWAS analyses are described in **Supplementary Note 2** and **Supplementary Table 1**. We identified 30 independent (defined in **Supplementary Note 3**) loci among the 1,672 SNPs that exceeded the genome-wide threshold for significance ($P < 5 \times 10^{-8}$; Manhattan plot in **Fig. 1**, regional association plots and forest plots in **Supplementary Figs. 2-31**, and a list of all independent genome-wide significant SNPs in **Table 1** with additional details in **Supplementary Tables 2** and **3**). The independence of the 30 lead SNPs was subsequently validated using conditional and joint analysis (GCTA-COJO)[14] (**Supplementary Table 4**). An analysis of the X-chromosome, conducted in a subset of the data for which this information was available (23andMe), yielded no significant associations (**Supplementary Note 4** and **Supplementary Figure 37e**). Of the 15 genome-wide significant loci previously reported in preprints[12,13], 13 were genome-wide significant in the current GWAS, with the remaining two showing suggestive significance ($P = 5.23 \times 10^{-8}$ and $P = 2.2 \times 10^{-7}$; **Supplementary Table 5**). Using MiXeR[15], we estimated that approximately 11,500 (s.e. = 607) causal variants account for 90% of the OCD SNP-based heritability.

No statistically significant heterogeneity was observed across individual cohorts for the 30 genome-wide significant loci, as assessed with Cochran's Q-test (**Supplementary Fig. 32**), the I^2 statistic, and GenomicSEM's Q_{SNP}-statistic[16] (**Supplementary Table 2**). Genome-wide analyses of samples grouped by clinical, comorbid, biobank, and 23andMe (**Supplementary Table 3** and **Supplementary Figs. 33-37**) showed evidence that sample ascertainment impacted results at a genome-wide scale, although not beyond what is observed with closely related psychiatric disorders[17,18]. We observed moderate to high genetic correlations across the subgroups (between 0.63, s.e. = 0.11 for biobanks and comorbid, and 0.92, s.e. = 0.07 for 23andMe and comorbid; **Supplementary Table 7**), and a satisfactory fit for a one-factor GenomicSEM model (**Supplementary Table 8** and **Supplementary Fig. 39**). A common-factor GWAS based on the one-factor GenomicSEM model resulted in 20 significant loci, all of which were also significant in the primary GWAS (**Supplementary Table 8** and **Supplementary Fig. 40**; analysis details in **Supplementary Note 5**). SNP-heritability (assuming a 1% population prevalence) was 6.7% (s.e. = 0.3%), with slightly higher estimates for the clinical (h^2_{SNP} = 16.4%, s.e. = 1.5%) and comorbid (h^2_{SNP} = 13.3%, s.e. = 1.7%) subgroups (**Supplementary Table 1**).

Gene-based findings

We prioritized putative risk genes for OCD using six positional and functional QTL gene-based mapping approaches. Positional mapping was performed with mBAT-combo[19]. Functional eQTL mapping was performed with transcriptome-wide association study (TWAS)[20] using PsychENCODE gene expression weights[21], and summary-based Mendelian randomization (SMR)[22] using whole blood eQTLGen[23] and MetaBrain[24] datasets. Functional protein QTL (pQTL) mapping was done using a protein-wide association study (PWAS) of human brain protein expression panels[25]. Finally, we used PsyOPS[26], which combines positional mapping with biological annotations, to further prioritize risk genes within genome-wide significant loci. We identified 207 significant genes (Bonferroni correction, $P < 2.67 \times 10^{-6}$) with mBAT-combo, and 24 genes using TWAS ($P < 4.76 \times 10^{-6}$), 14 of which were conditionally independent. The SMR-eQTLGen analysis identified 39 significant risk genes ($P < 4.28 \times 10^{-6}$), and the SMR-MetaBrain analysis identified 14 risk genes ($P < 9.23 \times 10^{-6}$). The PWAS identified 3 significant genes ($P < 3.39 \times 10^{-5}$), while PsyOPS prioritized 29 genes. In total, 251 genes were significantly associated with OCD through at least one gene-based approach, and 48 were implicated by at least two methods (details in **Online Methods**, **Supplementary Note 7**, and **Supplementary Tables 9-14**).

From the 48 genes implicated by at least two approaches, we prioritized likely causal genes for OCD using colocalization (TWAS-COLOC)[27,28] and SMR-heterogeneity in dependent instruments (SMR-HEIDI)[22] tests. Colocalization was used to identify significant TWAS associations where the underlying GWAS and eQTL summary statistics are likely to share a single causal variant. Similarly, HEIDI was used to select SMR associations where the same causal variant affects gene expression and trait variation. 25 of the 48 genes implicated by at least two gene-based tests were also significant in either the TWAS-COLOC or SMR-HEIDI tests, implying causality (**Fig. 2a**). Only two of these 25 genes were prioritized by both TWAS-COLOC and SMR-HEIDI, WD repeat domain 6 (*WDR6*) and DALR Anticodon Binding Domain Containing 3 (*DALRD3*). Another gene of interest, Catenin Delta 1 (*CTNND1*), was implicated by three of our five approaches (mBAT-combo, TWAS, PWAS) and showed evidence for colocalization. Only three genes were implicated in the PWAS; of these, *CTNND1* was the only gene also implicated in the TWAS. In the PWAS, downregulation of *CTNND1* protein expression in human dorsolateral prefrontal cortex was significantly associated with OCD risk ($Z = -4.49,\ P = 7.11 \times 10^{-6}$; **Supplementary Table 13**), consistent with the downregulation of *CTNND1* gene expression in prefrontal cortex seen in the TWAS ($Z = -6.86,\ P = 6.90 \times 10^{-12}$; **Supplementary Table 10**). For a discussion of the overlap between the gene-findings with rare-coding variants in OCD, see **Supplementary Table 6** and **Supplementary Note 7**.

Tissue and cell type enrichment analysis

After mapping significantly associated SNPs from the GWAS meta-analysis to likely causal genes, we explored which tissues or cell types showed enriched gene expression of OCD associated genetic signals, using a previously described approach[29] on published human gene expression datasets from bulk tissue RNA-seq data from GTEx and single-cell RNA-sequencing data from the adult mouse central and peripheral nervous systems (CNS and PNS)[30]. We found enrichment of OCD GWAS signals in 6 of 13 human brain tissue types in GTEx, but no enrichment in human peripheral tissues (**Fig. 2b** and **Supplementary Table 15**). In the adult mouse CNS and PNS, we found enrichment of OCD GWAS signals in 41 of 166 tested specific single cell types using the MAGMA gene-set enrichment test (**Supplementary Table 16**). When summarizing results of

individual single cell types into groups of cell types defined by the same region/tissue and cell type, 9 of 35 are enriched for OCD GWAS signals (top 20 shown in **Fig. 2c**). Strong enrichment of OCD GWAS signal was especially observed in excitatory neurons of hippocampus and cerebral cortex, as well as D1 and D2 medium spiny neurons.

Genetic relationship of OCD with other phenotypes

Using phenome-wide association analysis, we examined whether the 30 independent OCD-associated loci identified by our GWAS meta-analysis have previously been associated with other phenotypes (see **Supplementary Tables 17a-d** for look-ups in four, partially overlapping GWAS-databanks and **Table 1** for highlighted associations). We found that 22 of the 30 loci were associated with other phenotypes, including schizophrenia (seven loci), depression/major depressive disorder (two loci), bipolar disorder (one locus), neuroticism (seven loci), educational attainment (seven loci), and body fat mass or body mass index (eight loci).

We further used bivariate linkage-disequilibrium score regression (LDSC)[31] to investigate the extent of genetic correlations between OCD and 112 previously published GWASs encompassing psychiatric, substance use, and neurological phenotypes, among others (**Fig. 3**). We found that 65 phenotypes were significantly correlated with OCD after correcting for multiple testing using the Benjamini-Hochberg[32] procedure to control the false-discovery rate (FDR) at a threshold of 0.05. OCD was significantly positively correlated with all tested psychiatric phenotypes, the highest correlations with anxiety (ANX, $r_G = 0.70$), depression (DEP, $r_G = 0.60$), anorexia nervosa (AN, $r_G = 0.52$), Tourette syndrome (TS, $r_G = 0.47$), and post-traumatic stress disorder (PTSD, $r_G = 0.48$). Significant positive genetic correlations were also obtained for neuroticism ($r_G = 0.53$), in particular for the worry subcluster ($r_G = 0.64$), and all individual items in the worry subcluster, with slightly lower estimates for the depressive sub-cluster ($r_G = 0.35$). Suicide attempt ($r_G = 0.40$), history of childhood maltreatment ($r_G = 0.37$), and tiredness ($r_G = 0.36$) were also notable for strong positive associations with OCD. Of the assessed neurological disorders, OCD was only significantly correlated with migraine ($r_G = 0.15$). Some autoimmune disorders, such as Crohn's disease ($r_G = -0.13$), ulcerative colitis ($r_G = -0.14$), and inflammatory bowel disease ($r_G = -0.14$), showed negative correlations with OCD (see **Fig. 3**, **Supplementary Table 18** for all genetic correlation estimates, 95% CI and *P*-values, **Supplementary Note 6** for a more in depth discussion of all significant genetic correlations, and **Supplementary Table 19** and **Supplementary Figs. 41** and **42** for sub-group specific genetic-correlation estimates).

Discussion

The OCD GWAS reported here, comprising over 53,000 cases, identified 30 independent genome-wide significant loci. Common SNPs explained 6.7% of the variation in OCD risk in our meta-analysis (LDSC with an assumed population prevalence of 1%), a significant reduction from the 28% reported previously[10]. However, differences in the assumed population prevalence (where a lower assumed prevalence for LDSC heritability calculation results in a lower heritability estimate) and an increase in sample heterogeneity likely contributed to this discrepancy. The reduction in SNP-heritability is in line with previous observations for closely related psychiatric disorders such as ADHD[33,34] or depression[17,35–37], where expanding the phenotype definition increased genetic heterogeneity, potentially accounting for the observed decrease in SNP heritability. This aligns with the fact that heritability estimates for more homogeneous OCD

subgroups were higher: 16.4% for the clinically-ascertained subgroup and 13.3% for the comorbid subgroup (**Supplementary Note 10**). The current estimates are comparable to those of other psychiatric and substance use disorders, with SNP-heritability estimates ranging between 9% and 28%[38].

The most significant SNP rs78587207 ($P = 5.28 \times 10^{-12}$) identified in the GWAS is located on chr11q12.1 and has been previously associated with several traits, including neuropsychiatric phenotypes[39] such as depressive symptoms[40] and neuroticism[40]. Gene-based analyses identified four putative causal genes within this locus. The closest gene to rs78587207 is catenin delta 1 (*CTNND1*), which encodes the cell adhesion molecule p120-catenin. This gene was associated with OCD using three gene-based tests (mBAT-combo, TWAS, and PWAS) and we found strong evidence for colocalization of the TWAS signal for *CTNND1* in dorso-lateral prefrontal cortex (dlPFC). dlPFC has been consistently implicated in the neural circuitry of OCD as well as in compulsivity more broadly as part of the cortico-striatal-thalamo-cortical circuitry[41,42]. The protein product of *CTNND1* is a regulator of cell-cell adhesion[43] and plays a crucial role in gene transcription, Rho GTPase activity, and cytoskeletal organization[44–46]. Other credible causal genes in the locus include cleavage factor polyribonucleotide kinase subunit 1 (*CLP1*), thioredoxin related transmembrane protein 2 (*TMX2*), and zinc finger DHHC-type palmitoyltransferase 5 (*ZDHHC5*). Rare genetic mutations in *CLP1* are associated with pontocerebellar hypoplasia type 10, a very rare autosomal recessive neurodegenerative disease characterized by brain atrophy and delayed myelination resulting in intellectual disability[47]. *TMX2* is associated with increased risk of neurodevelopmental disorders with microcephaly, cortical malformations, spasticity, and congenital nervous system abnormalities[48]. *ZDHHC5* is broadly expressed in the brain, including the frontal cortex. *ZDHHC5* has not been implicated in brain development but has been linked to lung acinar adenocarcinoma and lung papillary adenocarcinoma in prior studies[49].

Our finding that approximately 11,500 (s.e. = 607) causal variants account for 90% of the SNP-based heritability of OCD suggests that OCD is more polygenic than other complex traits such as height (n_{causal} = 4,000), schizophrenia (n_{causal} = 9,600) and ADHD (n_{causal} = 5,600), but less polygenic than major depression (n_{causal} = 14,500) and educational attainment (n_{causal} = 13,200)[50].

We identified a total of 25 credible causal genes based on robust evidence using multiple positional and functionally-informed gene-based approaches. Notably, *DLGAP1*, which has been previously implicated in OCD pathogenesis[10,51], was not identified in either the GWAS or in the gene-based analyses. Of the 25 genes that were implicated, 15 were within 6.5 kb of a SNP that surpassed genome-wide significance in the meta-analysis. In addition to the four genes discussed above, several others are of particular interest, including *WDR6* and *DALRD3*, which had the strongest evidence from the gene-based analyses. These genes lie in a gene-rich region on chr3p21.31, which in addition to harboring multiple genome-wide significant SNPs, has been previously associated with a broad range of psychiatric disorders and related traits, including schizophrenia[39], well-being[52], and the worry-subcluster of neuroticism[53].

WDR6 (WD repeat domain 6) is broadly expressed in the brain, particularly the hypothalamus. Its protein product is involved in cell growth arrest[54], and recent studies have implicated it in anorexia nervosa[55] and Parkinson's disease[56]. *DALRD3* (DALR anticodon binding domain containing 3) is located on chromosome 3 in the same region as *WDR6. DALRD3*, when fully disrupted, is implicated in a form of epileptic encephalopathy with associated developmental delay[57]. Finally, a third gene in the 3p21 locus, *CELSR3* (cadherin EGF LAG seven-pass G-type receptor 3), encodes a protocadherin that is highly expressed in the developing basal ganglia[58]. Multiple loss-of-function mutations in *CELSR3* have been associated with Tourette syndrome[59,60], which co-occurs with OCD in 10-20% of patients.

Four other genes identified through these analyses are located in the MHC locus, a region on chromosome 6 that plays a major role in the adaptive immune system and has been repeatedly linked to major psychiatric disorders[61]. The newly identified MHC association for OCD is noteworthy given evidence linking OCD with autoimmune disorders[62–64]. Genetic pleiotropy may underlie this connection, with variants predisposing individuals to both autoimmune conditions and OCD[65]. Further, some OCD subtypes, such as pediatric acute-onset neuropsychiatric disorders associated with streptococcus (PANDAS) and pediatric acute-onset neuropsychiatric syndrome (PANS) may have autoimmune origins[66,67]. Nevertheless, we were surprised to discover several negative genetic correlations between OCD and autoimmune disorders such as Crohn's disease, ulcerative colitis, and inflammatory bowel disease in our analyses, suggesting that there is heterogeneity (and perhaps pleiotropy) in the genetic relationships between autoimmune disorders and OCD.

Tissue and cell type enrichment analysis revealed significant enrichment of OCD SNP-heritability in several tissues and cell types, with the strongest enrichment in excitatory neurons of hippocampus and cerebral cortex, and in dopamine D1 receptor (D1R)-positive and dopamine D2 receptor (D2R)-positive medium spiny neurons (MSNs) in the striatum. These findings are in line with traditional neural circuitry models of OCD, which focus on frontal cortical-striatal pathways[68,69]. These findings are consistent with and build on previous work linking various neuronal cell types to psychiatric and cognitive phenotypes[70].

Interestingly, frontal and anterior cingulate cortex, which were enriched in our tissue-based analyses, as well as hippocampus and striatum, which were implicated in our cell-type based analyses, are among the regions that are consistently implicated in neuroimaging studies of OCD[41,71–73]. Enrichment in MSNs in striatum is consistent with their role in the observed aberrant circuitry in OCD, where the medium spiny neurons D1 project to the globus pallidus interna and substantia nigra in the direct pathway, and the D2-type MSNs project to the globus pallidus externa in the indirect pathway[74]. However, medium spiny neurons are also enriched in major depressive disorder[75], schizophrenia[76], and intelligence[77], suggesting that the observed enrichment is not specific for OCD.

Our analyses of the shared genetic risk between OCD and other psychiatric disorders provides further insights into the etiology of OCD. In line with previous observations[38,78], OCD was significantly genetically correlated with multiple psychiatric disorders and traits. The strongest genetic correlations were observed for anxiety disorders, depression, and anorexia nervosa, all of which are highly comorbid with OCD[79]. This aligns with previous findings from cross-disorder analyses suggesting a shared genetic susceptibility among most psychiatric disorders[38,80,81]. A notable exception is our finding that risk variants for OCD are protective for alcohol dependence[82], which is at odds with epidemiological evidence strongly linking OCD and alcohol related disorders[83], but in line with a recent paper[79] reporting a lower than expected lifetime comorbidity of substance use disorders in OCD. The observed pattern of correlations with other phenotypes can be thought of as falling into two categories: compulsivity/impulsivity and rumination/worry/neuroticism. In both categories, the patterns of genetic correlations appear to follow a gradient across disorders/traits. For example, in the compulsivity/impulsivity category, strong positive correlations are seen with anorexia nervosa and Tourette syndrome, which are disorders with strong compulsive features, with less positive associations seen with ADHD, and negative correlations with alcohol dependence and risk-taking behaviors, which are all phenotypes characterized by impulsivity. A similar gradient is observed for the rumination/worry/neuroticism related phenotypes, with strong positive correlations with anxiety,

and other ruminative phenotypes such as worry, transitioning to less strong correlations with individual depression-related items.

This study marks the transition from the flat (sample building) phase of SNP discovery described for GWAS[84] (**Supplementary Fig. 20**), where few to no genome-wide significant loci are identified[10,12,51,85], to the linear phase of SNP discovery, where even relatively small increases in sample size identify additional genome-wide significant loci[18]. The strengths of the current study hence include the marked increase in the number of OCD cases and the rigorous analytic methods, including two multivariate approaches (MTAG and GenomicSEM) to control for potential overlapping study participants and to examine potential heterogeneity between the multiple ascertainment approaches. Potential weaknesses include the inability to document comorbid psychiatric disorders in the majority of cases that were not ascertained from clinical collections or electronic registries, the lack of inclusion of non-European ancestries, and the limited availability of sex-chromosome data. Due to the nature of our study, imputation references used in the different cohorts were heterogeneous and did not allow for confident analysis of rare variant associations. Future larger scale sequencing studies that are currently underway will be needed to identify associations in this allele frequency spectrum. We also note that the genetic correlation analyses are impacted by residual heterogeneity in genetic signals due to the employment of heterogeneous ascertainment strategies.

In summary, this work substantially advances the field of OCD genetics by identifying new OCD genetic risk loci and multiple credible candidate causal genes, including those expressed in brain regions and cell types previously implicated in OCD[86]. We have also shown that OCD is highly polygenic in nature, with many variants implicated not only in OCD but also in commonly comorbid disorders or traits, in particular anxiety, neuroticism, anorexia nervosa, and depression. The observation that common variants explain only a modest amount of the phenotypic variation in OCD suggests that other types of genetic variation may also contribute to the etiology of OCD. Notably, whole-exome sequencing studies have suggested that a substantial proportion of OCD cases (22%) may be influenced by rare de novo coding variants[87], especially in genes that are intolerant to loss of function[88]. Similarly, rare potentially damaging copy number variations represent part of the risk architecture for OCD[9]. These findings emphasize the need for a comprehensive exploration of the contribution of both common and rare genetic factors, as well as their interplay, to OCD risk. Finally, with the implication of the MHC complex, we provide additional evidence for potential shared genetic influences underlying both OCD and increased liability to autoimmune processes, although the directionality of those relationships remains to be definitively elucidated. In addition to continuing to increase sample sizes, future studies will require ancestrally diverse samples to further facilitate the discovery of additional OCD risk variants. Similarly, sex-specific analyses and additional clinical phenotyping will allow for the further elucidation of genetic and clinical relationships between OCD and co-occurring disorders. Finally, with the emergence of drug databases describing the relations between drugs and molecular phenotypes[89], our results may be useful for drug repurposing (i.e., identifying existing drugs targeting OCD risk genes), leading to new opportunities to find more effective treatments.

Consortium author lists and affiliations

23andMe Inc.
Chris German[268]
[268]23andMe, Inc., Sunnyvale, CA, USA.

Estonian Biobank
Andres Metspalu[98], Tõnu Esko[98], Reedik Mägi[98], Mari Nelis[98], Georgi Hudjashov[98]

Acknowledgements

We thank the research participants and employees of **all cohorts included in this study** for making this work possible**.** A list of members of the **23andMe** Research Team, **HUNT**, **CoGa**, and **MVP** that contributed to this study can be found in **Supplementary Note 1**. We thank the research participants and employees of **23andMe, Inc.** for making this work possible. The **Trøndelag Health Study (HUNT)** is a collaboration between HUNT Research Centre (Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology NTNU), Trøndelag County Council, Central Norway Regional Health Authority, and the Norwegian Institute of Public Health. The genotype quality control and imputation has been conducted by the K.G. Jebsen Center for Genetic Epidemiology, Department of Public Health and Nursing, Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology (NTNU). HUNT analyses were performed in digital labs at HUNT Cloud, HUNT Research Centre Department of Public Health and Nursing, Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology (NTNU), Trondheim, Norway. **NORDiC** is funded by NIMH R01 MH110427 (PI Crowley), NIMH R01 MH105500 (PI Crowley) and the Swedish Research Council grant # 2015–02271 (PI Mataix-Cols). NORDiC was further supported by the Swedish Research Council (grants 2012-07111 and 2018-02487), Swedish Research Council for Health, Working Life and Welfare 2018-00221 and Center for Innovative Medicine – CIMED. We are deeply grateful for the study participants contributing to the **NORDiC** research. We thank the collection team that worked to recruit them: Anders Juréus, Jessica Pege, Malin Rådström, Radja Satgunanthan-Dawoud, Milka Krestelica, and Birgitta Ohlander, as well as data manager Bozenna Iliadou. We also thank the National Quality Registry for Eating Disorders (RIKSÄT) for help with recruiting patients. We finally thank the BBMRI.se and KI Biobank at Karolinska Institutet for professional biobank service. Grant support for the **MoBa** team was also provided from RCN (273291, 262656, 248778, 223273) and the KG Jebsen Stiftelsen.). MoBa is supported by the Norwegian Ministry of Health and Care Services and the Ministry of Education and Research. We are grateful to all the participating families in Norway who take part in this on-going cohort study. The **AGDS** was primarily funded by National Health and Medical Research Council (NHMRC) of Australia grant 1086683. This work was further supported by NHMRC grants 1145645, 1078901 and 1087889. LCC is supported by a QIMR Berghofer Institute fellowship. We thank all the people who helped in the conception, implementation, beta testing, media campaign and data cleaning of the AGDS data. We specifically acknowledge Dale Nyholt for advice on using the PBS for research; Ken Kendler, Patrick Sullivan, Andrew McIntosh and Cathryn Lewis for input on the questionnaire; Lorelle Nunn, Mary Ferguson, Lucy Winkler and Natalie Garden for data and sample collection; Natalia Zmicerevska, Alissa Nichles and Candace Brennan for participant recruitment support; Jonathan Davies, Luke Lowrey and Valeriano Antonini for support with IT aspects; Vera Morgan and Ken Kirkby for help with the media campaign. We thank VIVA! Communications for their effort in promoting the study. We also acknowledge David Whiteman and Catherine Olsen from QSkin. The work done by the **EstBB** team has received funding from the European Union's Horizon 2020

Research and Innovation Programme under Grant agreement 847776 (CoMorMent). Data analysis for EstBB was carried out in part in the High-Performance Computing Center of University of Tartu. We thank participants, families and staff of primary and secondary schools who kindly contribute to this research (**Marta Ribasés**, **Metal-Cat and INSchool**). **EGOS** was supported by a grant from the Beatrice and Samuel A. Seaver Foundation to DEG. The genotyping of **HUNT** was financed by the National Institute of health (NIH), University of Michigan, The Norwegian Research council, and Central Norway Regional Health Authority and the Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology (NTNU). This research is based in part on data from the **Million Veteran Program**, Office of Research and Development, Veterans Health Administration, and was supported by awards CSP575b, I01CX001849-01, 1P1HX002375, and the National Center for PTSD Research. MVP was supported by funding from the Department of Veterans Affairs Office of Research and Development, USVA, grants CSP575B and I01CX001849, MVP-025, and the VA Cooperative Studies Program study, no. 575B; the VA National Center for PTSD Research, and the West Haven VA Mental Illness Research, Education and Clinical Center; and by NIH grant R01 AA026364 (JG). D.F.L. is supported by a Career Development Award CDA-2 from the Veterans Affairs Office of Research and Development (1IK2BX005058-01A2) and is Aimee Mann Fellow of Psychiatric Genetics. This publication does not represent the views of the Department of Veteran Affairs or the United States Government. The **EPOC** study was funded by the Deutsche Forschungsgemeinschaft (DFG; KA815/6-1 and WA731/10-1). **LifeGene** was supported by the Swedish Research Council, the Karolinska Institutet/Stockholm County Council research grants, AFA Insurance and the Torsten and Ragnar Söderbergs Foundation. **GENOS** was supported by the German Research Fundation (GR 1912/1-1). The OCD Collaborative Genetics Association Study (**OCGAS**) is a collaborative research study and was funded by the following NIMH Grant Numbers: MH071507, MH079489, MH079487, MH079488 and MH079494. This work (**OCGAS and IOCDF**) is supported by the Netherlands Organization for Scientific Research-Gravitation project 'BRAINSCAPES: a Roadmap from Neurogenetics to Neurobiology' (024.004.012) and the European Research Council advanced grant 'From GWAS to Function' (ERC- 2018-ADG 834057). The **OCGAS and IOCDF** samples are supported through NIMH Grant Numbers: MH071507 (G.N.), MH079489 (D.A.G.), MH079487 (J.M.), MH079488 (A.F.), and MH079494 (J.K.). The **iPSYCH** team was supported by grants from the Lundbeck Foundation (R102-A9118, R155-2014-1724, and R248-2017-2003), NIH/NIMH (1R01MH124851-01 to A.D.B.) and the Universities and University Hospitals of Aarhus and Copenhagen. The Danish National Biobank resource was supported by the Novo Nordisk Foundation. High-performance computer capacity for handling and statistical analysis of iPSYCH data on the GenomeDK HPC facility was provided by the Center for Genomics and Personalized Medicine and the Centre for Integrative Sequencing, iSEQ, Aarhus University, Denmark (grant to A.D.B.). **Anders D. Borglum** was also supported by the EU's HORIZON-HLTH-2021-STAYHLTH-01programme, project number 101057385: Risk and Resilience in Developmental Diversity and Mental Health (R2D2-MH). **Mental-Cat and INSchool** was supported by the Agència de Gestió d'Ajuts Universitaris i de Recerca (AGAUR, 2017SGR-1461, 2021SGR-00840), the Instituto de Salud Carlos III (PI20/00041, PI23/00404 and PI23/00026), the European Regional Development Fund (ERDF); the ECNP Network 'ADHD across the Lifespan'; "la Marató de TV3" (202228-30 and 202228-31). **BioVU**: CTSA (SD, Vanderbilt Resources) was supported by the National Center for Research Resources, Grant UL1 RR024975-01, and is now at the National Center for Advancing Translational Sciences, Grant 2 UL1 TR000445-06. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.The dataset(s) used for the analyses (BioVU) described were obtained from Vanderbilt University Medical Center's BioVU which is supported by numerous sources: institutional funding, private agencies, and federal grants. These include the NIH funded Shared Instrumentation Grant S10RR025141; and CTSA grants UL1TR002243, UL1TR000445, and UL1RR024975. Genomic data are also supported by

investigator-led projects that include U01HG004798, R01NS032830, RC2GM092618, P50GM115305, U01HG006378, U19HL065962, R01HD074711; and additional funding sources listed at https://victr.vumc.org/biovu-funding/. **Silvia Alemany** acknowledges Miguel Servet contract (CP22/00026) awarded by the Instituto de Salud Carlos III and co-funded by the European Union Found: Fondo Social Europeo Plus, FSE +.HYPERGENES and InterOmics cohorts provided controls of Italian origin for the present study. **Judit Cabana-Dominguez** acknowledges her contract from the Network Center for Biomedical Research (CIBER). **Richard Delorme** acknowledges the Clinical Investigation Centre, Robert Debré Hospital. INSERM @ APHP granted the study. **Bengt Fundin** acknowledges The Anorexia Nervosa Genetics Initiative (ANGI), an initiative of the Klarman Family Foundation. **Jan Haavik** acknowledges the Trond Mohn Foundation, Bergen, Norway. **Christine Lochner** acknowledges the South African Medical Research Council and the National Research Foundation for their support. **Thomas V Fernandez:** Research reported in this publication was supported by the National Institute of Mental Health of the National Institutes of Health under Award Number R01MH114927. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. **Nicholas G. Martin** has received funding from a project grant from Australian NHMRC. The Research Council of Norway supported **H. Ask, A. Havdahl** and **T. Reichborn-Kjennerud** (274611). A. Havdahl was also supported by South East Norway Health Authority (2020022). **Zachary F. Gerring** is supported by an Australian NHMRC EL1 Investigator Grant (2034743) and NIH/NIA Grant AG068026. **Marco Galimberti** received support from the following grants (Joel Gelernter): CSP575b, I01CX001849-01, 1P1HX002375, National Center for PTSD Research, 5R01DA054869-01. **Abdel Abdellaoui** was supported by the Foundation Volksbond Rotterdam. **Tim Bigdeli** is supported by NIMH grant 7R01MH103657 (GPC-OCD). **Jonathan Coleman:** This study represents independent research partly funded by the National Institute for Health Research (NIHR) Maudsley Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care. **Christina Barlassina** was supported by grant EU FP7-HEALTH-2007-A-201550, and grant MIUR-CNR PB05. **Enda Byrne** was supported by the NHMRC Project Grant 1145645; University of Queensland Health Research Accelerator Program (HeRA). **Carolina Cappi** was supported by grant K99MH128540-01A1. **Valentina Ciullo** was supported by the Italian Ministry of Health grant RC-18-19-20-21/A. **Marco A. Grados** was supported by NIMH K23 MH066284. **Jan Haavik** was supported by Stiftelsen KG Jebsen (SKGJ MED-02). **Kristen Hagen** was supported by the Trond Mohn Foundation. **Elinor K. Karlsson** was supported by NIH R21 MH109938. **Paul S. Nestadt** was supported by R01MH071507. **Fabrizio and Federica Piras** are supported by the Italian Ministry of Health RC18-19-20-21/A grant. This work was in part supported by the German Research Foundation (DFG) grants: [RA1971/8-1], [RA1971/7-1]; and by the Bundesministerium für Bildung und Forschung (BMBF) grant: 01ED2007A to **Alfredo Ramirez. Stephan Ripke** was supported by research grant 1R01MH124873-01. **Maria Soler Artigas** was supported by The Instituto de Salud Carlos III (P19/01224, PI22/00464 and CP22/00128) and the European Regional Development Fund (ERDF). **Arpana Agrawal** was supported by grant U10AA008401. **Pino Alonso** was supported by the Spanish Ministry of Science, Innovation and Universities (ISCIII PI22/00752) and Fundació La Marató 202201-30. **Cynthia M. Bulik** was supported by R01 MH124871 (Sullivan/Bulik) PGC4. **Howard Edenberg** was supported by grant U10AA008401. **Dan A. Geller** was supported by NIMH (OCGAS and OCGS). **Gregory L. Hanna** was supported by the National Institute of Mental Health (R01 MH58376), National Institute of Mental Health (K20 MH01065), National Institute of Mental Health (R01 MH101493), National Institute of Mental Health (R01 MH085321) **Norbert Kathmann** has received funding from Deutsche Forschungsgemeinschaft (DFG) KA815/6-1. **Sarah E. Medland** is supported by an Australian NHMRC Investigator Grant (APP1172917). **Benjamin M. Neale** is funded by grant R01MH124851. **Michele Pato and Carlos Pato** have received support from

R01MH103657 and R01MH079494 from the National Institutes of Mental Health (NIMH) and the Della Martin Foundation, Los Angeles CA. **John Piacentini** has received support through the National Institute of Mental Health: R01MH50214: Collaborative OCD Genetics Study (G. Nestadt, PI; J. McCracken, UCLA PI). **Margaret A. Richter** was supported by funding from the Canadian Institutes for Health Research and the Ontario Mental Health Foundation. **David R. Rosenberg** was supported by NIMH R01MH059299. **Jack F. Samuels** was supported by NIMH Grant Number: MH071507. **Gianfranco Spalletta** is supported by the Italian Ministry of Health RC18-19-20-21/A grant. **Eric A. Storch** collected data as part of the following NIH grant: 1R01MH093381. **Ole A. Andreassen** (MoBa) has received grant support from RCN (324499,273291,262656,248778,223273), KG Jebsen Stiftelsen, NordForsk #164218. **Jaakko Kaprio** has been supported by the Academy of Finland (grant 336823). **Paul D. Arnold** is supported by the Alberta Innovates Translational Health Chair in Child and Youth Mental Health. **Dorothy E. Grice** is supported by the grant MH124679-01. **James A. Knowles** is supported through the following grants: R01MH103657 and R01MH079494 from the National Institutes of Mental Health (NIMH) and the Della Martin Foundation, Los Angeles CA. **Karin J. H. Verweij** is supported by the Foundation Volksbond Rotterdam. **Lea K. Davis** was supported by grants from the National Institutes of Health including R01NS102371, R01MH113362, R01MH118223, R01NS105746, and R56MH120736. J.S. was supported by a NIH training grant in Human Genetics, 2T32GM080178. **James Crowley** was supported by NIH grants R01MH105500 and R01MH110427. **Murray B. Stein** has been funded by the Veterans Affairs Administration (United States VA).

Author Contributions

J.M. Scharf, M.B.S., J. Gelernter, C.A.M., E.M.D. and M. Mattheisen designed the study. N.I.S., Z.F.G., M.G., D.Y. and M.W.H. conducted data analysis. N.I.S., Z.F.G., M.G., D.Y., M.W.H., A. Abdellaoui, C.R.-F., J.M. Sealock, T.B., J.R.C., B. Mahjani, J.G.T., K.B., C.L.B., J.J.L., G.Z., S.A., C.A., K.D.A., J.B., N.B., C.B., J.B.N., O.J.B., D.B., M.H.B., S.B., R.B., M.B., B.P.B., H.B., J.D.B., J.B.-G., E.M.B., J.C.-D., B.C., A. Camarena, C.C., A. Carracedo, M.C., M.C.C., V.C., E.H.C., J.C., B.A.C., E.J.D.S., R.D., S.D., J.A.E., X.E., M.J.F., B.T.F., L.G., C.G., F.S.G., M.A.G., J. Grove, W.G., J.H., K. Hagen, K. Harrington, A.H., K.D.H., A.G.H., D.H., C.H., M.J., E.J., E.K.K., K. Kelley, J. Klawohn, J.E.K., K. Krebs, C. Lange, N.L., D. Levey, K.L.-T., F.M., B. Maher, B. Mathes, E.M., N.M., N.C.M., S.M., E.C.M., M. Mulhem, P.S.N., E.L.N., K.S.O., L.O., O.T.O., T.P., N.L.P., Fabrizio Piras, Federica Piras, S.P., R.R., A. Ramirez, S. Rauch, A. Reichenberg, M.A. Riddle, S. Ripke, M.C.R., A.S.S., M.A.S., A.H.S., L.G.S., J.S., M.S.A., L.F.T., E.T., H.V., N.v.K., J.V.-V., N.N.V., C.P.W., Y.W., J.R.W., B.S.W., Y.Y., H.Z., A. Agrawal, P.A., G. Berberich, K.K.B., C.M.B., D.C., D.D., V.E., H.E., P.F., T.V.F., A.J.F., J.M.G., D.A.G., H.J.G., B.D.G., G.L.H., I.B.H., D.M.H., N.K., J. Kennedy, D. Lai, M. Landén, S.L.H., M. Leboyer, C. Lochner, J.T.M., S.E.M., P.B.M., B.M.N., H.N., M.N., M.P., C. Pato, D.L.P., J.P., C. Pittenger, D.P., J.A.R.-Q., S.A.R., M.A. Richter, D.R.R., S. Ruhrmann, J.F.S., S.S., P.S., G.S., D.J. Stein, S.E.S., E.A.S., B.E.S., M.T., T.W., O.A.A., A.D.B., S.W., K. Hveem, B.K.H., C.R., N.G.M., L.M., O.M., T.R.-K., M.R., G.K., D.M.-C., K.D., E.G., M.W., J.-A.Z., G. Breen, G.N., J. Kaprio, P.D.A., D.E.G., J.A.K., H.A., K.J.V., L.K.D., D.J. Smit, J.J.C., J.M. Scharf, M.B.S., J. Gelernter, C.A.M., E.M.D. and M. Mattheisen provided samples and/or processed individual cohort data. N.I.S., Z.F.G., M.G., D.Y., M.W.H., J.M. Scharf, M.B.S., J. Gelernter, C.A.M., E.M.D. and M. Mattheisen wrote the paper and formed the core revision group. J.M. Scharf, M.B.S., J. Gelernter, C.A.M., E.M.D. and M. Mattheisen supervised and directed the study. All authors discussed the results and approved the final version of the manuscript.

Competing interests

Chris German is employed by and holds stock or stock options in 23andMe, Inc. **Erika L. Nurmi** is on the Scientific Advisory Board for Myriad Genetics and Medical Advisory Board for Tourette Association of America and received Clinical trial funding from Emalex and Octapharma Pharmaceuticals. **Jeremy Veenstra-VanderWeele** has served on advisory boards or consulted with Roche, Novartis, and SynapDx; received research funding from Roche, Novartis, SynapDx, Seaside Therapeutics, Forest, Janssen, Acadia, Yamo, and MapLight; received stipends for editorial work from Wiley and Springer. **Jens R. Wendland** is a current employee and shareholder of Takeda Pharmaceuticals and a past employee and shareholder of F. Hoffmann-La Roche, Pfizer and Nestle Health Science. **Cynthia M. Bulik** reports: Pearson (author, royalty recipient). **Peter Falkai** reports no conflict of interest regarding this study and reports to have received financial support and Advisory Board: Richter, Recordati, Boehringer-Ingelheim, Otsuka, Janssen and Lundbeck. **Hans J. Grabe** has received travel grants and speakers honoraria from Fresenius Medical Care, Neuraxpharm, Servier and Janssen Cilag as well as research funding from Fresenius Medical Care. **Ian B. Hickie** is the Co-Director, Health and Policy at the Brain and Mind Centre (BMC) University of Sydney, Australia. The BMC operates an early-intervention youth services at Camperdown under contract to headspace. Professor Hickie has previously led community-based and pharmaceutical industry-supported (Wyeth, Eli Lily, Servier, Pfizer, AstraZeneca, Janssen Cilag) projects focused on the identification and better management of anxiety and depression. He is the Chief Scientific Advisor to, and a 3.2% equity shareholder in, InnoWell Pty Ltd which aims to transform mental health services through the use of innovative technologies. **Benjamin M. Neale** is a member of the scientific advisory board at Deep Genomics and Neumora. **Christopher Pittenger** consults and/or receives research support from Biohaven Pharmaceuticals, Freedom Biosciences, Ceruvia Lifesciences, Transcend Therapeutics, UCB BioPharma, and F-Prime Capital Partners. He owns equity in Alco Therapeutics. These relationships are not related to the current work. **Dan J. Stein** has received consultancy honoraria from Discovery Vitality, Johnson & Johnson, Kanna, L'Oreal, Lundbeck, Orion, Sanofi, Servier, Takeda and Vistagen. **Eric A. Storch** reports receiving research funding to his institution from the Ream Foundation, International OCD Foundation, and NIH. He was formerly a consultant for Brainsway and Biohaven Pharmaceuticals in the past 12 months. He owns stock less than $5000 in NView/Proem for distribution related to the YBOCS scales. He receives book royalties from Elsevier, Wiley, Oxford, American Psychological Association, Guildford, Springer, Routledge, and Jessica Kingsley. **Ole A. Andreasson** reports to be a consultant to Cortechs.ai, Precision Health AS, speakers honorarium from Otsuka, Lundbeck, Sunovion, Janssen. **Anders D. Børglum** has received a speaker fee from Lundbeck. **David Mataix-Cols** receives royalties for contributing articles to UpToDate, Wolters Kluwer Health, and personal fees for editorial work from Elsevier, all unrelated to the current work. **Murray B. Stein** has in the past 3 years received consulting income from Acadia Pharmaceuticals, BigHealth, Biogen, Bionomics, Boehringer Ingelheim, Clexio, Eisai, EmpowerPharm, Engrail Therapeutics, Janssen, Jazz Pharmaceuticals, NeuroTrauma Sciences, Otsuka, PureTech Health, Sage Therapeutics, Sumitomo Pharma, and Roche/Genentech. Dr. Stein has stock options in Oxeia Biopharmaceuticals and EpiVario. He has been paid for his editorial work on *Depression and Anxiety* (Editor-in-Chief), *Biological Psychiatry* (Deputy Editor), and UpToDate (Co-Editor-in-Chief for Psychiatry). **Joel Gelernter** is paid for editorial work by the journal *Complex Psychiatry*. Pino Alonso has received funding from Biohaven, Boston Scientific, Medtronic. All other authors report no conflicts of interest.

Tables

Table 1 | Genome-wide-significant loci associated with OCD. Shown are the lead SNP, the chromosome (CHR), base pair position on the genome (BP), *P*-value (*P*), effect estimate as an odds ratio (OR), standard error of the effect estimate (s.e.), effect allele and non-effect allele (A1/A2), frequency of A1 in cases (FRQCA) and in controls (FRQCO), imputation quality score (INFO), number of genes in a region of 6.5 kb around the SNP (*n* genes), and a curated list of phenotypes that also showed a genome-wide significant association with this SNP (in one or more of the following four databases: causalDB[90], GenomeAtlas[52], NHGRI-EBI GWAS catalog[91], IEU Open GWAS project[92]. If fewer than four traits are significant across all four databases, all four traits are shown. If more than five traits are significant across the databases, neuropsychiatric traits are prioritized (closely related traits are summarized into one trait category). Previously identified GWAS hits for OCD (or SNPs in high LD with a previously identified SNP) are in italics. For a full list of associations in the four databases, see **Supplementary Table 18a-**d. A more detailed list of the significant loci can be found in **Supplementary Table 2**. Abbreviations in the last column are: schizophrenia (SCZ), bipolar disorder (BP), major depressive disorder (MDD), and body-mass index (BMI).

SNP	Position	*P*	OR	s.e.	A1/A2	FRQCA	FRQCO	INFO	*n* genes	Selection of other significant traits
rs78587207	11q12.1	5.28×10^{-12}	0.9522	0.0071	T/G	0.681	0.677	0.987	11	SCZ, well-being, neuroticism, educational attainment
rs13262595	8q24.3	1.31×10^{-11}	0.9566	0.0066	G/A	0.552	0.561	0.989	2	SCZ, well-being, neuroticism, educational attainment
rs4990036	6p21.33	1.45×10^{-11}	0.9299	0.0108	T/C	0.119	0.123	0.985	118	SCZ, depression, blood cell count, lung function
rs10877425	12q14.1	1.62×10^{-11}	0.9526	0.0072	G/A	0.513	0.516	0.985	0	-
rs7626445	3p21.31	1.74×10^{-11}	0.9551	0.0068	T/C	0.647	0.654	0.994	32	neuroticism, smoking, blood cell count, height
rs2564930	3p21.1	3.41×10^{-11}	0.9546	0.007	T/C	0.339	0.345	0.988	12	SCZ, neuroticism, blood cell count, BMI
rs4702	15q26.1	9.07×10^{-10}	1.0414	0.0066	G/A	0.455	0.449	0.984	5	SCZ, BP, MDD, risk taking behavior
rs35518360	4q24	1.39×10^{-9}	1.0757	0.0121	T/A	0.0828	0.0756	0.947	1	SCZ, neuroticism, educational attainment, BMI
rs4904738	14q21.1	1.48×10^{-9}	0.9605	0.0067	T/C	0.558	0.553	0.984	1	MDD
rs254779	5q14.3	1.53×10^{-9}	0.9606	0.0067	T/C	0.419	0.421	0.988	4	educational attainment, cognition, fat mass
rs2198140	8p12	2.18×10^{-9}	0.9590	0.007	T/C	0.496	0.513	0.979	1	cognition
rs12516488	5p12	3.79×10^{-9}	1.0531	0.0088	G/A	0.825	0.826	0.994	1	neuroticism, age at first birth, age at first sexual intercourse
rs3899258	5q11.1	4.94×10^{-9}	1.0509	0.0085	G/A	0.782	0.792	0.989	2	-
rs3027160	17p13.1	5.35×10^{-9}	1.0497	0.0083	T/C	0.775	0.782	0.996	19	sleep, height
rs203768	2q33.1	6.14×10^{-9}	0.9513	0.0086	T/C	0.824	0.817	0.992	5	SCZ
rs11263940	1p34.3	7.23×10^{-9}	0.9578	0.0074	T/C	0.689	0.69	0.991	0	neuroticism, well-being
rs67839857	5q14.3	7.63×10^{-9}	1.0423	0.0072	G/A	0.692	0.691	0.994	0	-
rs1555466	20p11.23	8.42×10^{-9}	1.0490	0.0083	T/C	0.218	0.224	0.996	0	ease of skin tanning
rs9886111	7q21.13	8.59×10^{-9}	0.9598	0.0071	G/C	0.701	0.711	0.992	2	-
rs9287859	2q24.3	9.83×10^{-9}	0.9595	0.0072	G/A	0.39	0.39	0.994	1	-
rs2087319	4q12	1.59×10^{-8}	0.9579	0.0076	C/A	0.74	0.744	0.968	6	height, blood pressure
rs11125759	2p16.1	1.79×10^{-8}	0.9690	0.0071	G/A	0.569	0.556	0.991	1	BMI, sleep
rs6474628	9p23	1.89×10^{-8}	1.0380	0.0066	T/G	0.579	0.585	0.999	0	-
rs11768238	7q33	2.28×10^{-8}	0.9601	0.0073	G/A	0.661	0.661	0.998	2	educational attainment, age at first sexual intercourse
rs9479138	6q25.1	2.41×10^{-8}	1.0397	0.007	T/G	0.339	0.34	0.975	1	educational attainment, age at first sexual intercourse, age at first birth, lung function
rs1567288	4q22.3	3.80×10^{-8}	0.9643	0.0066	G/A	0.548	0.55	0.981	1	-
rs4831130	3q13.31	3.93×10^{-8}	1.0427	0.0076	T/G	0.753	0.74	0.984	2	-

rs17718444	3p13	4.25×10^{-8}	0.9622	0.007	T/C	0.323	0.32	0.984	2	educational attainment, lung function, use of sun/UV protection
rs6660196	1q24.1	4.86×10^{-8}	1.0403	0.0072	T/G	0.638	0.642	0.989	1	blood cell count
rs4931	20q13.12	5.00×10^{-8}	0.9609	0.0073	C/A	0.278	0.279	0.993	5	blood cell count, height, BMI

Figure legends

Figure 1 | Manhattan plot of OCD GWAS meta-analysis. The *y*-axis represents $-\log_{10}$ *P*-values (two-sided, not adjusted for multiple testing) for the association of variants with OCD using an inverse-variance weighted fixed effects model (n_{cases} = 53,660 and $n_{controls}$ = 2,044,417). The *x*-axis shows chromosomes 1 to 22. The horizontal red line represents the threshold for genome-wide significance ($P = 5 \times 10^{-8}$). Index variants of genome-wide significant loci are highlighted as a green diamond.

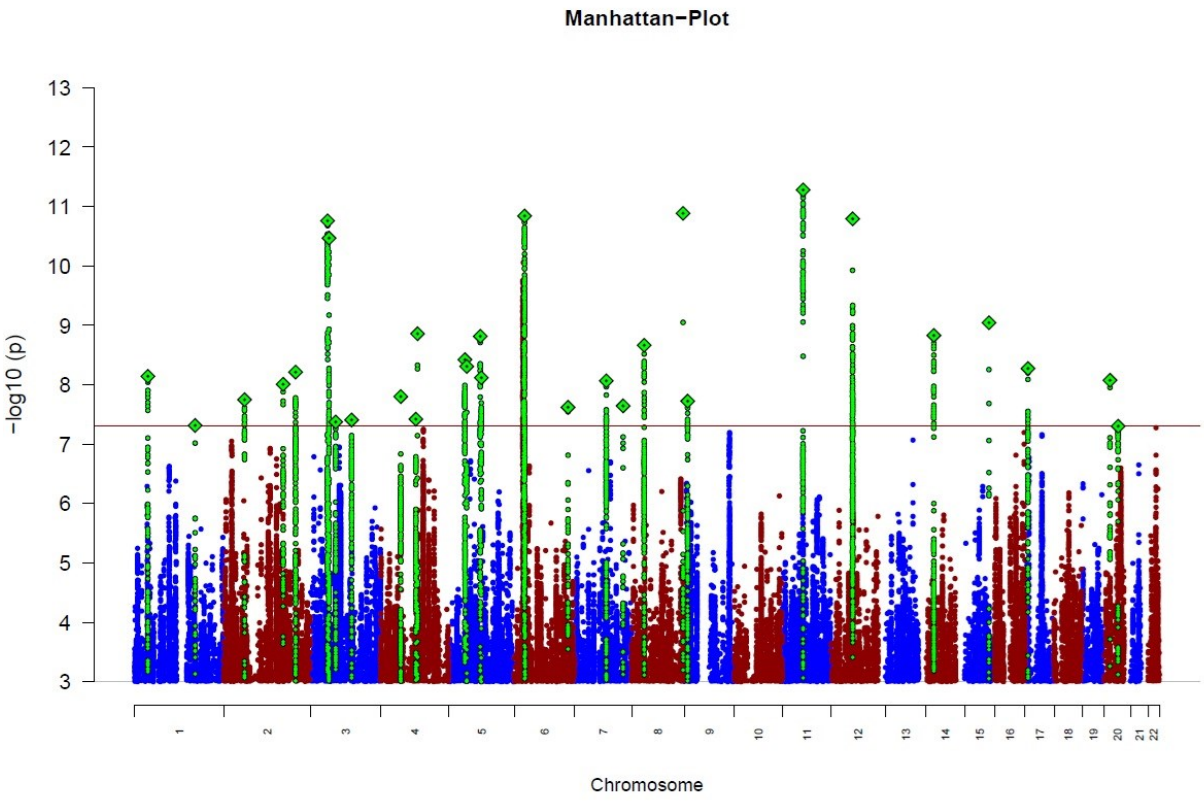


Figure 2 | Gene-based, tissue, and cell type enrichment analyses. a, List of 25 genes that were implicated in at least two of the five different gene-based tests (significance indicated by grey dots) and passed the TWAS colocalization and/or SMR-HEIDI filters (significance indicated by orange dots). Conditionally independent genes within each locus are indicated by blue dots. **b**, Enrichment of OCD GWAS signal in human brain-related tissues from GTEx (v8). No significant enrichment was observed in the peripheral tissues (not included in the figure). The horizontal bar size represents the significance of the enrichment measured using the MAGMA gene set enrichment test or partitioned LDSC. **c**, Top 20 groups of brain cell types (n = 35 total tested) enriched with OCD GWAS signal using MAGMA. Dots represent $-\log_{10}$ P-values from MAGMA gene set enrichment tests of individual neuronal cell types from Zeisel et al.[30]. Vertical crosses represent the mean $-\log_{10}$ P-value observed for each brain cell type group. Blue crosses represent a significant enrichment of OCD GWAS signals (false discovery rate across 35 groups, FDR < 0.05), while pink crosses indicate non-significant enrichment. Grey points represent the association ($-\log_{10}$ P-value) for each single cell cluster ("level 5" analysis defined by Zeisel et al.[30]) in a given cell type (e.g., excitatory neurons, cerebral cortex).

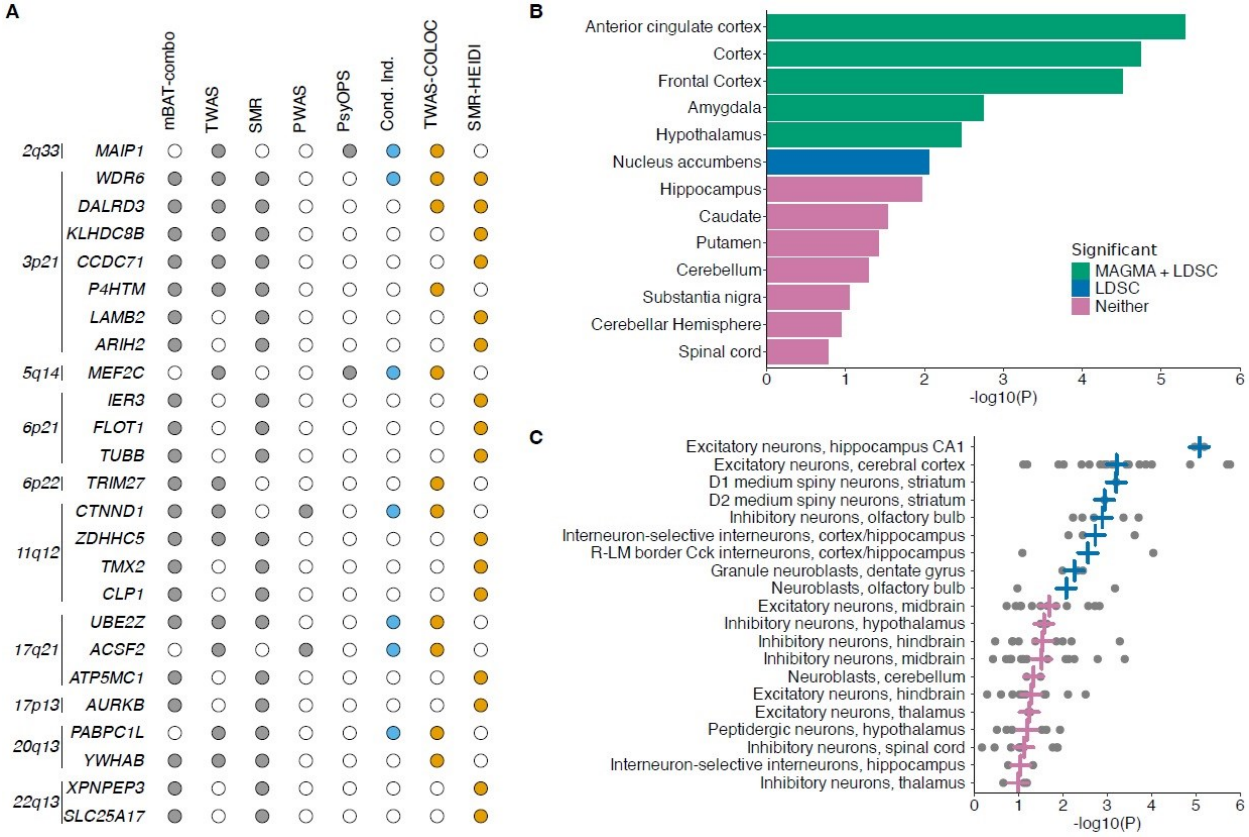
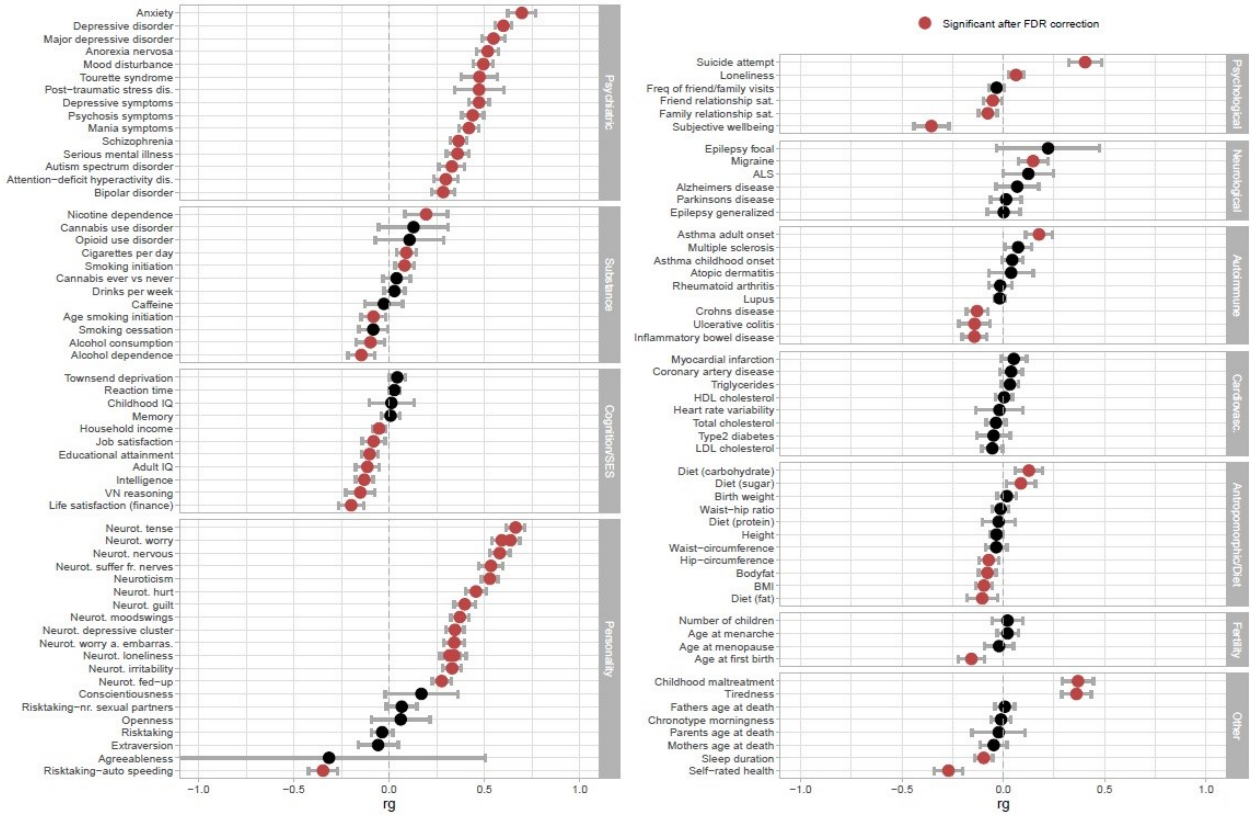


Figure 3 | Genetic correlations (r_G) between OCD and 112 phenotypes. This includes psychiatric, substance use, cognition/socioeconomic status (SES), personality, psychological, neurological, autoimmune, cardiovascular, anthropomorphic/diet, fertility, and other phenotypes. References and sample sizes of the corresponding summary statistics of the GWAS studies can be found in **Supplementary Table 18**. The OCD summary statistics are of the main meta-analysis (n_{cases} = 53,660 and $n_{controls}$ = 2,044,417). Error bars represent the 95% confidence intervals for the genetic correlation estimates (r_G). Red circles indicate significant associations with a *P*-value adjusted for multiple testing with the Benjamini-Hochberg procedure to control the FDR (< 0.05). Black circles indicate associations that are not significant.



References

1. Fawcett, E. J., Power, H. & Fawcett, J. M. Women are at greater risk of OCD than men: a meta-analytic review of OCD prevalence worldwide. *J. Clin. Psychiatry* **81**, 19r13085 (2020).
2. World Health Organization. *The Global Burden of Disease: 2004 Update*. Geneva: WHO Press (2008).
3. Fernández de la Cruz, L. *et al.* Suicide in obsessive–compulsive disorder: a population-based study of 36 788 Swedish patients. *Mol. Psychiatry* **22**, 1626–1632 (2017).
4. Meier, S. M., Mattheisen, M., Mors, O., Schendel, D. E., Mortensen, P. B. & Plessen, K. J. Mortality among persons with obsessive-compulsive disorder in Denmark. *JAMA Psychiatry* **73**, 268–274 (2016).
5. Blanco-Vieira, T., Radua, J., Marcelino, L., Bloch, M., Mataix-Cols, D. & do Rosário, M. C. The genetic epidemiology of obsessive-compulsive disorder: a systematic review and meta-analysis. *Trans.l Psychiatry* **13**, 230 (2023).
6. Burton, C. L. *et al.* Heritability of obsessive–compulsive trait dimensions in youth from the general population. *Transl. Psychiatry* **8**, 191 (2018).
7. Pauls, D. L. The genetics of obsessive compulsive disorder: a review of the evidence. *Am. J. Med. Genet. C Semin. Med. Genet.* **148**, 133–139 (2008).
8. van Grootheest, D. S., Cath, D. C., Beekman, A. T. & Boomsma, D. I. Twin studies on obsessive–compulsive disorder: a review. *Twin Res. Hum. Genet.* **8**, 450–458 (2005).
9. Mahjani, B. *et al.* The genetic architecture of obsessive-compulsive disorder: contribution of liability to OCD from alleles across the frequency spectrum. *Am. J. Psychiatry* **179**, 216–225 (2022).
10. International Obsessive Compulsive Disorder Foundation Genetics Collaborative (IOCDF-GC) and OCD Collaborative Genetics Association Studies (OCGAS). Revealing the complex genetic architecture of obsessive-compulsive disorder using meta-analysis. *Mol. Psychiatry* **23**, 1181–1188 (2018).
11. Davis, L. K. *et al.* Partitioning the heritability of Tourette syndrome and obsessive compulsive disorder reveals differences in genetic architecture. *PLoS Genet.* 9, e1003864 (2013).
12. Strom, N. I. *et al.* Genome-wide association study identifies new locus associated with OCD. *medRxiv* (https://doi.org/10.1101/2021.10.13.21261078).
13. Strom, N. I. *et al.* Genome-wide association study identifies new loci associated with OCD. *medRxiv* (https://doi.org/10.1101/2024.03.06.24303776).
14. Yang, J., Lee, S. H., Goddard, M. E. & Visscher, P. M. GCTA: a tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* **88**, 76–82 (2011).
15. Holland, D. *et al.* Beyond SNP heritability: polygenicity and discoverability of phenotypes estimated with a univariate Gaussian mixture model. *PLoS Genet.* **16**, e1008612 (2020).
16. Grotzinger, A. D. *et al.* Genomic structural equation modelling provides insights into the multivariate genetic architecture of complex traits. *Nat. Hum. Behav.* **3**, 513–525 (2019).
17. Wray, N. R. *et al.* Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nat. Genet.* **50**, 668–681 (2018).
18. Strom, N. I. *et al.* Genome-wide association study of major anxiety disorders in 122,341 European-ancestry cases identifies 58 loci and highlights GABAergic signaling. *medRxiv* (https://doi.org/10.1101/2024.07.03.24309466).
19. Li, A. *et al.* mBAT-combo: a more powerful test to detect gene-trait associations from GWAS data. *Am. J. Hum. Genet.* **110**, 30–43 (2023).
20. Gusev, A. *et al.* Integrative approaches for large-scale transcriptome-wide association studies. *Nat. Genet.* **48**, 245–252 (2016).
21. Gandal, M. J. *et al.* Shared molecular neuropathology across major psychiatric disorders parallels polygenic overlap. *Science* **359**, 693–697 (2018).
22. Zhu, Z. *et al.* Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nat. Genet.* **48**, 481–487 (2016).
23. Võsa, U. *et al.* Large-scale *cis*- and *trans*-eQTL analyses identify thousands of genetic loci and polygenic scores that regulate blood gene expression. *Nat. Genet.* **53**, 1300–1310 (2021).
24. Qi, T. *et al.* Identifying gene targets for brain-related traits using transcriptomic and methylomic data from blood. *Nat. Commun.* **9**, 2282 (2018).
25. Wingo, T. S. *et al.* Brain proteome-wide association study implicates novel proteins in depression pathogenesis. *Nat. Neurosci.* **24**, 810–817 (2021).
26. Wainberg, M., Merico, D., Keller, M. C., Fauman, E. B. & Tripathy, S. J. Predicting causal genes from psychiatric genome-wide association studies using high-level etiological knowledge. *Mol. Psychiatry* **27**, 3095–3106 (2022).
27. Giambartolomei, C. *et al.* Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. *PLoS Genet.* **10**, e1004383 (2014).

28. Wallace, C. Eliciting priors and relaxing the single causal variant assumption in colocalisation analyses. *PLoS Genet.* **16**, e1008720 (2020).
29. Bryois, J. *et al.* Genetic identification of cell types underlying brain complex traits yields insights into the etiology of Parkinson's disease. *Nat. Genet.* **52**, 482–493 (2020).
30. Zeisel, A. *et al.* Molecular architecture of the mouse nervous system. *Cell* **174**, 999–1014.e22 (2018).
31. Bulik-Sullivan, B. K. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* **47**, 291–295 (2015).
32. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Statist. Soc. B* **57**, 289–300 (1995).
33. Demontis, D. *et al.* Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. *Nat. Genet.* **51**, 63–75 (2019).
34. Demontis, D. *et al.* Genome-wide analyses of ADHD identify 27 risk loci, refine the genetic architecture and implicate several cognitive domains. *Nat. Genet.* **55**, 198–208 (2023).
35. Hall, L. S. *et al.* Genome-wide meta-analyses of stratified depression in Generation Scotland and UK Biobank. *Transl. Psychiatry* **8**, 9 (2018).
36. Als, T. D. *et al.* Depression pathophysiology, risk prediction of recurrence and comorbid psychiatric disorders using genome-wide analyses. *Nat. Med.* **29**, 1832–1844 (2023).
37. Cai, N. *et al.* Minimal phenotyping yields genome-wide association signals of low specificity for major depression. *Nat. Genet.* **52**, 437–447 (2020).
38. Derks, E. M., Thorp, J. G. & Gerring, Z. F. Ten challenges for clinical translation in psychiatric genetics. *Nat. Genet.* **54**, 1457–1465 (2022).
39. Pardiñas, A. F. *et al.* Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. *Nat. Genet.* **50**, 381–389 (2018).
40. Baselmans, B. M. L. *et al.* Multivariate genome-wide analyses of the well-being spectrum. *Nat. Genet.* **51**, 445–451 (2019).
41. van den Heuvel, O. A. *et al.* (Brain circuitry of compulsivity. *Eur. Neuropsychopharmacol.* **26**, 810–827 (2016).
42. Shephard, E. *et al.* Toward a neurocircuit-based taxonomy to guide treatment of obsessive-compulsive disorder. *Mol. Psychiatry* **26**, 4583–4604 (2021).
43. Davis, M. A., Ireton, R. C. & Reynolds, A. B. A core function for p120-catenin in cadherin turnover. *J. Cell Biol.* **163**, 525–534 (2003).
44. Yanagisawa, M. *et al.* A p120 catenin isoform switch affects Rho activity, induces tumor cell invasion, and predicts metastatic disease. *J. Biol. Chem.* **283**, 18344–18354 (2008).
45. Daniel, J. M. & Reynolds, A. B. The catenin p120(ctn) interacts with Kaiso, a novel BTB/POZ domain zinc finger transcription factor. *Mol. Cell. Biol.* **19**, 3614–3623 (1999).
46. Ishiyama, N. *et al.* Dynamic and static interactions between p120 catenin and E-cadherin regulate the stability of cell-cell adhesion. *Cell* **141**, 117–128 (2010).
47. Schaffer, A. E. *et al. CLP1* founder mutation links tRNA splicing and maturation to cerebellar development and neurodegeneration. *Cell* **157**, 651–663 (2014).
48. Vandervore, L. V. *et al.* TMX2 is a crucial regulator of cellular redox state, and its dysfunction causes severe brain developmental abnormalities. *Am. J. Hum. Genet.* **105**, 1126–1147 (2019).
49. Zhang, Y., Li, F., Fu, K., Liu, X., Lien, I.-C. & Li, H. Potential role of S-palmitoylation in cancer stem cells of lung adenocarcinoma. *Front. Cell Dev. Biol.* **9**, 734897 (2021).
50. Hindley, G. *et al.* Charting the landscape of genetic overlap between mental disorders and related traits beyond genetic correlation. *Am. J. Psychiatry* **179**, 833–843 (2022).
51. Mattheisen, M. *et al.* Genome-wide association study in obsessive-compulsive disorder: results from the OCGAS. *Mol. Psychiatry* **20**, 337–344 (2015).
52. Watanabe, K. *et al.* A global overview of pleiotropy and genetic architecture in complex traits. *Nat. Genet.* **51**, 1339–1348 (2019).
53. Nagel, M., Watanabe, K., Stringer, S., Posthuma, D. & van der Sluis, S. Item-level analyses reveal genetic heterogeneity in neuroticism. *Nat. Commun.* **9**, 905 (2018).
54. Xie, X., Wang, Z. & Chen, Y. Association of LKB1 with a WD-repeat protein WDR6 is implicated in cell growth arrest and p27Kip1 induction. *Mol. Cell. Biochem.* **301**, 115–122 (2007).
55. Adams, D. M., Reay, W. R. & Cairns, M. J. Multiomic prioritisation of risk genes for anorexia nervosa. *Psychol. Med.* **53**, 6754–6762 (2023).
56. Kia, D. A. *et al.* Identification of candidate Parkinson disease genes by integrating genome-wide association study, expression, and epigenetic data sets. *JAMA Neurol.* **78**, 464–472 (2021).
57. Lentini, J. M., Alsaif, H. S., Faqeih, E., Alkuraya, F. S. & Fu, D. DALRD3 encodes a protein mutated in epileptic encephalopathy that targets arginine tRNAs for 3-methylcytosine modification. *Nat. Commun.* **11**, 2510 (2020).
58. Wu, J., Poppi, L. A. & Tischfield, M. A. Planar cell polarity and the pathogenesis of Tourette disorder: new hypotheses and perspectives. *Dev. Biol.* **489**, 14–20 (2022).

59. Willsey, A. J. *et al.* De novo coding variants are strongly associated with Tourette disorder. *Neuron* **94**, 486–499.e9 (2017).
60. Zhao, X., Wang, S., Hao, J., Zhu, P., Zhang, X. & Wu, M. A whole-exome sequencing study of Tourette disorder in a Chinese population. *DNA Cell Biol.* **39**, 63–68 (2020).
61. Debnath, M., Berk, M., Leboyer, M. & Tamouza, R. The MHC/HLA gene complex in major psychiatric disorders: emerging roles and implications. *Curr. Behav. Neurosci. Rep.* **5**, 179–188 (2018).
62. Mataix-Cols, D. *et al.* A total-population multigenerational family clustering study of autoimmune diseases in obsessive–compulsive disorder and Tourette's/chronic tic disorders. *Mol. Psychiatry* **23**, 1652–1658 (2018).
63. Tylee, D. S. *et al.* Genetic correlations among psychiatric and immune-related phenotypes based on genome-wide association data. *Am J Med Genet B Neuropsychiatr Genet* **177**, 641–657 (2018).
64. Westwell-Roper, C. *et al.* Immune-related comorbidities in childhood-onset obsessive compulsive disorder: lifetime prevalence in the Obsessive Compulsive Disorder Collaborative Genetics Association Study. *J. Child Adolesc. Psychopharmacol.* **29**, 615–624 (2019).
65. Zhang, T. *et al.* Prenatal and early childhood infections and subsequent risk of obsessive-compulsive disorder and tic disorders: a nationwide, sibling-controlled study. *Biol. Psychiatry* **93**, 1023–1030 (2023).
66. Swedo, S. E. *et al.* Pediatric autoimmune neuropsychiatric disorders associated with streptococcal infections: clinical description of the first 50 cases. *Am. J. Psychiatry* **155**, 264–271 (1998).
67. Wilbur, C. *et al.* PANDAS/PANS in childhood: controversies and evidence. *Paediatr. Child Health* **24**, 85–91 (2019).
68. Ahmari, S. E. *et al.* Repeated cortico-striatal stimulation generates persistent OCD-like behavior. *Science* **340**, 1234–1239 (2013).
69. Ahmari, S. E. & Rauch S. L. The prefrontal cortex and OCD. *Neuropsychopharmacol.* **47**, 211–224 (2022).
70. Olislagers, M., Rademaker, K., Adan, R. A. H., Lin, B. D. & Luykx, J. J. Comprehensive analyses of RNA-seq and genome-wide data point to enrichment of neuronal cell type subsets in neuropsychiatric disorders. *Mol. Psychiatry* **27**, 947–955 (2022).
71. Boedhoe, P. S. W. *et al.* Cortical abnormalities associated with pediatric and adult obsessive-compulsive disorder: findings from the ENIGMA Obsessive-Compulsive Disorder Working Group. *Am. J. Psychiatry* **175**, 453–462 (2018).
72. Bruin, W. B. *et al.* Structural neuroimaging biomarkers for obsessive-compulsive disorder in the ENIGMA-OCD consortium: medication matters. *Transl. Psychiatry* **10**, 342 (2020).
73. van den Heuvel, O. A. *et al.* An overview of the first 5 years of the ENIGMA obsessive–compulsive disorder working group: the power of worldwide collaboration. *Human Brain Mapping* **43**, 23–36 (2022).
74. Haber, S. N. Corticostriatal circuitry. *Dialogues Clin. Neurosci.* **18**, 7–21 (2016).
75. Major Depressive Disorder Working Group of the Psychiatric Genomics Consortium. Trans-ancestry genome-wide study of depression identifies 697 associations implicating cell types and pharmacotherapies. *Cell* **188**, 640–652.e9 (2025).
76. Trubetskoy, V. *et al.* Mapping genomic loci implicates genes and synaptic biology in schizophrenia. *Nature* **604**, 502–508 (2022).
77. Savage, J. E. *et al.* Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. *Nat. Genet.* **50**, 912–919 (2018).
78. Romero, C. *et al.* Exploring the genetic overlap between twelve psychiatric disorders. *Nat. Genet.* **54**, 1795–1802 (2022).
79. Sharma, E. *et al.* Comorbidities in obsessive-compulsive disorder across the lifespan: a systematic review and meta-analysis. *Front. Psychiatry* **12**, 703701 (2021).
80. Lee, P. H. *et al.* Genomic relationships, novel loci, and pleiotropic mechanisms across eight psychiatric disorders. *Cell* **179**, 1469–1482.e11 (2019).
81. Grotzinger, A. D. Shared genetic architecture across psychiatric disorders. *Psychol. Med.* **51**, 2210–2216 (2021).
82. Walters, R. K. *et al.* Transancestral GWAS of alcohol dependence reveals common genetic underpinnings with psychiatric disorders. *Nat. Neurosci.* **21**, 1656–1669 (2018).
83. Virtanen, S. *et al.* Association of obsessive-compulsive disorder and obsessive-compulsive symptoms with substance misuse in 2 longitudinal cohorts in Sweden. *JAMA Network Open* **5**, e2214779 (2022).
84. Sullivan, P. F. *et al.* Psychiatric genomics: an update and an agenda. *Am. J. Psychiatry* **175**, 15–27 (2018).
85. Stewart, S. E. *et al.* Genome-wide association study of obsessive-compulsive disorder. *Mol. Psychiatry* **18**, 788–798 (2013).
86. Piantadosi SC, McClain LL, Klei L, Wang J, Chamberlain BL, Springer SA, *et al.* (2021): Transcriptome alterations are enriched for synapse-associated genes in the striatum of subjects with obsessive-compulsive disorder [no. 1]. *Transl Psychiatry* 11: 1–11.
87. Cappi, C. *et al.* De novo damaging DNA coding mutations are associated with obsessive-compulsive disorder and overlap with Tourette's disorder and autism. *Biol. Psychiatry* **87**, 1035–1044 (2020).
88. Halvorsen, M. *et al.* Exome sequencing in obsessive-compulsive disorder reveals a burden of rare damaging coding variants. *Nat. Neurosci.* **24**, 1071–1076 (2021).

89. Knox, C. *et al.* DrugBank 6.0: the DrugBank Knowledgebase for 2024. *Nucleic Acids Res.* **52**, D1265–D1275 (2024).
90. Wang, J. *et al.* CAUSALdb: a database for disease/trait causal variants identified using summary statistics of genome-wide association studies. *Nucleic Acids Res.* **48**, D807–D816 (2020).
91. Buniello, A. *et al.* The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res.* **47**, D1005–D1012 (2019).
92. Elsworth, B. *et al.* The MRC IEU OpenGWAS data infrastructure. *bioRxiv* (https://doi.org/10.1101/2020.08.10.244293).

Methods

Ethics

All relevant ethics approvals have been obtained by the respective cohort's institutions and a list of all respective approvals can be found in **Supplementary Note 2**.

Study participants

We analyzed genomic data from 28 OCD case-control cohorts including 53,660 OCD cases and 2,044,417 controls of European ancestry. **Supplementary Table 1** provides an overview of the individual cohorts. A subset of the cases and controls have been included in previous studies[10,51,85] and preprints[12,13], as described in **Supplementary Note 2**. Among all included individuals, 323 cases were part of a parent-proband trio; in these cases, parents were used as pseudo-controls. A total of 20,427 cases met DSM-5[93] or ICD-10[94] criteria for OCD as assessed by a healthcare professional or derived from (electronic) health records, while the remaining 32,233 cases were based on self-reported OCD diagnosis (23andMe, AGDS, and parts of UKBB). Cohort-specific sample and analytic details can be found in **Supplementary Note 2**. Data collections were approved by the relevant institutional review boards at all participating sites, and all participants provided written informed consent.

Individual GWAS analyses and harmonizing of results

First, the data of each participating cohort were analyzed individually (see **Supplementary Note 2** for details). Genetic data were imputed using either the Haplotype Reference Consortium (HRC)[95] or 1000 Genomes Project Phase 3 reference panels[96]. The resulting GWAS summary statistics were then harmonized before a conjoint meta-analysis of all autosomes was conducted. Each summary statistic data set was transformed to 'daner' file format following RICOPILI[97] specifications. All variants had to meet the following criteria for inclusion: minor allele frequency (MAF) > 1% in cases and controls, imputation quality (INFO) score > 0.8 and < 1.2. If the effect measure, *P*-value or standard error (s.e.) was missing or was out of bounds (infinite), the SNP was removed. Once cleaned summary statistics were produced, all datasets were aligned to the HRC reference panel. If variants were reported on different strands, they were flipped to the orientation in the HRC-reference. Furthermore, strand-ambiguous A/T and C/G SNPs were removed if their MAF was > 0.4. In the case that A/T and C/G SNPs showed a MAF < 0.4, allele frequencies were compared to frequencies in the HRC-reference. If an allele frequency match

was found, i.e., minor alleles were the same in the summary statistics and the HRC reference, the same strand orientation was assumed. If an allele mismatch was found, i.e., the allele had a frequency > 0.5 in HRC, it was assumed that alleles were reported on different strands and alleles were flipped subsequently. Marker-names were uniformly switched to those present in the HRC reference. If a variant did not overlap with the variants in the HRC reference, it was removed.

GWAS meta-analysis

Inverse variance weighted meta-analysis was conducted on 28 European cohorts using METAL[98]. Weighting was based on standard error primarily to account for the large case/control imbalances in cohorts that used linear mixed model approaches in their primary GWAS. Heterogeneity was assessed with Cochran's *Q* statistic and I^2 statistic[99,100] (see **Supplementary Note 5** for details). The genomic control factor lambda (λ) was calculated for each individual GWAS and for the overall meta-analysis to identify residual population stratification or systematic technical artifact. GWAS summary statistics were subjected to linkage disequilibrium (LD) score regression (LDSC) analyses on high-quality common SNPs (INFO score > 0.9) to examine the LDSC intercept to distinguish polygenicity from other types of inflation, and to estimate the genetic heritability from the meta-analysis and genetic correlations between cohorts. The genomic inflation factor lambda (λ) was estimated at 1.330 with an λ_{1000} of 1.033, while the LDSC intercept was 1.0155 $(s.e. = 0.0085)$, indicating that the inflation was mostly due to polygenic signal and unlikely to be significantly confounded by population structure. The genome-wide significance threshold for the GWAS was set at a *P*-value of 5.0×10^{-8}. The 23andMe data included information on the X-chromosome; as this information was not present for all other cohorts, analysis of the X-chromosome was only conducted in this sub-cohort (see **Supplementary Note** 4 for details).

We further conducted GWAS meta-analyses on four subgroups, defined by differences in their sample ascertainment, either (i) clinical-OCD cases diagnosed by a health care professional in a clinical setting ($n_{cases} = 9{,}089, n_{controls} = 21{,}077$; including IOCDF, IOCDF_trio, EPOC, NORDiC-nor, NORDiC-swe, EGOS, OCGAS, OCGAS-ab, OCGAS-gh, OCGAS-nes, Psych_Broad, WWF, MVP, Michigan/Toronto IGS, YalePenn, Chop, CoGa), (ii) comorbid-individuals that were primarily ascertained for another comorbid psychiatric disorder ($n_{cases} = 5{,}266, n_{controls} = 43{,}760$; AGDS, IPSYCH), (iii) biobank data from large-scale biobanks or registries with ICD or DSM codes ($n_{cases} = 9{,}138, n_{controls} = 1{,}049{,}776$; BioVU, EstBB, FinnGen, HUNT, MoBa, UKBB) or (iv) 23andMe data ($n_{cases} = 30{,}167, n_{controls} = 929{,}804$). While these groups are not exclusive (e.g., diagnoses in health records were originally given in a clinical setting, or comorbid cases were also assessed in a clinical setting or derived from health records), we defined these groups by the cohort's primary characteristic. We also conducted one meta-analysis including all clinical, comorbid, and biobank subgroups, while excluding the 23andMe data, resulting in 23,493 cases and 1,114,613 controls. As 23andMe is the only consumer-based dataset, we intended to compare this dataset to all others.

Number of Trait specific causal Variants (MiXeR analysis)

We applied MiXeR v1.3[15] to quantify the polygenicity of OCD (i.e., estimate the total number of trait-influencing genetic variants). MiXeR fits a Gaussian mixture model assuming that common genetic effects on a trait are a mixture of causal variants and noncausal variants. Polygenicity is reported as the number of the causal variants that explain 90% of SNP heritability of OCD (to avoid extrapolating model parameters into the area of infinitesimally small effects).

SNP-based fine-mapping - GCTA-COJO

We performed a conditional-and-joint analysis (GCTA-COJO)[14] to identify independent signals within significant OCD loci. This approach performs a conditional and joint analysis on the basis of conditional *P*-values before calculating the joint effects of all selected SNPs. We used the stepwise model selection procedure to select independently associated SNPs. The linkage disequilibrium reference sample was created from 73,005 individuals from the QIMR Berghofer Medical Research Institute genetic epidemiology cohort. The distance assumed for complete linkage disequilibrium was 10 Mb, and we used the default *P*-value threshold of 5×10^{-8} to define a genome-wide significant hit.

Multi-trait analysis of ascertainment subgroups

We used multi-trait analysis of GWAS (MTAG)[101] to conduct multivariable GWAS analyses, reporting GWAS results for each of the ascertainment-specific sub-groups. Through this approach, we aimed to address potential concerns about heterogeneity in genetic liability for individual sub-groups following different ascertainment strategies. MTAG is a multi-trait analysis that is usually used to combine different but related traits into one meta-analysis by leveraging the shared heritability among the different traits and thereby gaining power. In this case, our aim was to generate ascertainment-specific estimates, while boosting power by leveraging the high shared heritability between the subgroups. The MTAG analysis resulted in four different GWAS summary statistics, one for each subgroup (clinical, comorbid, biobanks, 23andMe). We performed maxFDR analyses to approximate the upper bound on the FDR of MTAG results.

GenomicSEM

Similarly, we used genomic structural equation modeling (GenomicSEM[16]) to model the joint genetic architecture of the four subgroups. First, we ran a common-factor model without individual SNP effects, following the tutorial 'Models without individual SNP effects' on the GenomicSEM github website (see web resources). Second, we ran a multivariate GWAS of the common factor (see **Supplementary Note 5** for details). We specified the model using unit-variance identification, for which the latent factor variance was fixed to 1 and the loadings of the traits are estimated freely. This ensures that we capture how much of each subgroup contributes to the latent factor. GenomicSEM also generates Q_{SNP}-values, which indicate possible heterogeneous effects across the subgroups. The Q_{SNP} statistic is mathematically similar to the Q-statistic from standard meta-analysis and is a X^2-distributed test statistic with larger values indexing a violation of the null hypothesis that the SNP acts entirely through the common factor.

SNP heritability estimation

The proportion of the phenotypic variance that could be explained by the aggregated effect of all included SNPs (SNP-based heritability, h^2_{SNP}) was estimated using LDSC[31]. The analysis was performed using pre-computed LDscores from samples restricted to European-ancestry in the

1000 Genomes Project[96], filtered for SNPs included in the HapMap3 reference panel[102]. SNP heritability was estimated based on the slope of the LDSC, with heritability on the liability scale calculated assuming a 1% population prevalence of OCD[1]. To omit a downward bias in our estimates of liability scale heritability, following Grotzinger *et al.*[103], we accounted for varying levels of ascertainment across cohorts in our meta-analysis by summing the effective sample sizes across the contributing cohorts and using that as the input sample size for LDSC. For the conversion to the liability scale (1%), the sample prevalence was then specified as 0.5. The SNP-heritability was calculated for the whole OCD sample as well as for ascertainment-specific sub-groups.

Genetic correlations

We used cross-trait LDSC[31], a method that computes genetic correlations between GWASs without bias from ancestry differences or sample overlap to calculate genetic correlations between the primary OCD meta-analysis and other phenotypes of interest. The selection of traits was based on phenotypic relevance and/or prior report of a genetic relationship with OCD. The genetic correlation between traits is based on the estimated slope from the regression of the product of *Z*-scores from two GWASs on the LD score and represents the genetic covariation between two traits based on all polygenic effects captured by the included SNPs. The genome-wide LD information used by these methods was based on European populations from the HapMap 3 reference panel[102], and GWAS summary statistics were filtered to only include SNPs that were part of the 1,290,028 HapMap 3 SNPs.

To ensure the internal consistency of the datasets included in our meta-analysis, we calculated genetic correlations between all cohorts we considered to have a sample size large enough for LDSC (effective sample size of $\geq$ 1,000) and between the four ascertainment-specific subgroups.

We further calculated genetic correlations between OCD and 112 other disorders and traits. The source studies of the GWAS summary statistics can be found in **Supplementary Table 18**. As a follow-up, we also calculated genetic correlations between the 112 phenotypes and each ascertainment-specific sub-cohort and compared the genetic correlation patterns between the four groups. For all cross-phenotype genetic correlation analyses, we adjusted *P*-values for multiple testing using the Benjamini-Hochberg procedure to control for FDR (< 0.05).

Gene-based analyses

To match the significant SNPs to the genes whose function they likely influence, we conducted a series of positional and functional gene mapping analyses. The positional mapping employed MBAT-combo[19], while the functional mapping tested whether genetic variants associated with OCD were also associated with differential expression of nearby genes (within 1-Mb window) using (i) transcriptome-wide association study (TWAS)[20] utilizing PsychENCODE data and included colocalization with COLOC[27,28] and (ii) summary-based Mendelian randomization (SMR[22]) using whole-blood eQTL information and brain tissues from MetaBrain, alongside the HEIDI test which tests for heterogeneity in GWAS signal and eQTL association. Further, a protein-wide association study (PWAS) was conducted. As a final step, genes within each locus were prioritized using PsyOPS[26], which integrates both positional and functional information. The details of each method are described below.

Positional gene mapping (MBAT-combo)

A gene-based analysis was conducted using multivariate Set-Based Association Test (mBAT-combo)[19] within GCTA version 1.94.1[14]. The European subsample (n = 503 individuals) from Phase 3 of the 1000 Genomes Project[96] was used as the LD reference panel with the fastBAT default LD cut-off of 0.9 applied. After filtering SNPs with MAF > 0.01, there were 6,629,124 SNPs for analysis in our sample. A gene list consisting of 19,899 protein coding genes was used to map the base pair position of genes using genome build hg19 (see **Supplementary Note 7** for details).

Functional gene mapping

Transcriptome-wide association study (TWAS)

We used TWAS FUSION[20] to perform a transcriptome-wide association study of OCD. We used brain gene expression weights from the PsychENCODE (104) and LD information from the 1000 Genomes Project Phase 3[96]. TWAS FUSION uses reference LD and reference gene expression panels with GWAS summary statistics to estimate the association between gene expression and OCD risk. These data were processed with the test statistics from the OCD GWAS to estimate the expression-GWAS association statistic. We corrected for multiple testing using Bonferroni correction.

We performed colocalization analyses using the COLOC R function[27,28] implemented in TWAS FUSION. Colocalization is a Bayesian method used to calculate the posterior probabilities (PP) that individual lead SNPs within a significant TWAS locus are (i) independent (e.g., two causal SNPs in LD, one affecting transcription, and one affecting OCD; posterior probability (PP3)) or (ii) share the same associated variant (e.g., a single causal SNP affects both transcription and OCD (PP4)). We also performed a conditional analysis to determine whether identified associations represented independent associations. This was performed using the FUSION software, which jointly estimates the effect of all significant features within each locus by using residual SNP associations with OCD after accounting for the predicted expression of other features.

Summary-based Mendelian randomization (SMR)

SMR[22] was performed using default settings and eQTL meta-analysis summary statistics from European populations for whole blood from eQTLGen[23], and all five nervous system tissues from MetaBrain (basal ganglia, cerebellum, cortex, hippocampus and spinal cord)[105]. The HEIDI (heterogeneity in dependent instruments) test is performed alongside SMR to test for effect size heterogeneity between the GWAS and eQTL summary statistics. Both SMR and TWAS have a number of important assumptions and limitations, which we discuss in **Supplementary Note 9**.

Psychiatric omnilocus prioritization score

We used the gene prioritization method PsyOPS (Psychiatric Omnilocus Prioritization Score)[26] to rank genes within genome-wide significant loci. This supervised approach integrates biological annotations on mutational intolerance, brain-specific expression, and involvement in

neurodevelopmental disorder for genes within significant loci. Genes with the top PsyOPS score within each locus were used for further gene prioritization (see “Gene prioritization” below). In the instance where two genes in the same locus had the same PsyOPS score, the gene nearest the index SNP was prioritized.

Protein-wide association study (PWAS)

We performed a protein-wide association study (PWAS) using protein expression data from human brain samples. Human brain proteome reference weight data were obtained using the Religious Orders Study and Rush Memory and Aging Project (ROS/MAP) and Banner Sun Health Research Institute (Banner) study. The ROS/MAP proteomes were generated from the dorsolateral prefrontal cortex (DLPFC) of 376 participants of European ancestry and included 1,476 proteins with significant SNP-based heritability ($P < 0.01$). The Banner PWAS weights were generated from 152 individuals of European ancestry and included 1,147 proteins with significant SNP-based heritability. The PWAS was performed using the TWAS FUSION software[20] with LD reference information from the 1000 Genomes Project Phase 3[96]. We corrected for multiple testing using Bonferroni correction.

Gene prioritization

We created a list of prioritized genes using both gene-based tests and colocalization/HEIDI filters. Results from each gene-based test were first restricted to protein coding genes with unique gene identifiers based on the release from GENCODE (v40) for hg19. The following criteria were then used to prioritize genes: (i) a significant (Bonferroni corrected) association from at least two gene-based tests (mBAT-combo, TWAS FUSION, SMR, or PsyOPS), and (ii) evidence of colocalization (COLOC PP4 > 0.8) and/or significant SMR association with HEIDI $P > 0.05$. Joint/conditional tests of association and significant PWAS associations were used as ancillary approaches to further annotate the prioritized gene list.

Tissue and cell-type enrichment analysis

An analysis of tissue and cell-type enrichment of OCD GWAS association signals was conducted using MAGMA (v1.08)[106] and partitioned LD score regression[107]. We employed Bryois et al.'s approach[29] to determine gene expression specificity in bulk tissue RNA-seq data from 37 tissues in GTEx (v8) and single-cell RNA sequencing data from 19 regions in the mouse central and peripheral nervous systems[30]. The analysis was limited to protein-coding genes with 1:1 orthologs between mice and humans. Gene expression in each tissue or cell type was calculated relative to total expression across all tissues or cell types. Enrichment analysis was performed on genes with the top 10% specificity values in each tissue or cell type, as previously defined[29].

To evaluate the enrichment of tissue and cell type specific genes in OCD genetic association signals, we applied MAGMA and partitioned LDSC. We restricted the analysis to summary statistics for SNPs with a high INFO score (> 0.6) and frequency in the entire cohort (MAF > 0.01). Using MAGMA (v1.08), we tested if genes with the top 10% specificity in a tissue or cell type showed enrichment in gene-level genetic associations for OCD, with the 1000 Genomes Phase 3 European sample genotypes serving as the LD reference panel. We used standard gene boundaries (35 kb upstream of the transcription start site to 10 kb downstream of the transcription

stop site). Partitioned LDSC was used to examine whether SNPs within 100-kb regions of the top 10% specifically expressed genes were enriched for SNP-based heritability for OCD. All results were corrected for multiple testing with an FDR threshold of 0.05.

SNP and gene findings in the context of previous analyses

Previously reported associations for significant SNPs (PheWAS)

Multiple resources were used to identify previously reported associations of our 30 significant SNPs with other phenotypes. We used the IEU open gwas project[92], PheWAS analysis of gwasATLAS[52], the NHGRI-EBI GWAS Catalog[91], and identified credible SNPs through causaldb[90]. Causaldb estimates causal probabilities of all genetic variants in GWAS significant loci using three state-of-the-art fine-mapping tools including PAINTOR, CAVIARBF and FINEMAP[108–111]. We used default settings for our causaldb queries.

Lookup of previous OCD GWAS findings

We performed a look-up of SNPs, identified to be significantly associated with OCD-related phenotypes in previous GWASs. Note that this is not an independent replication as previous studies partially overlap with the cohorts included in this GWAS.

Overlap of previous rare coding variants in OCD and GWAS gene-findings

We performed a bi-directional look-up, assessing (i) whether gene-findings from our GWAS showed evidence for rare variant involvement and (ii) vice versa, whether findings from rare variant testing showed evidence of common variant association in our GWAS.

First, we comprehensively assessed the overlap between 251 genes that we highlighted in our manuscript as carrying common risk variation for OCD (**Supplementary Table 14**), and current gene-based summary statistics from OCD exome sequencing data. We utilized data from Halvorsen *et al.*[88] since it is the largest published exome sequencing study of OCD presently. The supplemental materials from that paper include *de novo* variant calls from 771 case trios and 1,911 controls (Table S14 in Halvorsen *et al.*[88]). We compared the burden of *de novo* variants, partitioned by variant annotation (synonymous, missense, loss of function) in trio cases versus trio controls within these 251 GWAS-prioritized genes. As was done in the Halvorsen et al.[88], we only included de novo variants that were in loci well-covered in both case and control data (In_Jointly_Covered_Loci==TRUE). We also excluded all calls from quartet samples in Halvorsen et al. (Cohort!="OCD_JHU_quartets"). For each of the four variant annotation classes, we compared the proportion of cases that have at least one qualifying *de novo* variant to the proportion of controls using a two-sided Fisher's exact test.

Second, since Halvorsen *et al.*[88] describe an overall excess of loss of function variants in OCD cases relative to controls specifically within loss of function intolerant genes (Table S4 in Halvorsen *et al.*[88]), we analyzed the overlap between those genes and our GWAS-derived genes. We looked up 200 genes with a probability of loss of function intolerance > 0.995 (derived from Lek *et al.*[112]) and effect size estimate > 1. We further tested for a difference in the proportion of

these pLI > 0.995 genes with effect size estimate >1 vs ≤1 within the set of genes highlighted in the OCD GWAS (*n* = 251) vs. outside this set using a two-sided Fisher's exact test.

Data availability

The meta-analyzed summary statistics (not including 23andMe data) are available from the Psychiatric Genomics Consortium Download page (https://www.med.unc.edu/pgc/download-results/). In line with 23andMe regulations, 10,000 SNPs from the full GWAS including 23andMe are also being made available at the https://www.med.unc.edu/pgc/download-results/.

The full GWAS summary statistics for the 23andMe discovery data set will be made available through 23andMe to qualified researchers under an agreement with 23andMe that protects the privacy of the 23andMe participants. Datasets will be made available at no cost for academic use. Please visit https://research.23andme.com/collaborate/#dataset-access/ for more information and to apply to access the data.

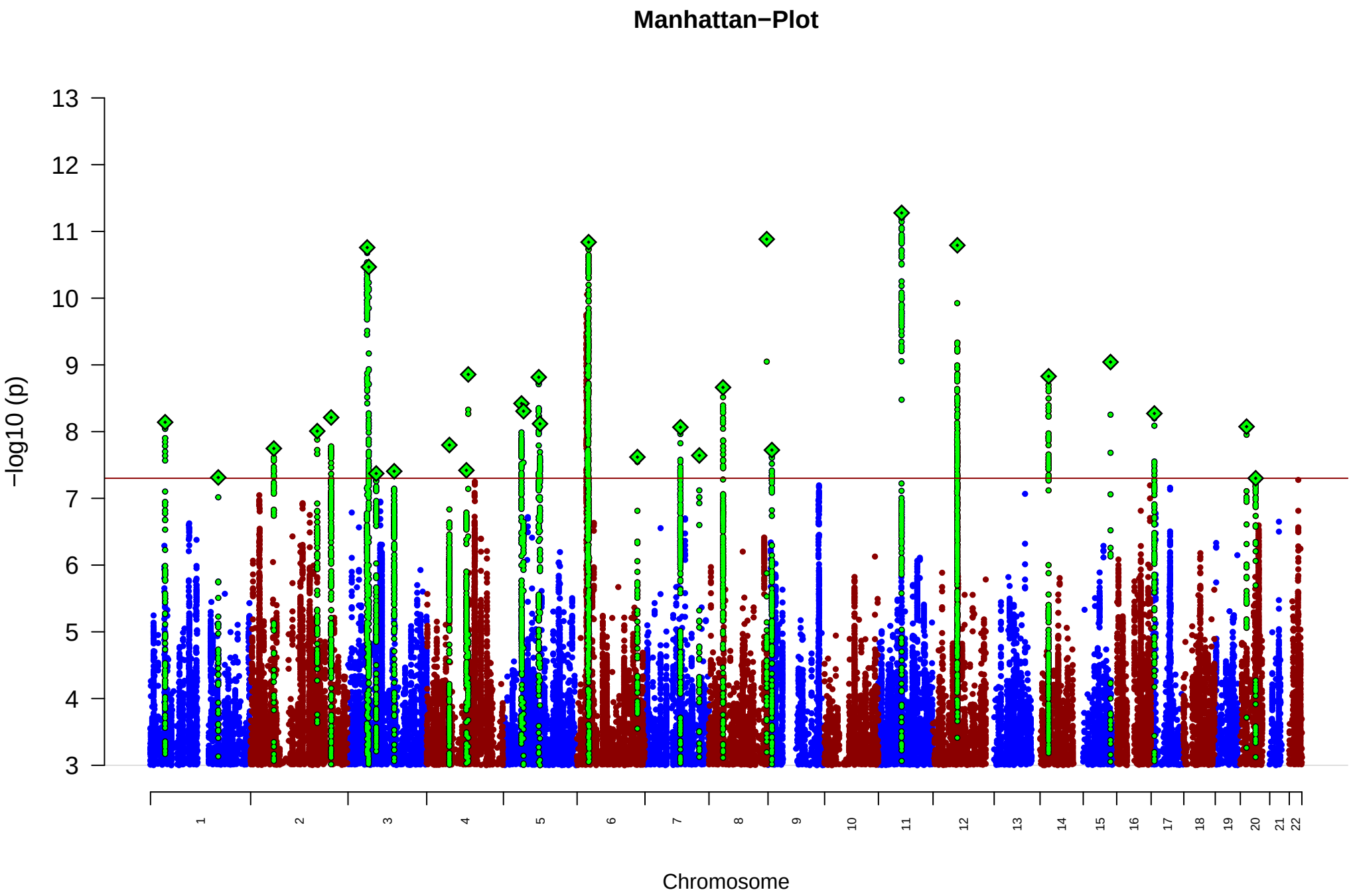
MVP summary statistics are made available through dbGAP request under accession phs001672.v12.p1.

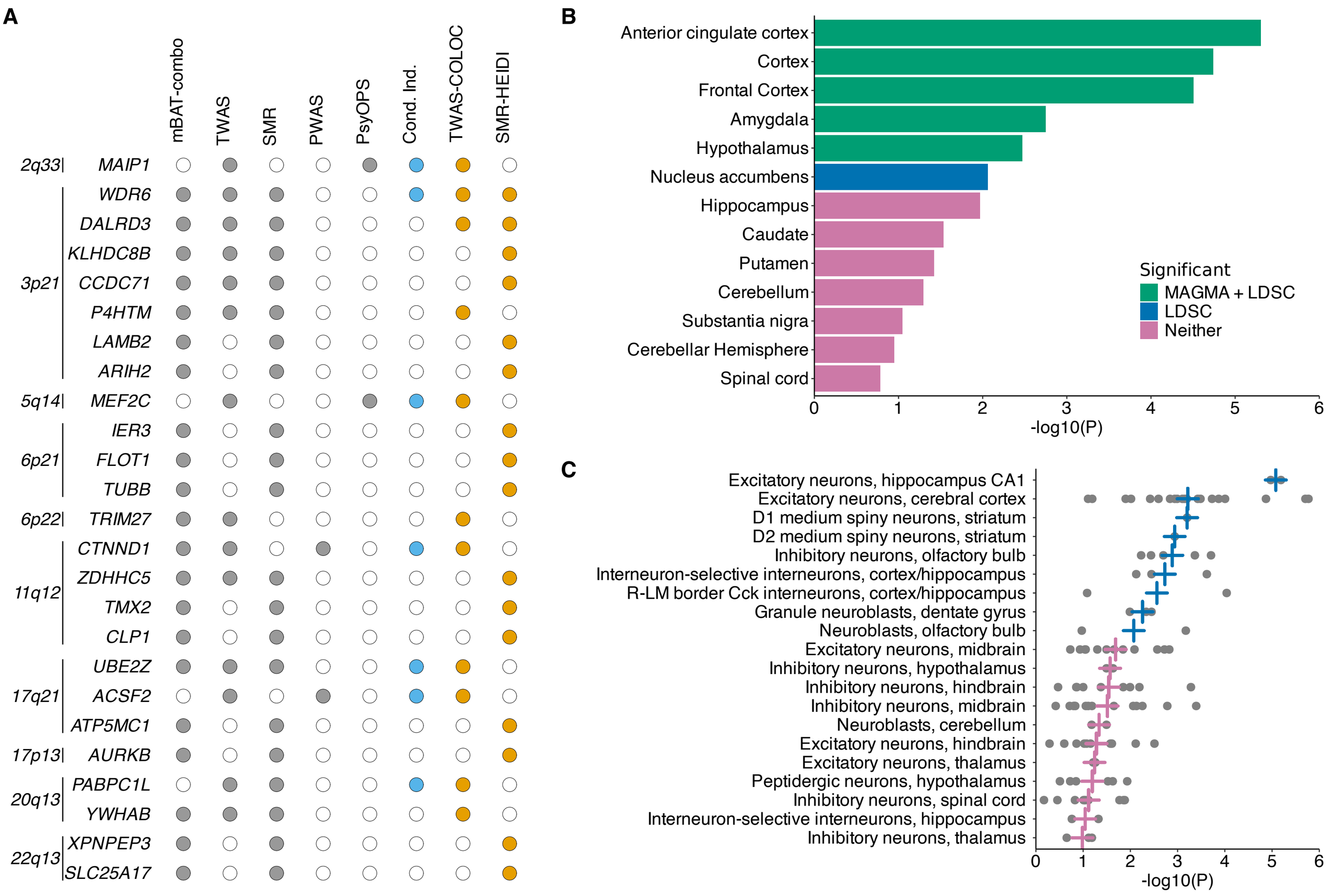
Code availability

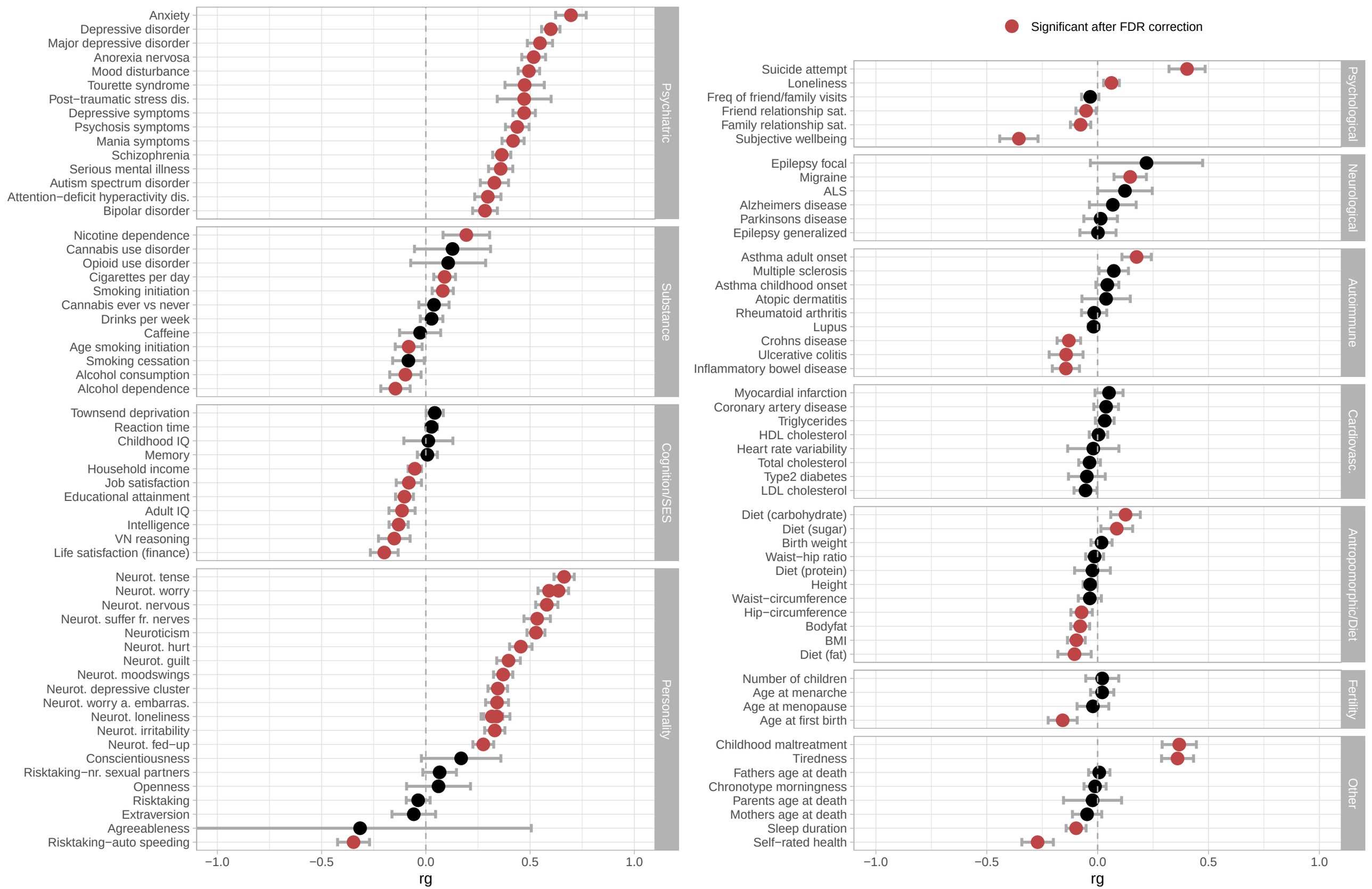
Core analysis code for RICOPILI can be found at https://sites.google.com/a/broadinstitute.org/ricopili/. This includes PLINK (https://www.cog-genomics.org/plink2/), EIGENSOFT (https://www.hsph.harvard.edu/alkes-price/software/), Eagle2 (https://alkesgroup.broadinstitute.org/Eagle/), Minimac3 (https://genome.sph.umich.edu/wiki/Minimac3), SHAPEIT3 (https://mathgen.stats.ox.ac.uk/genetics_software/shapeit/shapeit.html), METAL (https://genome.sph.umich.edu/wiki/METAL_Documentation) and LDSR (https://github.com/bulik/ldsc). MAGMA can be found at https://ctg.cncr.nl/software/magma . Genomic SEM, specifically the tutorial 'Models without Individual SNP effects' can be found here: https://github.com/GenomicSEM/GenomicSEM/wiki/3.-Models-without-Individual-SNP-effects
TWAS FUSION: http://gusevlab.org/projects/fusion/
PWAS: For access to the protein weights, see: https://www.synapse.org/#!Synapse:syn24872746
GCTA (mBAT-combo and COJO): https://yanglab.westlake.edu.cn/software/gcta/#Overview
LDSC and partitioned heritability: https://github.com/bulik/ldsc
Additional code for data processing (e.g., harmonization of summary statistics) can be found at https://www.doi.org/10.6084/m9.figshare.28451894

Methods-only references

93. DSM-5. Diagnostic and statistical manual of mental disorders: DSM-5. 237-242. ISBN 978-0-89042-555–8 (2013).
94. World Health Organization. ICD-11: International classification of diseases (11th revision), https://icd.who.int/ (2019).
95. McCarthy, S. *et al.* A reference panel of 64,976 haplotypes for genotype imputation. *Nat. Genet.* **48**, 1279–1283 (2016).
96. The 1000 Genomes Project Consortium. A global reference for human genetic variation. *Nature* **526**, 68–74 (2015).
97. Lam, M. *et al.* RICOPILI: Rapid Imputation for COnsortias PIpeLIne. *Bioinformatics* **36**, 930–933 (2020).
98. Willer, C. J., Li, Y. & Abecasis, G. R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190–2191 (2010).
99. Higgins, J. P. T. & Thompson, S. G. Quantifying heterogeneity in a meta-analysis. *Stat.in Med.* **21**, 1539–1558 (2002).
100. Higgins, J. P. T., Thompson, S. G., Deeks, J. J. & Altman, D. G. Measuring inconsistency in meta-analyses. *BMJ* **327**, 557–560 (2003).
101. Turley, P. *et al.* Multi-trait analysis of genome-wide association summary statistics using MTAG. *Nat. Genet.* **50**, 229–237 (2018).
102. Altshuler, D. M. *et al.* Integrating common and rare genetic variation in diverse human populations. *Nature* **467**, 52–58 (2010).
103. Grotzinger, A. D., de la Fuente, J., Privé, F., Nivard, M. G. & Tucker-Drob, E. M. Pervasive downward bias in estimates of liability-scale heritability in genome-wide association study meta-analysis: a simple solution. *Biol. Psychiatry* **93**, 29–36 (2023).
104. Wang, D. *et al.* Comprehensive functional genomic resource and integrative model for the human brain. *Science* **362**, eaat8464 (2018).
105. de Klein, N. *et al.* Brain expression quantitative trait locus and network analyses reveal downstream effects and putative drivers for brain-related diseases. *Nat. Genet.* **55**, 377–388 (2023).
106. de Leeuw, C. A., Mooij, J. M., Heskes, T. & Posthuma, D. MAGMA: generalized gene-set analysis of GWAS data. *PLoS Comput. Biol.* **11**, e1004219 (2015).
107. Finucane, H. K. *et al.* Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nat. Genet.* **47**, 1228–1235 (2015).
108. Benner, C., Spencer, C. C. A., Havulinna, A. S., Salomaa, V., Ripatti, S. & Pirinen, M. FINEMAP: efficient variable selection using summary data from genome-wide association studies. *Bioinformatics* **32**, 1493–1501 (2016).
109. Chen, W. *et al.* Fine mapping causal variants with an approximate Bayesian method using marginal test statistics. *Genetics* **200**, 719–736 (2015).
110. Kichaev, G. *et al.* Integrating functional data to prioritize causal variants in statistical fine-mapping studies. *PLoS Genet.* **10**, e1004722 (2014).
111. Kichaev, G. & Pasaniuc, B. Leveraging functional-annotation data in trans-ethnic fine-mapping studies. *Am J Hum Genet* **97**, 260–271 (2015).
112. Lek, M. *et al.* Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **536**, 285–291 (2016).


Manhattan−Plot
−log10 (p)
13
12
11
10
9
8
7
6
5
4
3
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
Chromosome


A
mBAT-combo
TWAS
SMR
PWAS
PsyOPS
Cond. Ind.
TWAS-COLOC
SMR-HEIDI
2q33
MAIP1
3p21
WDR6
DALRD3
KLHDC8B
CCDC71
P4HTM
LAMB2
ARIH2
5q14
MEF2C
6p21
IER3
FLOT1
TUBB
6p22
TRIM27
11q12
CTNND1
ZDHHC5
TMX2
CLP1
17q21
UBE2Z
ACSF2
ATP5MC1
17p13
AURKB
20q13
PABPC1L
YWHAB
22q13
XPNPEP3
SLC25A17
B
Anterior cingulate cortex
Cortex
Frontal Cortex
Amygdala
Hypothalamus
Nucleus accumbens
Hippocampus
Caudate
Putamen
Cerebellum
Substantia nigra
Cerebellar Hemisphere
Spinal cord
0
1
2
3
4
5
6
-log10(P)
Significant
MAGMA + LDSC
LDSC
Neither
C
Excitatory neurons, hippocampus CA1
Excitatory neurons, cerebral cortex
D1 medium spiny neurons, striatum
D2 medium spiny neurons, striatum
Inhibitory neurons, olfactory bulb
Interneuron-selective interneurons, cortex/hippocampus
R-LM border Cck interneurons, cortex/hippocampus
Granule neuroblasts, dentate gyrus
Neuroblasts, olfactory bulb
Excitatory neurons, midbrain
Inhibitory neurons, hypothalamus
Inhibitory neurons, hindbrain
Inhibitory neurons, midbrain
Neuroblasts, cerebellum
Excitatory neurons, hindbrain
Excitatory neurons, thalamus
Peptidergic neurons, hypothalamus
Inhibitory neurons, spinal cord
Interneuron-selective interneurons, hippocampus
Inhibitory neurons, thalamus
0
1
2
3
4
5
6
-log10(P)


Significant after FDR correction
Psychiatric
Anxiety
Depressive disorder
Major depressive disorder
Anorexia nervosa
Mood disturbance
Tourette syndrome
Post−traumatic stress dis.
Depressive symptoms
Psychosis symptoms
Mania symptoms
Schizophrenia
Serious mental illness
Autism spectrum disorder
Attention−deficit hyperactivity dis.
Bipolar disorder
Substance
Nicotine dependence
Cannabis use disorder
Opioid use disorder
Cigarettes per day
Smoking initiation
Cannabis ever vs never
Drinks per week
Caffeine
Age smoking initiation
Smoking cessation
Alcohol consumption
Alcohol dependence
Cognition/SES
Townsend deprivation
Reaction time
Childhood IQ
Memory
Household income
Job satisfaction
Educational attainment
Adult IQ
Intelligence
VN reasoning
Life satisfaction (finance)
Personality
Neurot. tense
Neurot. worry
Neurot. nervous
Neurot. suffer fr. nerves
Neuroticism
Neurot. hurt
Neurot. guilt
Neurot. moodswings
Neurot. depressive cluster
Neurot. worry a. embarras.
Neurot. loneliness
Neurot. irritability
Neurot. fed−up
Conscientiousness
Risktaking−nr. sexual partners
Openness
Risktaking
Extraversion
Agreeableness
Risktaking−auto speeding
−1.0
−0.5
0.0
0.5
1.0
rg
Psychological
Suicide attempt
Loneliness
Freq of friend/family visits
Friend relationship sat.
Family relationship sat.
Subjective wellbeing
Neurological
Epilepsy focal
Migraine
ALS
Alzheimers disease
Parkinsons disease
Epilepsy generalized
Autoimmune
Asthma adult onset
Multiple sclerosis
Asthma childhood onset
Atopic dermatitis
Rheumatoid arthritis
Lupus
Crohns disease
Ulcerative colitis
Inflammatory bowel disease
Cardiovasc.
Myocardial infarction
Coronary artery disease
Triglycerides
HDL cholesterol
Heart rate variability
Total cholesterol
Type2 diabetes
LDL cholesterol
Antropomorphic/Diet
Diet (carbohydrate)
Diet (sugar)
Birth weight
Waist−hip ratio
Diet (protein)
Height
Waist−circumference
Hip−circumference
Bodyfat
BMI
Diet (fat)
Fertility
Number of children
Age at menarche
Age at menopause
Age at first birth
Other
Childhood maltreatment
Tiredness
Fathers age at death
Chronotype morningness
Parents age at death
Mothers age at death
Sleep duration
Self−rated health
−1.0
−0.5
0.0
0.5
1.0
rg