DOI:
10.1038/s41588-019-0439-2

Document Version
Peer reviewed version

Abstract: 150
Text: 1804
Refs: 32
Tables: 1
Figures: 2
Supp Text: 1
Supp Tables: 20
Supp Figs: 16

Genome-wide Association Study Identifies Eight Risk Loci and Implicates Metabo-Psychiatric Origins for Anorexia Nervosa

Correspondence should be addressed to: C.M.B. (cbulik@med.unc.edu).

Hunna J Watson[1,2,3], Zeynep Yilmaz[1,4,a], Laura M Thornton[1,a], Christopher Hübel[5,6,a], Jonathan RI Coleman[5,7,a], Héléna A Gaspar[5,7], Julien Bryois[6], Anke Hinney[8], Virpi M Leppä[6], Manuel Mattheisen[9,10,11,12], Sarah E Medland[13], Stephan Ripke[14,15,16], Shuyang Yao[6], Paola Giusti-Rodríguez[4], Anorexia Nervosa Genetics Initiative[b], Ken B Hanscombe[17], Kirstin L Purves[5], Eating Disorders Working Group of the Psychiatric Genomics Consortium[b], Roger AH Adan[18,19,20], Lars Alfredsson[21], Tetsuya Ando[22], Ole A Andreassen[23], Jessica H Baker[1], Wade H Berrettini[24], Ilka Boehm[25], Claudette Boni[26], Vesna Boraska Perica[27,28], Katharina Buehren[29], Roland Burghardt[30], Matteo Cassina[31], Sven Cichon[32], Maurizio Clementi[31], Roger D Cone[33], Philippe Courtet[34], Scott Crow[35], James J Crowley[4,10], Unna N Danner[19], Oliver SP Davis[36,37], Martina de Zwaan[38], George Dedoussis[39], Daniela Degortes[40], Janiece E DeSocio[41], Danielle M Dick[42], Dimitris Dikeos[43], Christian Dina[44], Monika Dmitrzak-Weglarz[45], Elisa Docampo[46,47,48], Laramie E Duncan[49], Karin Egberts[50], Stefan Ehrlich[25], Geòrgia Escaramís[46,47,48], Tõnu Esko[51,52], Xavier Estivill[46,47,48,53], Anne Farmer[5], Angela Favaro[40], Fernando Fernández-Aranda[54,55], Manfred M Fichter[56,57], Krista Fischer[51], Manuel Föcker[8], Lenka Foretova[58], Andreas J Forstner[32,59,60,61,62], Monica Forzan[31], Christopher S Franklin[27], Steven Gallinger[63], Ina Giegling[64], Johanna Giuranna[8], Fragiskos Gonidakis[65], Philip Gorwood[26,66], Monica Gratacos Mayora[46,47,48], Sébastien Guillaume[34], Yiran Guo[67], Hakon Hakonarson[67,68], Konstantinos Hatzikotoulas[27,69], Joanna Hauser[70], Johannes Hebebrand[8], Sietske G Helder[5,71], Stefan Herms[32,60,62], Beate Herpertz-Dahlmann[29], Wolfgang Herzog[72], Laura M Huckins[27,73], James I Hudson[74], Hartmut Imgart[75], Hidetoshi Inoko[76], Vladimir Janout[77], Susana Jiménez-Murcia[54,55], Antonio Julià[78], Gursharan Kalsi[5], Deborah Kaminská[79], Jaakko Kaprio[80,81], Leila Karhunen[82], Andreas Karwautz[83], Martien JH Kas[18,84], James L Kennedy[85,86,87], Anna Keski-Rahkonen[80], Kirsty Kiezebrink[88], Youl-Ri Kim[89], Lars Klareskog[90], Kelly L Klump[91], Gun Peggy S Knudsen[92], Maria C La Via[1], Stephanie Le Hellard[93,94,95], Robert D Levitan[85,86,87], Dong Li[67], Lisa Lilenfeld[96], Bochao Danae Lin[18], Jolanta Lissowska[97], Jurjen Luykx[18], Pierre J Magistretti[98,99], Mario Maj[100], Katrin Mannik[51,101], Sara Marsal[78], Christian R Marshall[102], Morten Mattingsdal[23], Sara McDevitt[103,104], Peter McGuffin[5], Andres Metspalu[51,105], Ingrid Meulenbelt[106], Nadia Micali[107,108,109], Karen Mitchell[110], Alessio Maria Monteleone[100], Palmiero Monteleone[111], Melissa A Munn-Chernoff[l], Benedetta Nacmias[112], Marie Navratilova[58], Ioanna Ntalla[39], Julie K O'Toole[113], Roel A Ophoff[18,114], Leonid Padyukov[90], Aarno Palotie[52,81,115], Jacques Pantel[26], Hana Papezova[79], Dalila Pinto[73], Raquel Rabionet[116,117,118], Anu Raevuori[80], Nicolas Ramoz[26], Ted Reichborn-Kjennerud[92,119], Valdo Ricca[112,120], Samuli Ripatti[52,80,121], Franziska Ritschel[25,122], Marion Roberts[5,123,124], Alessandro Rotondo[125], Dan Rujescu[56,64], Filip Rybakowski[126], Paolo Santonastaso[127], André Scherag[128], Stephen W Scherer[129], Ulrike Schmidt[7,130], Nicholas J Schork[131], Alexandra Schosser[132], Jochen Seitz[29], Lenka Slachtova[133], P.

Eline Slagboom[106], Margarita CT Slof-Op 't Landt[134,135], Agnieszka Slopien[136], Sandro Sorbi[112,137], Beata Świątkowska[138], Jin P Szatkiewicz[4], Ioanna Tachmazidou[27], Elena Tenconi[40], Alfonso Tortorella[139,140], Federica Tozzi[141], Janet Treasure[7,130], Artemis Tsitsika[142], Marta Tyszkiewicz-Nwafor[136], Konstantinos Tziouvas[143], Annemarie A van Elburg[19,144], Eric F van Furth[134,135], Gudrun Wagner[83], Esther Walton[25], Elisabeth Widen[81], Eleftheria Zeggini[27,69], Stephanie Zerwas[1], Stephan Zipfel[145], Andrew W Bergen[146,147], Joseph M Boden[148], Harry Brandt[149], Steven Crawford[149], Katherine A Halmi[150], L. John Horwood[148], Craig Johnson[151], Allan S Kaplan[85,86,87], Walter H Kaye[152], James Mitchell[153], Catherine M Olsen[13], John F Pearson[154], Nancy L Pedersen[6], Michael Strober[155,156], Thomas Werge[157], David C Whiteman[13], D. Blake Woodside[86,87,158,159], Garret D Stuber[1,160], Scott Gordon[13], Jakob Grove[9,161,162,163], Anjali K Henders[164], Anders Juréus[6], Katherine M Kirk[13], Janne T Larsen[161,165,166], Richard Parker[13], Liselotte Petersen[161,165,166], Jennifer Jordan[123,167], Martin Kennedy[168], Grant W Montgomery[13,164,169], Tracey D Wade[170], Andreas Birgegård[10,11], Paul Lichtenstein[6], Claes Norring[10,11], Mikael Landén[6,171,a], Nicholas G Martin[13,a], Preben Bo Mortensen[161,165,166,a], Patrick F Sullivan[1,4,6,a], Gerome Breen[5,7,c], Cynthia M Bulik[1,6,172,c]

1 Department of Psychiatry, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, US
2 School of Psychology, Curtin University, Perth, Australia
3 School of Paediatrics and Child Health, University of Western Australia, Perth, Australia
4 Department of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, US
5 Institute of Psychiatry, Psychology and Neuroscience, Social, Genetic and Developmental Psychiatry (SGDP) Centre, King's College London, London, UK
6 Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden
7 National Institute for Health Research Biomedical Research Centre, King's College London and South London and Maudsley National Health Service Foundation Trust, London, UK
8 Department of Child and Adolescent Psychiatry, University Hospital Essen, University of Duisburg-Essen, Essen, Germany
9 Department of Biomedicine, Aarhus University, Aarhus, Denmark
10 Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden
11 Center for Psychiatry Research, Stockholm Health Care Services, Stockholm City Council, Stockholm, Sweden
12 Department of Psychiatry, Psychosomatics and Psychotherapy, University of Würzburg, Würzburg, Germany
13 QIMR Berghofer Medical Research Institute, Brisbane, Australia
14 Analytic and Translational Genetics Unit, Massachusetts General Hospital, Boston, Massachusetts, US
15 Stanley Center for Psychiatric Research, Broad Institute of the Massachusetts Institute of Technology and Harvard University, Cambridge, Massachusetts, US
16 Department of Psychiatry and Psychotherapy, Charité - Universitätsmedizin, Berlin, Germany
17 Department of Medical and Molecular Genetics, King's College London, Guy's Hospital, London, UK
18 Brain Center Rudolf Magnus, Department of Translational Neuroscience, University Medical Center Utrecht, Utrecht, The Netherlands

19 Center for Eating Disorders Rintveld, Altrecht Mental Health Institute, Zeist, The Netherlands
20 Institute of Neuroscience and Physiology, Sahlgrenska Academy at the University of Gothenburg, Gothenburg, Sweden
21 Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden
22 Department of Behavioral Medicine, National Institute of Mental Health, National Center of Neurology and Psychiatry, Tokyo, Japan
23 NORMENT KG Jebsen Centre, Division of Mental Health and Addiction, University of Oslo, Oslo University Hospital, Oslo, Norway
24 Department of Psychiatry, Center for Neurobiology and Behavior, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, Pennsylvania, US
25 Division of Psychological and Social Medicine and Developmental Neurosciences, Faculty of Medicine, Technische Universität Dresden, Dresden, Germany
26 INSERM 1266, Institute of Psychiatry and Neuroscience of Paris, Paris, France
27 Wellcome Sanger Institute, Wellcome Genome Campus, Hinxton, Cambridge, UK
28 Department of Medical Biology, School of Medicine, University of Split, Split, Croatia
29 Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, RWTH Aachen University, Aachen, Germany
30 Department of Child and Adolescent Psychiatry, Klinikum Frankfurt/Oder, Frankfurt, Germany
31 Clinical Genetics Unit, Department of Woman and Child Health, University of Padova, Padova, Italy
32 Institute of Medical Genetics and Pathology, University Hospital Basel, Basel, Switzerland
33 Life Sciences Institute and Department of Molecular and Integrative Physiology, University of Michigan, Ann Arbor, Michigan, US
34 Department of Emergency Psychiatry and Post-Acute Care, CHRU Montpellier, University of Montpellier, Montpellier, France
35 Department of Psychiatry, University of Minnesota, Minneapolis, Minnesota, US
36 MRC Integrative Epidemiology Unit, University of Bristol, Bristol, UK
37 School of Social and Community Medicine, University of Bristol, Bristol, UK
38 Department of Psychosomatic Medicine and Psychotherapy, Hannover Medical School, Hannover, Germany
39 Department of Nutrition and Dietetics, Harokopio University, Athens, Greece
40 Department of Neurosciences, University of Padova, Padova, Italy
41 College of Nursing, Seattle University, Seattle, Washington, US
42 Department of Psychology, Virginia Commonwealth University, Richmond, Virginia, US
43 Department of Psychiatry, Athens University Medical School, Athens University, Athens, Greece
44 L'institut du thorax, INSERM, CNRS, UNIV Nantes, CHU Nantes, Nantes, France
45 Department of Psychiatric Genetics, Poznan University of Medical Sciences, Poznan, Poland
46 Barcelona Institute of Science and Technology, Barcelona, Spain
47 Universitat Pompeu Fabra, Barcelona, Spain
48 Centro de Investigación Biomédica en Red en Epidemiología y Salud Pública (CIBERESP), Barcelona, Spain
49 Department of Psychiatry and Behavioral Sciences, Stanford University, Stanford, California, US

50 Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, University Hospital of Würzburg, Centre for Mental Health, Würzburg, Germany
51 Estonian Genome Center, University of Tartu, Tartu, Estonia
52 Program in Medical and Population Genetics, Broad Institute of the Massachusetts Institute of Technology and Harvard University, Cambridge, Massachusetts, US
53 Genomics and Disease, Bioinformatics and Genomics Programme, Centre for Genomic Regulation, Barcelona, Spain
54 Department of Psychiatry, University Hospital of Bellvitge –IDIBELL and CIBERobn, Barcelona, Spain
55 Department of Clinical Sciences, School of Medicine, University of Barcelona, Barcelona, Spain
56 Department of Psychiatry and Psychotherapy, Ludwig-Maximilians-University (LMU), Munich, Germany
57 Schön Klinik Roseneck affiliated with the Medical Faculty of the University of Munich (LMU), Munich, Germany
58 Department of Cancer, Epidemiology and Genetics, Masaryk Memorial Cancer Institute, Brno, Czech Republic
59 Institute of Human Genetics, University of Bonn, School of Medicine & University Hospital Bonn, Bonn, Germany
60 Department of Genomics, Life and Brain Center, University of Bonn, Bonn, Germany
61 Department of Psychiatry (UPK), University of Basel, Basel, Switzerland
62 Department of Biomedicine, University of Basel, Basel, Switzerland
63 Department of Surgery, Faculty of Medicine, University of Toronto, Toronto, Canada
64 Department of Psychiatry, Psychotherapy and Psychosomatics, Martin Luther University of Halle-Wittenberg, Halle, Germany
65 1st Psychiatric Department, National and Kapodistrian University of Athens, Medical School, Eginition Hospital, Athens, Greece
66 CMME, Hôpital Sainte-Anne (GHU Paris Psychiatrie et Neurosciences), Paris Descartes University, Paris, France
67 Center for Applied Genomics, Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, US
68 Department of Pediatrics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, US
69 Institute of Translational Genomics, Helmholtz Zentrum München, Neuherberg, Germany
70 Department of Adult Psychiatry, Poznan University of Medical Sciences, Poznan, Poland
71 Zorg op Orde, Leidschendam, The Netherlands
72 Department of General Internal Medicine and Psychosomatics, Heidelberg University Hospital, Heidelberg University, Heidelberg, Germany
73 Department of Psychiatry, and Genetics and Genomics Sciences, Division of Psychiatric Genomics, Icahn School of Medicine at Mount Sinai, New York, New York, US
74 Biological Psychiatry Laboratory, McLean Hospital/Harvard Medical School, Boston, Massachusetts, US
75 Eating Disorders Unit, Parklandklinik, Bad Wildungen, Germany
76 Department of Molecular Life Science, Division of Basic Medical Science and Molecular Medicine, School of Medicine, Tokai University, Isehara, Japan
77 Faculty of Health Sciences, Palacky University, Olomouc, Czech Republic

78 Rheumatology Research Group, Vall d'Hebron Research Institute, Barcelona, Spain
79 Department of Psychiatry, First Faculty of Medicine, Charles University, Prague, Czech Republic
80 Department of Public Health, University of Helsinki, Helsinki, Finland
81 Institute for Molecular Medicine Finland, Helsinki Institute of Life Science, University of Helsinki, Helsinki, Finland
82 Institute of Public Health and Clinical Nutrition, Department of Clinical Nutrition, University of Eastern Finland, Kuopio, Finland
83 Eating Disorders Unit, Department of Child and Adolescent Psychiatry, Medical University of Vienna, Vienna, Austria
84 Groningen Institute for Evolutionary Life Sciences, University of Groningen, Groningen, The Netherlands
85 Centre for Addiction and Mental Health, Toronto, Canada
86 Institute of Medical Science, University of Toronto, Toronto, Canada
87 Department of Psychiatry, University of Toronto, Toronto, Canada
88 Institute of Applied Health Sciences, University of Aberdeen, Aberdeen, UK
89 Department of Psychiatry, Seoul Paik Hospital, Inje University, Seoul, Korea
90 Rheumatology Unit, Department of Medicine, Center for Molecular Medicine, Karolinska Institutet and Karolinska University Hospital, Stockholm, Sweden
91 Department of Psychology, Michigan State University, East Lansing, Michigan, US
92 Department of Mental Disorders, Norwegian Institute of Public Health, Oslo, Norway
93 Department of Clinical Science, K.G. Jebsen Centre for Psychosis Research, Norwegian Centre for Mental Disorders Research (NORMENT), University of Bergen, Bergen, Norway
94 Dr. Einar Martens Research Group for Biological Psychiatry, Center for Medical Genetics and Molecular Medicine, Haukeland University Hospital, Bergen, Norway
95 Department of Clinical Medicine, Laboratory Building, Haukeland University Hospital, Bergen, Norway
96 American School of Professional Psychology, Argosy University, Northern Virginia, Arlington, Virginia, US
97 Department of Cancer Epidemiology and Prevention, M Skłodowska-Curie Cancer Center - Oncology Center, Warsaw, Poland
98 BESE Division, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia
99 Department of Psychiatry, University of Lausanne-University Hospital of Lausanne (UNIL-CHUV), Lausanne, Switzerland
100 Department of Psychiatry, University of Campania "Luigi Vanvitelli", Naples, Italy
101 Center for Integrative Genomics, University of Lausanne, Lausanne, Switzerland
102 Department of Paediatric Laboratory Medicine, The Hospital for Sick Children, Toronto, Canada
103 Department of Psychiatry, University College Cork, Cork, Ireland
104 HSE National Clinical Programme for Eating Disorders, Cork, Ireland
105 Institute of Molecular and Cell Biology, University of Tartu, Tartu, Estonia
106 Department of Biomedical Data Science, Leiden University Medical Centre, Leiden, The Netherlands
107 Department of Psychiatry, Faculty of Medicine, University of Geneva, Geneva, Switzerland
108 Division of Child and Adolescent Psychiatry, Geneva University Hospital, Geneva, Switzerland

109 Great Ormond Street Institute of Child Health, University College London, London, UK
110 National Center for PTSD, VA Boston Healthcare System, Department of Psychiatry, Boston University School of Medicine, Boston, Massachusetts, US
111 Department of Medicine, Surgery and Dentistry "Scuola Medica Salernitana", University of Salerno, Salerno, Italy
112 Department of Neuroscience, Psychology, Drug Research and Child Health (NEUROFARBA), University of Florence, Florence, Italy
113 Kartini Clinic, Portland, Oregon, US
114 Center for Neurobehavioral Genetics, Semel Institute for Neuroscience and Human Behavior, University of California Los Angeles, Los Angeles, California, US
115 Center for Human Genome Research at the Massachusetts General Hospital, Boston, Massachusetts, US
116 Saint Joan de Déu Research Institute, Saint Joan de Déu Barcelona Children's Hospital, Barcelona, Spain
117 Institute of Biomedicine (IBUB), University of Barcelona, Barcelona, Spain
118 Department of Genetics, Microbiology and Statistics, University of Barcelona, Barcelona, Spain
119 Institute of Clinical Medicine, University of Oslo, Oslo, Norway
120 Department of Health Science, University of Florence, Florence, Italy
121 Institute for Molecular Medicine Finland (FIMM), HiLIFE Unit, University of Helsinki, Helsinki, Finland
122 Eating Disorders Research and Treatment Center, Department of Child and Adolescent Psychiatry, Faculty of Medicine, Technische Universität Dresden, Dresden, Germany
123 Department of Psychological Medicine, University of Otago, Christchurch, New Zealand
124 Faculty of Medicine & Health Sciences, University of Aukland, Aukland, New Zealand
125 Department of Psychiatry, Neurobiology, Pharmacology, and Biotechnologies, University of Pisa, Pisa, Italy
126 Department of Psychiatry, Poznan University of Medical Sciences, Poznan, Poland
127 Department of Neurosciences, Padua Neuroscience Center, University of Padova, Padova, Italy
128 Institute of Medical Statistics, Computer and Data Sciences, Jena University Hospital, Jena, Germany
129 Department of Genetics and Genomic Biology, The Hospital for Sick Children, Toronto, Canada
130 Institute of Psychiatry, Psychology and Neuroscience, Department of Psychological Medicine, King's College London, London, UK
131 J. Craig Venter Institute (JCVI), La Jolla, California, US
132 Department of Psychiatry and Psychotherapy, Medical University of Vienna, Vienna, Austria
133 Department of Pediatrics and Center of Applied Genomics, First Faculty of Medicine, Charles University, Prague, Czech Republic
134 Center for Eating Disorders Ursula, Rivierduinen, Leiden, The Netherlands
135 Department of Psychiatry, Leiden University Medical Centre, Leiden, The Netherlands
136 Department of Child and Adolescent Psychiatry, Poznan University of Medical Sciences, Poznan, Poland
137 IRCSS Fondazione Don Carlo Gnocchi, Florence, Italy

138 Department of Environmental Epidemiology, Nofer Institute of Occupational Medicine, Lodz, Poland
139 Department of Psychiatry, University of Naples SUN, Naples, Italy
140 Department of Psychiatry, University of Perugia, Perugia, Italy
141 Brain Sciences Department, Stremble Ventures, Limassol, Cyprus
142 Adolescent Health Unit, Second Department of Pediatrics, "P. & A. Kyriakou" Children's Hospital, University of Athens, Athens, Greece
143 Pediatric Intensive Care Unit, "P. & A. Kyriakou" Children's Hospital, University of Athens, Athens, Greece
144 Faculty of Social and Behavioral Sciences, Utrecht University, Utrecht, The Netherlands
145 Department of Internal Medicine VI, Psychosomatic Medicine and Psychotherapy, University Medical Hospital Tuebingen, Tuebingen, Germany
146 BioRealm, LLC, Walnut, California, US
147 Oregon Research Institute, Eugene, Oregon, US
148 Christchurch Health and Development Study, University of Otago, Christchurch, New Zealand
149 The Center for Eating Disorders at Sheppard Pratt, Baltimore, Maryland, US
150 Department of Psychiatry, Weill Cornell Medical College, New York, New York, US
151 Eating Recovery Center, Denver, Colorado, US
152 Department of Psychiatry, University of California San Diego, San Diego, California, US
153 Department of Psychiatry and Behavioral Science, University of North Dakota School of Medicine and Health Sciences, Fargo, North Dakota, US
154 Biostatistics and Computational Biology Unit, University of Otago, Christchurch, New Zealand
155 Department of Psychiatry and Biobehavioral Science, Semel Institute for Neuroscience and Human Behavior, University of California Los Angeles, Los Angeles, California, US
156 David Geffen School of Medicine, University of California Los Angeles, Los Angeles, California, US
157 Department of Clinical Medicine, University of Copenhagen, Copenhagen, Denmark
158 Centre for Mental Health, University Health Network, Toronto, Canada
159 Program for Eating Disorders, University Health Network, Toronto, Canada
160 Department of Cell Biology and Physiology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, US
161 The Lundbeck Foundation Initiative for Integrative Psychiatric Research (iPSYCH), Aarhus, Denmark
162 Centre for Integrative Sequencing, iSEQ, Aarhus University, Aarhus, Denmark
163 Bioinformatics Research Centre, Aarhus University, Aarhus, Denmark
164 Institute for Molecular Bioscience, University of Queensland, Brisbane, Australia
165 National Centre for Register-Based Research, Aarhus BSS, Aarhus University, Aarhus, Denmark
166 Centre for Integrated Register-based Research (CIRRAU), Aarhus University, Aarhus, Denmark
167 Canterbury District Health Board, Christchurch, New Zealand
168 Department of Pathology and Biomedical Science, University of Otago, Christchurch, New Zealand
169 Queensland Brain Institute, University of Queensland, Brisbane, Australia

170 School of Psychology, Flinders University, Adelaide, Australia
171 Department of Psychiatry and Neurochemistry, Institute of Neuroscience and Physiology, Sahlgrenska Academy at the University of Gothenburg, Gothenburg, Sweden
172 Department of Nutrition, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, US

a These authors contributed equally to this work
b The members of this consortium are listed in the Supplementary Note
c These authors jointly directed this project
* e-mail: cynthia_bulik@med.unc.edu

335 ***Genome-wide Association Study Identifies Eight Risk Loci and Implicates***

336 ***Metabo-Psychiatric Origins for Anorexia Nervosa***

337

338 **Characterized primarily by low BMI, anorexia nervosa is a complex and serious illness[1],**

339 **affecting 0.9-4% of women and 0.3% of men[2-4], with twin-based heritability estimates of**

340 **50-60%[5]. Mortality rates are higher than other psychiatric disorders[6], and outcomes are**

341 **unacceptably poor[7]. Combining data from the Anorexia Nervosa Genetics Initiative**

342 **(ANGI)[8,9] and the Eating Disorders Working Group of the Psychiatric Genomics**

343 **Consortium (PGC-ED), we conducted a genome-wide association study (GWAS) of 16,992**

344 **anorexia nervosa cases and 55,525 controls, identifying eight significant loci. The genetic**

345 **architecture of anorexia nervosa mirrors its clinical presentation showing significant**

346 **genetic correlations with psychiatric disorders, physical activity, metabolic (including**

347 **glycemic), lipid, and anthropometric traits, independent of the effects of common variants**

348 **associated with BMI. Results further encourage a reconceptualization of anorexia nervosa**

349 **as a metabo-psychiatric disorder. Explicating the metabolic component is a critical**

350 **direction, and attention to both psychiatric and metabolic components may be key to**

351 **improving outcomes.**

352 The first PGC-ED GWAS (3,495 cases, 10,982 controls) estimated the common genetic

353 variant-based heritability of anorexia nervosa as ~20%, identified the first genome-wide

354 significant locus, and reported significant genetic correlations (r_g) between anorexia nervosa and

355 psychiatric and metabolic/anthropometric phenotypes[10]. These r_g pointed toward metabolic

356 etiological factors, as they are robust to reverse causation although they could be mediated

357 associations[11] or reflect confounding processes[12]. To advance genomic discovery in anorexia

nervosa and further explore genetic correlations, we combined samples from ANGI[8,9], the Genetic Consortium for Anorexia Nervosa (GCAN)/Wellcome Trust Case Control Consortium-3 (WTCCC-3)[13], and the UK Biobank[14], quadrupling our sample size.

Our GWAS meta-analysis included 33 datasets comprising 16,992 cases and 55,525 controls of European ancestry from 17 countries (**Supplementary Tables 1-4**). We had 80% power to detect an odds ratio (OR) of 1.09-1.19 (additive model, 0.9% lifetime risk, $\alpha = 5 \times 10^{-8}$, MAF 0.05–0.5). Typical of complex trait GWAS, we observed test statistic inflation ($\lambda = 1.22$) consistent with polygenicity, with no evidence of significant population stratification according to the LD intercept and attenuation ratio (**Supplementary Results; Supplementary Fig. 1**). Meta-analysis results were completed for autosomes and the X chromosome. We identified eight loci exceeding genome-wide significance ($P < 5 \times 10^{-8}$; **Table 1** for loci; **Fig. 1** for the Manhattan plot; **Supplementary Figs. 2a-h** and **3a-h** for the forest and region plots). Many were near the threshold for significance, and no significant heterogeneity of SNP associations across cohorts was detected ($P = 0.15$-0.64; **Supplementary Figs 2a-h**). Conditional and joint analysis (GCTA-COJO)[15] confirmed independence of the lead SNPs within the significant loci (**Supplementary Table 5**). The eight loci were annotated to identify known protein-coding genes (**Supplementary Table 6; Supplementary Table 7** reports a gene look-up restricted to the single-gene loci). The previously reported PGC-ED genome-wide significant variant (rs4622308)[10] on 12q13.2 did not reach genome-wide significance ($P = 7.02 \times 10^{-5}$); however, between-cohort heterogeneity was apparent ($I^2 = 53.7$; **Supplementary Fig. 4** and **Supplementary Results**). The OR was in the same direction in 22 (67%) of the cohorts ($z = 2.00$, $P = 0.05$, 2-tailed).

Although GWAS findings are informative genome-wide, identifying strong hypotheses about their connections to specific genes is not straightforward. We evaluated three ways to "connect" anorexia nervosa GWAS loci to genes: regulatory chromatin interactions; relationship to brain expression QTLs (eQTLs; using a superset of CommonMind[16] and GTEx[17]) and the standard approach of gene location within a GWAS locus. The significant anorexia nervosa loci implicated 121 brain-expressed genes, 74% by location, 55% by adult brain eQTL, 93% by regulatory chromatin interaction, and 58 genes by all three methods. **Supplementary Figs. 5a-h** show the eight GWAS loci, GENCODE gene models, adult brain regulatory chromatin interactions, brain eQTLs, and functional genomic annotations.

Four single-gene loci were confirmed by eQTL, chromatin interaction, or both. These were the locus-intersecting genes *CADM1* (locus 2 chr11:114.9-115.4 Mb, **Supplementary Fig. 5b**), *MGMT* (locus 4, chr10:131.2-131.4 Mb, **Supplementary Fig. 5d**), *FOXP1* (locus 5, chr3:70.6-71.0 Mb, **Supplementary Fig. 5e**) and *PTBP2* (locus 6, chr1:96.6-97.2 Mb, **Supplementary Fig. 5f**). For locus 5, eQTL data implicated a distal gene, *GPR27*. One intergenic locus (locus 7, chr5:24.9-25.3 Mb, **Supplementary Fig. 5g**) had no eQTL or chromatin interactions whereas the other intergenic locus (locus 8, chr3:93.9-95.0 Mb, **Supplementary Fig. 5h**) had eQTL connections to *PROS1* and *ARL13B*. Two complex multigenic loci had many brain-expressed genes and dense chromatin and eQTL interactions that precluded identification of any single gene (locus 1, chr3:47.5-51.3 Mb; locus 3, chr2:53.8-54.3 Mb, **Supplementary Figs. 5a** and **5c**). The clearest evidence and connections were for the single-gene loci intersecting *CADM1*, *MGMT*, *FOXP1,* and *PTBP2* and we conclude these genes may play a role in anorexia nervosa etiology (**Supplementary Results**).

Supplementary Table 8 presents multi-trait analysis (GCTA-mtCOJO[18] conditioning our genome-wide significant SNPs on associated variants in GWAS of BMI, type 2 diabetes, education years, HDL cholesterol, neuroticism, and schizophrenia. Seven loci appear to be independent. Locus 2 on chr11 may not be unique to anorexia nervosa and may be driven by genetic variation also associated with type 2 diabetes.

Liability-scale SNP heritability (SNP-h^2) was estimated with LD score regression (LDSC)[19,20]. Assuming a lifetime prevalence of 0.9-4%[2-4], SNP-h^2 was 11-17% (s.e. = 1%), supporting the polygenic nature of anorexia nervosa. Polygenic risk score (PRS) analyses using a leave-one-out approach indicated that the PRS captures ~1.7% of the phenotypic variance on the liability scale for discovery $P = 0.5$. We did not observe differences in polygenic architecture between anorexia nervosa subtypes with binge eating (2,381 cases, 10,249 controls) or without (2,262 cases, 10,254 controls) or between males (447 cases, 20,347 controls) and females (14,898 cases, 27,545 controls) (**Methods**, **Supplementary Results**, **Supplementary Fig. 6, Supplementary Table 9**). Similar to females, males in the highest PRS decile had 4.13 (95% CI: 2.58-6.62) times the odds of anorexia nervosa than those in the lowest decile. Confirmation of these results requires larger samples.

We tested SNP-based genetic correlations (SNP-r_g) with external traits using bivariate LDSC[19,20]. Bonferroni-significant SNP-r_g assorted into five trait categories: psychiatric and personality; physical activity; anthropometric; metabolic; and educational attainment (**Supplementary Table 10**). **Fig. 2** presents Bonferroni-corrected positive SNP-r_g with OCD (SNP-$r_g \pm$ s.e. = 0.45 ± 0.08; $P = 4.97 \times 10^{-9}$), MDD (0.28 ± 0.07; $P = 8.95 \times 10^{-5}$), anxiety disorders (0.25 ± 0.05; $P = 8.90 \times 10^{-8}$), and schizophrenia (0.25 ± 0.03; $P = 4.61 \times 10^{-18}$). This pattern reflects observed comorbidities in clinical and epidemiological studies[21,22]. The newly-

identified positive SNP-r_g with physical activity (0.17 ± 0.05; $P = 1.00 \times 10^{-4}$) encourages further exploration of the refractory symptom of pathologically elevated activity in anorexia nervosa[23]. We note that the significant SNP-r_g of anorexia nervosa with educational attainment (0.25 ± 0.03; $P = 1.69 \times 10^{-15}$) and related constructs was not seen for IQ[24].

Expanding our previous observations[10], we present a palette of metabolic and anthropometric r_g with anorexia nervosa more pronounced than in other psychiatric disorders. We observed significant negative SNP-r_g with fat mass (-0.33 ± 0.03; $P = 7.23 \times 10^{-25}$), fat-free mass (-0.12 ± 0.03; $P = 4.65 \times 10^{-5}$), BMI (-0.32 ± 0.03; $P = 8.93 \times 10^{-25}$), obesity (-0.22 ± 0.03; $P = 2.96 \times 10^{-11}$), type 2 diabetes (-0.22 ± 0.05; $P = 3.82 \times 10^{-5}$), fasting insulin (-0.24 ± 0.06; $P = 2.31 \times 10^{-5}$), insulin resistance (-0.29 ± 0.07; $P = 2.83 \times 10^{-5}$), and leptin (-0.26 ± 0.06; $P = 4.98 \times 10^{-5}$), and a significant positive SNP-r_g with HDL cholesterol (0.21 ± 0.04; $P = 3.08 \times 10^{-7}$).

Systems biology analyses of our results revealed preliminarily interesting results (**Methods, Supplementary Tables 11-13, Supplementary Figs. 7-15**). Gene-wise analysis with MAGMA prioritized 79 Bonferroni-significant genes, most within the multigenic locus on chr3 (**Supplementary Table 11**). MAGMA indicated an association with *NCAM1* **(Supplementary Table 11)** the expression of which increases in response to food restriction in a rodent activity-based anorexia nervosa model[25]. Partitioned heritability analysis showed, as with other GWAS[26], considerable enrichment of SNP-h^2 in conserved regions (fold enrichment = 24.97, s.e. = 3.29, $P = 3.32 \times 10^{-11}$; **Supplementary Fig. 7)**[27]. Cell type group-specific annotations revealed that the overall SNP-h^2 is significantly enriched for CNS tissue (**Supplementary Fig. 8**). One biological pathway was significant: GO:positive_regulation_of_embryonic_development (32 genes, $P = 1.39 \times 10^{-7}$; **Supplementary Table 12**), which contains two Bonferroni-significant genes on

448 chr3, *CTNNB1* and *DAG1*. *CTNNB1* encodes catenin beta-1, which is part of adherens junctions,
449 and *DAG1* encodes dystroglycan, a receptor which binds extracellular matrix proteins[28]. *DAG1*
450 falls within locus 1 (47.5-51.3 Mb). This pathway points to a potential role of developmental
451 processes in the etiology of this complex phenotype (although this is currently speculative).
452 Genes associated with anorexia nervosa were enriched for expression in most brain tissues,
453 particularly the cerebellum, which has a notably high proportion of neurons[29] (**Supplementary**
454 **Fig. 9**). Among 24 brain cell types from mouse brain, significant enrichment was found for
455 medium spiny neurons and pyramidal neurons from hippocampal CA1 (**Supplementary Fig.**
456 **10**). Both medium spiny and pyramidal neurons are linked to feeding behaviors including food
457 motivation and reward[30,31] (**Supplementary Results**). Using PrediXcan (**Supplementary**
458 **Methods)**, 36 genes were predicted to be differentially expressed in GTEx tissues or blood
459 (**Supplementary Table 13**) with the expression of *MGMT* predicted to be downregulated in the
460 caudate. We cautiously note that these results represent the first indications of specific pathways,
461 tissues, and cell types that may mediate genetic risk for anorexia nervosa.
462 Because low BMI is pathognomonic of anorexia nervosa, we investigated the extent to
463 which genetic variants associated with BMI accounted for genetic correlations with metabolic
464 and anthropometric traits. First, covarying for the genetic associations of BMI (**Methods**) led to
465 a mild but statistically non-significant attenuation of the SNP-r_g between anorexia nervosa and
466 fasting insulin, leptin, insulin resistance, type 2 diabetes, and HDL cholesterol (**Supplementary**
467 **Tables 14-15**), suggesting that anorexia nervosa shares genetic variation with these metabolic
468 phenotypes that may be independent of BMI. Second, we investigated bidirectional causality
469 using generalized summary data-based Mendelian randomization[18]. GSMR analyses indicate a
470 significant bidirectional causal relationship such that anorexia nervosa risk-increasing alleles

may increase risk for low BMI and BMI-lowering alleles may increase the risk of anorexia nervosa (**Supplementary Table 16**). It is important to note that having only eight genome-wide significant loci for anorexia nervosa render this analysis marginally powered in the direction of anorexia nervosa to BMI, although this analysis is well powered in the direction of BMI to anorexia nervosa.

Replication is challenging with GWAS of low prevalence conditions like anorexia nervosa, as replication samples must be sufficiently powered to detect the initial findings. We included all available samples in our analysis to maximize chances of reaching the GWAS inflection point, after which there might be a linear increase in "hits"[32]. The PRS leave-one-out analyses provide evidence of replication by demonstrating a higher burden of anorexia nervosa common risk variants in cases, compared with controls, across all the cohorts (**Supplementary Fig. 16**).

In conclusion, we report multiple genetic loci alongside promising clinical and functional analyses and enrichments. The increased sample size in the present GWAS has allowed us to characterize more fully the metabolic contribution to anorexia nervosa than our previous report[10] by revealing significant r_g with metabolism related phenotypes including glycemic and anthropometric traits and by demonstrating that the effect is robust to correction for the effects of common variants significantly associated with BMI. Low BMI has traditionally been viewed as a consequence of the psychological features of anorexia nervosa (i.e., drive for thinness and body dissatisfaction). This perspective has failed to yield interventions that reliably lead to sustained weight gain and psychological recovery[7]. Fundamental metabolic dysregulation may contribute to the exceptional difficulty that individuals with anorexia nervosa have in maintaining a healthy BMI (even after therapeutic renourishment). Our results encourage consideration of both

metabolic and psychological drivers of anorexia nervosa when exploring new avenues for treating this frequently lethal illness.

URLs. GCTA, http://cnsgenomics.com/software/gcta; GSMR, http://cnsgenomics.com/software/gsmr; LDSC, https://github.com/bulik/ldsc; MAGMA, http://ctg.cncr.nl/software/magma.

Acknowledgements

Grant support for ANGI, the PGC-ED, and its component groups is shown in **Supplementary Table 17**. We thank all study volunteers, study coordinators, and research staff who enabled this study. ANGI: The Anorexia Nervosa Genetics Initiative was an initiative of the Klarman Family Foundation. Additional support was offered by the National Institute of Mental Health. We acknowledge support from the North Carolina Translational and Clinical Sciences Institute (NC TraCS), the Carolina Data Warehouse, and the Foundation of Hope, Raleigh, North Carolina. PGC: We are deeply indebted to the investigators who +comprise the PGC, and to the hundreds of thousands of individuals who have shared their life experiences with PGC investigators and the contributing studies. We are grateful to the Children's Hospital of Philadelphia (CHOP), the Price Foundation Collaborative Group (PFCG), Genetic Consortium for Anorexia Nervosa (GCAN), Wellcome Trust Case-Control Consortium-3 (WTCCC-3), the Lundbeck Foundation Initiative for Integrative Psychiatric Research (iPSYCH), the QSkin Sun and Health Study, Riksät (Swedish National Quality Register for Eating Disorders), the Stockholm Center for Eating Disorders (SCÄ), LifeGene, the UK Biobank, and all PGC-ED members for their support in providing individual samples used in this study. We thank SURFsara (http://www.surf.nl) for support in using the Lisa Compute Cluster. We thank M. Lam for Ricopili consultation. This

study also represents independent research partly funded by the English National Institute for Health Research (NIHR) Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the English Department of Health and Social Care. High performance computing facilities were funded with capital equipment grants from the GSTT Charity (TR130505) and Maudsley Charity (980). Research reported in this publication was supported by the National Institute of Mental Health of the US National Institutes of Health under Award Number U01MH109514. The content is solely the responsibility of the authors and does not necessarily represent the official views of the US National Institutes of Health.

Author contributions

C.M.B. and P.F.S. conceived and designed the study. L.T., C.M.B., and G.B. performed overall study coordination. C.M.B. was lead PI of ANGI. P.F.S. was Co-Investigator of ANGI. N.G.M., M.L., and P.B.M. were site PIs of ANGI. H.J.W., Z.Y., J.R.I.C., C.H., J.B., H.A.G., S.Y., V.M.L., M.M., P.G-R. and S.E.M. performed the statistical analyses. H.J.W., Z.Y., C.H., J.R.I.C., H.A.G., J.B., A.H., P.G-R., P.F.S., G.B. and C.M.B. comprised the writing group. C.M.B. and G.B. were PGC-ED co-chairs. S.R. provided statistical consultation. A.H. assisted with data interpretation. A.W.B., C.M.B., J.J., M.K., K.M.K., P.L., G.M., C.N., R.P., L.T., and T.D.W. collected and managed the ANGI samples at sites and assisted with site-specific study co-ordination. A.W.B., J.M.B., H.B., S.C., K.A.H., L.J.H., C.J., A.S.K., W.K., J.M., C.M.O., J.F.P., N.L.P., M.S., T.W., D.C.W., and D.B.W. provided ANGI controls and extra samples. L.E.D provided data expertise. S.G., J.G., A.K.H., A.J., K.M.K., J.T.L., R.P., and L.P. contributed to the ANGI study. S.G., J.G., K.K., J.T.L., M.M., S.M., and L.P. were ANGI site

analysts. K.B.H. and K.L.P. provided additional analysis for some secondary analyses. G.W.M., T.D.W., A.B., P.L., and C.N. were ANGI investigators. J.J. and M.K. assisted with ANGI recruitment in NZ. C.M.B., G.B., and P.F.S. supervised the study. H.J.W., C.M.B., Z.Y., C.H., G.B., J.R.I.C., H.A.G., S.Y., J.B., P.F.S., and P.G. wrote the manuscript. PGC-ED members and other individuals contributed to sample acquisition and made individual data from subjects available: R.A.H.A., L.A. T.A., O.A.A., J.H.B., A.W.B., W.H.B., A.B., I.B., C.B., J.M.B., H.B., G.B., K. B., C.M.B., R.B., M.C., S.C., M.C., J.R.I.C., R.D.C., P.C., S.C., S.C., J.C., U.N.D., O.S.P.D, M.D, G.D., D.D., J.E.D., D.M.D., D.D., C.D., M.D., E.D.M., K.E., S.E., G.E., T.E., X.E., A.F., A.F., F.F., M.M.F., K.F., M.F., L.F., A.J.F., M.F., S.G., I.G., J.G., F.G., S.G., P.G., M.G.M., J.G., S.G., K.A.H., K.H., J.H., J.H., S.G.H., A.K.H., S.H., B.H., W.H., A.H., L.J.H., J.I.H., H.I., H.I., V.J., S.J., C.J., J.J., A.J., A.J., G.K., D.K., A.S.K., J.K,. L.K., A.K., M.J.H.K., W.K., J.L.K., M.K., A.K., K.K., Y.K., L.K., G.S.K., M.C.L, M.L., S.L., R.D.L., P.L., L.L., B.L., J.L., J.L., P.M., M.M., K.M., S.M., C.M., N.G.M., M.M., S.M., P.M., A.M., I.M., N.M., J.M., A.M.M., P.M., P.M., M.A.M., B.N., M.N., C.N., I.N., C.M.O., J.K.O., R.A.O., L.P., A.P., J.P., H.P., N.L.P., J.F.P., D.P., R.R., A.R., N.R., T.R., V.R., S.R., F.R., M.R., A.R., D.R., F.R., P.S., S.W.S., U.S., A.S., J.S., L.S., P.E.S., M.C.T.S.L., A.S., S.S., M.S., P.F.S., B.Ś., J.P.S., I.T., E.T., A.T., F.T., J.T., A.T., M.T., K.T., A.A.V, E.F.V., T.D.W., G.W., E.W., H.J.W., T.W., D.C.W., E.W., D.B.W., G.S., S.Z., and S.Z. All authors critically reviewed the manuscript.

Competing interests

The authors report the following potential competing interests. O.A.A. received a speaker's honorarium from Lundbeck. G.B. received grant funding and consultancy fees in preclinical genetics from Eli Lilly, consultancy fees from Otsuka and has received honoraria from Illumina.

C.M.B. is a grant recipient from Shire Pharmaceuticals and served on Shire Scientific Advisory Board; she receives author royalties from Pearson. D.D. served as a speaker and on advisory boards, and has received consultancy fees for participation in research from various pharmaceutical industry companies including: AstraZeneca, Boehringer, Bristol Myers Squibb, Eli Lilly, Genesis Pharma, GlaxoSmithKline, Janssen, Lundbeck, Organon, Sanofi, UniPharma, and Wyeth; he has received unrestricted grants from Lilly and AstraZeneca as director of the Sleep Research Unit of Eginition Hospital (National and Kapodistrian University of Athens, Greece). J.I.H. has received grant support from Shire and Sunovion, and has received consulting fees from DiaMentis, Shire, and Sunovion. A.S.K. is a member of the Shire Canadian BED Advisory Board and is on the steering committee for the Shire B/educated Educational Symposium: June 15-16, 2018. J.L.K. served as an unpaid member of the scientific advisory board of AssurexHealth Inc. M.L. declares that, over the past 36 months, he has received lecture honoraria from Lundbeck and served as scientific consultant for EPID Research Oy. No other equity ownership, profit-sharing agreements, royalties, or patent. P.F.S. is on the Lundbeck advisory committee and is a Lundbeck grant recipient; he has served on the scientific advisory board for Pfizer, has received a consultation fee from Element Genomics, and a speaker reimbursement fee from Roche. J.T. has received an honorarium for participation in an EAP meeting and has received royalties from several books from Routledge, Wiley, and Oxford University press. T.W. has acted as a lecturer and scientific advisor to H. Lundbeck A/S. All other authors have no conflicts of interest to disclose.

Additional information

Correspondence and requests for materials should be addressed to C.M.B. or G.B.

References

1. Schaumberg, K. *et al.* The science behind the Academy for Eating Disorders' nine truths about eating disorders. *Eur. Eat. Disord. Rev.* **25**, 432-450 (2017).
2. Keski-Rahkonen, A. & Mustelin, L. Epidemiology of eating disorders in Europe: prevalence, incidence, comorbidity, course, consequences, and risk factors. *Curr. Opin. Psychiatry* **29**, 340-345 (2016).
3. Hudson, J.I., Hiripi, E., Pope, H.G. & Kessler, R.C. The prevalence and correlates of eating disorders in the National Comorbidity Survey Replication. *Biol. Psychiatry* **61**, 348-358 (2007).
4. Micali, N. *et al.* Lifetime and 12-month prevalence of eating disorders amongst women in mid-life: a population-based study of diagnoses and risk factors. *BMC Med.* **15**, 12 (2017).
5. Yilmaz, Z., Hardaway, J.A. & Bulik, C.M. Genetics and epigenetics of eating disorders. *Adv. Genomics Genet.* **5**, 131-150 (2015).
6. Arcelus, J., Mitchell, A.J., Wales, J. & Nielsen, S. Mortality rates in patients with anorexia nervosa and other eating disorders: a meta-analysis of 36 studies. *Arch. Gen. Psychiatry* **68**, 724-731 (2011).
7. Watson, H. & Bulik, C. Update on the treatment of anorexia nervosa: review of clinical trials, practice guidelines and emerging interventions. *Psychol. Med.* **43**, 2477-2500 (2013).

8. Kirk, K.M. *et al.* The Anorexia Nervosa Genetics Initiative: study description and sample characteristics of the Australian and New Zealand arm. *Aust. N. Z. J. Psychiatry* **51**, 583-594 (2017).
9. Thornton, L., Munn-Chernoff, M., Baker, J., Juréus, A. & al., e. The Anorexia Nervosa Genetics Initiative (ANGI): Overview and methods. *Contemp. Clin. Trials* **74**, 61-69 (2018).
10. Duncan, L. *et al.* Significant locus and metabolic genetic correlations revealed in genome-wide association study of anorexia nervosa. *Am. J. Psychiatry* **173**, 850-858 (2017).
11. Pickrell, J.K. *et al.* Detection and interpretation of shared genetic influences on 42 human traits. *Nat. Genet.* **48**, 709-717 (2016).
12. Martin, J., Taylor, M.J. & Lichtenstein, P. Assessing the evidence for shared genetic risks across psychiatric disorders and traits. *Psychol. Med.* **48**, 1759-1774 (2018).
13. Boraska, V. *et al.* A genome-wide association study of anorexia nervosa. *Mol. Psychiatry* **19**, 1085-1094 (2014).
14. Sudlow, C. *et al.* UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med.* **12**, e1001779 (2015).
15. Yang, J. *et al.* Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nat. Genet.* **44**, 369-375 (2012).
16. Fromer, M. *et al.* Gene expression elucidates functional impact of polygenic risk for schizophrenia. *Nat. Neurosci.* **19**, 1442-1453 (2016).
17. GTEx Consortium. Genetic effects on gene expression across human tissues. *Nature* **550**, 204-213 (2017).

18. Zhu, Z. *et al.* Causal associations between risk factors and common diseases inferred from GWAS summary data. *Nat. Commun.* **9**, 224 (2018).

19. Bulik-Sullivan, B.K. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* **47**, 291-295 (2015).

20. Bulik-Sullivan, B. *et al.* An atlas of genetic correlations across human diseases and traits. *Nat. Genet.* **47**, 1236-1241 (2015).

21. Cederlöf, M. *et al.* Etiological overlap between obsessive-compulsive disorder and anorexia nervosa: a longitudinal cohort, multigenerational family and twin study. *World Psychiatry* **14**, 333-338 (2015).

22. Kaye, W.H., Bulik, C.M., Thornton, L., Barbarich, N. & Masters, K. Comorbidity of anxiety disorders with anorexia and bulimia nervosa. *Am. J. Psychiatry* **161**, 2215-2221 (2004).

23. Dalle Grave, R., Calugi, S. & Marchesini, G. Compulsive exercise to control shape or weight in eating disorders: prevalence, associated features, and treatment outcome. *Compr. Psychiatry* **49**, 346-352 (2008).

24. 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).

25. Ho, E.V., Klenotich, S.J., McMurray, M.S. & Dulawa, S.C. Activity-based anorexia alters the expression of BDNF transcripts in the mesocorticolimbic reward circuit. *PLoS One* **11**, e0166756 (2016).

26. Finucane, H.K. *et al.* Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nat. Genet.* **47**, 1228-1235 (2015).

27. Lindblad-Toh, K. *et al.* A high-resolution map of human evolutionary constraint using 29 mammals. *Nature* **478**, 476 (2011).

28. Bello, V. *et al.* The dystroglycan: nestled in an adhesome during embryonic development. *Dev. Biol.* **401**, 132-142 (2015).

29. Azevedo, F.A.C. *et al.* Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain. *J. Comp. Neurol.* **513**, 532-541 (2009).

30. O'Connor, E.C. *et al.* Accumbal D1R neurons projecting to lateral hypothalamus authorize feeding. *Neuron* **88**, 553-564 (2015).

31. Kim, J., Zhang, X., Muralidhar, S., LeBlanc, S.A. & Tonegawa, S. Basolateral to central amygdala neural circuits for appetitive behaviors. *Neuron* **93**, 1464-1479.e5 (2017).

32. Levinson, D.F. *et al.* Genetic studies of major depressive disorder: why are there no genome-wide association study findings and what can we do about it? *Biol. Psychiatry* **76**, 510-512 (2014).

Figure Titles and Captions

Figure 1. **The Manhattan plot for the primary genome-wide association meta-analysis of anorexia nervosa with 33 case-control samples (16,992 cases and 55,525 controls of European descent).** The $-\log_{10}(P)$ values for the association tests (two-tailed) are shown on the y-axis and the chromosomes are ordered on the x-axis. Eight genetic loci surpassed genome-wide significance ($-\log_{10}(P) > 7.3$). The lead variant is indicated by a diamond and green circles show the variants in linkage-disequilibrium. The blue and red colors differentiate adjacent chromosomes.

Figure. 2. Bonferroni-significant genetic correlations (SNP-r_gs) and standard errors (error bars) between anorexia nervosa and other phenotypes as estimated by LD score regression. Only traits with significant *P* values following Bonferroni correction are shown. Correlations with 447 phenotypes were tested (Bonferroni-corrected significance threshold $P > 1.11 \times 10^{-4}$). Complete results are shown in Table S10. PGC = Psychiatric Genomics Consortium, UKB = UK Biobank, HOMA-IR = Homeostatic model assessment - insulin resistance.

Table 1. Newly associated genome-wide significant loci for anorexia nervosa

Locus	Chr	Basepair region		Lead SNP	BP	*P*	A1/A2	OR	s.e.	Freq	Type	Number of genes	Nearest gene
		range left	range right										
1	3	47588253	51368253	rs9821797	48718253	6.99E-15	A/T	1.17	0.02	0.12	multigenic	111	*NCKIPSD*
2	11	114997256	115424956	rs6589488	115096956	6.31E-11	A/T	1.14	0.02	0.13	single-gene	1	*CADM1*
3	2	53881813	54362813	rs2287348	54039813	5.62E-09	T/C	1.11	0.02	0.16	multigenic	13	*ASB3, ERLEC1*
4	10	131269764	131463964	rs2008387	131448764	1.73E-08	A/G	1.08	0.01	0.33	single-gene	2	*MGMT*
5	3	70670750	71074150	rs9874207	71019750	2.05E-08	C/T	1.08	0.01	0.49	single-gene	2	*FOXP1*
6	1	96699455	97284455	rs10747478	96901455	3.13E-08	T/G	1.08	0.01	0.41	single-gene	2	*PTBP2*
7	5	24945845	25372845	rs370838138	25081845	3.17E-08	G/C	1.08	0.01	0.56	intergenic	0	*CDH10*
8	3	93968107	95059107	rs13100344	94605107	4.21E-08	T/A	1.08	0.01	0.54	intergenic	2	*NSUN3*

Note. Shown are the results of the GWAS meta-analysis of anorexia nervosa (16,992 cases and 55,552 controls) which detected eight genome-wide significant loci. All of the eight loci are novel. Chr (chromosome) and Region (hg19) are shown for SNPs with $P < 1e\text{-}05$ and linkage-disequilibrium (LD) $r^2 > 0.1$ with the most associated "lead" SNP, the location of which is given in BP (basepair). A1/A2 refers to Allele 1/Allele 2 and OR and s.e. are the odds ratio and standard error for the association between A1 and the phenotype. Freq is the frequency of A1 in controls. Number of genes was determined by genomic location, adult brain eQTL, regulatory chromatin interactions, and MAGMA gene-wise analysis (see Methods). Nearest gene is the nearest gene within the region of LD "friends" of the lead variant (LD-$r^2 > 0.6$ +/- 500 Kb). The meta-analysis was restricted to variants with minor allele frequency (MAF) ≥ 0.01 and information quality (INFO) score ≥ 0.70. All loci were confirmed via forest plots based on consistent direction of effect in the majority of cohorts and via region plots whereby neighboring LD "friends" were required to show a similar effect. Chromosome X was analyzed but had no loci that reached genome-wide significance. Note that although lead variants are annotated to the nearest gene, this does not mean that the gene listed is a causal gene.

Methods

Samples and study design. Thirty-three datasets with 16,992 anorexia nervosa cases and 55,525 controls were included in the primary GWAS. We included individuals from the Eating Disorders Working Group of the Psychiatric Genomics Consortium (PGC-ED) Freeze 1[10]; newly collected samples from the Anorexia Nervosa Genetics Initiative (ANGI)[8,9]; archived samples from the Genetic Consortium for Anorexia Nervosa (GCAN)/Wellcome Trust Case Control Consortium-3 (WTCCC3)[13]; anorexia nervosa samples from UK Biobank[14]; and additional controls from Poland. Case definitions established a lifetime diagnosis of anorexia nervosa via hospital or register records, structured clinical interviews, or on-line questionnaires based on standardized criteria (DSM-III-R, DSM-IV, ICD-8, ICD-9, or ICD-10), whereas in the UK Biobank cases self-reported a diagnosis of anorexia nervosa. Controls were carefully matched for ancestry, and some, but not all control cohorts were screened for lifetime eating and/or some or all psychiatric disorders. Given the relative rarity of anorexia nervosa, large unscreened control cohorts were deemed appropriate for inclusion[33].

The cohorts are detailed in the Supplement. Ethical approvals and consent forms were reviewed and archived for all participating cohorts (see Supplementary Methods ANGI-DK for Danish methods). Summary details about ascertainment (Supplementary Table 2), the genotyping platforms used (Supplementary Table 3), and genotype availability (Supplementary Table 4) can be accessed in the Supplement.

Statistical analysis. Data processing and analysis were done on the Lisa Compute Cluster hosted by SURFsara (http://www.surfsara.nl) and the GenomeDK high-performance computing cluster (http://genome.au.dk).

Meta-analysis of genome-wide association data. Quality control (QC), imputation, GWAS, and meta-analysis followed the standardized pipeline of the PGC, Ricopili (Rapid Imputation Consortium Pipeline). Ricopili versions used were 2017_Oct_11.002 and 2017_Nov_30.003. QC included SNP and sample QC, population stratification and ancestry outliers, and familial and cryptic relatedness. Further information about the Ricopili pipeline is available from the website (https://sites.google.com/a/broadinstitute.org/ricopili) and GitHub repository (https://github.com/Nealelab/ricopili/tree/master/rp_bin). Further details of the QC procedures can be found in the Supplementary Methods.

Imputation. Imputation of SNPs and insertions-deletions was based on the 1000 Genomes Phase 3 (http://www.internationalgenome.org) data[34].

GWAS. GWASs were conducted separately for each cohort using imputed variant dosages and an additive model. Covariates nominally associated with the phenotype in univariate analysis ($P < 0.05$) and five ancestry PCs were included in GWAS (Supplementary Table 18). These analyses used the tests and methods programmed in the Ricopili pipeline. Genomic inflation factors (λ) of the final datasets indicated no evidence of inflation of the test statistics due to population stratification or other sources (Supplementary Table 1). The 33 cohorts were meta-analyzed with the Ricopili pipeline which uses an inverse-variance weighted fixed-effect model. We filtered our GWAS results with minor allele frequency (MAF) ≥ 0.01 and INFO score ≥ 0.70 (indicating "high-quality").

Analysis of chrX. Several cohorts in the primary GWAS did not have X chromosome variant data, specifically, some GCAN-based cohorts (*fre1*, *ukd1*, *usa1*, *gns2*) and were excluded. Imputation was performed separately from the autosome[35]. ChrX variants in the pseudoautosomal regions were excluded prior to imputation. SNPs exceeding MAF and INFO score thresholds of 0.01 and 0.70 were retained and analysis was performed with PLINK v1.9 (https://www.cog-genomics.org/plink2) and Ricopili.

Female-only GWAS. A supplementary GWAS analysis was conducted on females only to determine the similarity of the results to the primary GWAS analysis which included both females and males. The cohorts that did not have chrX variants to verify sex could not be included (*fre1*, *ukd1*, *usa1*, *gns2*).

Distance- and LD-based clumping. The GWAS results implicate genomic regions ("loci"). To define a locus, (1) SNPs that met the genome-wide significant threshold of $P < 5 \times 10^{-8}$ were identified; (2) clumping was used to convert significant SNPs to regions. The SNP with the smallest *P* value in a genomic window was kept as the index SNP and SNPs in high linkage disequilibrium (LD) with the index SNP defined the left and right end of the region (SNPs with $P < 0.0001$ and $r^2 > 0.1$ within 3 Mb windows); (3) partially or wholly overlapping clumps within 50 Kb were identified and merged into one region; (4) only loci with additional evidence of association from variants in high LD as depicted by regional plots were retained; further, forest plots needed to confirm the associations based on the majority of cohorts; and (5) conditional analyses were conducted to identify SNPs with associations independent of the top SNP within the genomic chunk of interest.

752

753 *Annotation.* Genome-wide significant loci were annotated with RegionAnnotator
754 (https://github.com/ivankosmos/RegionAnnotator) to identify known protein-coding genes
755 within loci (Supplementary Table 6).

756

757 *Conditional and joint analysis.* Conditional and joint analysis was conducted using GCTA-
758 COJO[15]. GCTA-COJO investigates every locus with a joint combination of independent markers
759 via a genome-wide SNP selection procedure. It takes into account the LD correlations between
760 SNPs and runs a conditional and joint analysis on the basis of conditional *P* values. After a
761 model optimizing process, the joint effects of all selected SNPs are calculated. The largest
762 subsample from our GWAS (*sedk*) was used to approximate the underlying LD structure of the
763 investigated lead SNPs. The conditional regression was performed in a stepwise manner using
764 the GCTA software[36]. We analyzed SNPs that had a $P < 5 \times 10^{-8}$ (Supplementary Table 5).

765

766 *Multi-trait-based conditional and joint analysis.* To separate marginal effects from conditional
767 effects (i.e., the effect of a risk factor on an outcome controlling for the effect of another risk
768 factor), we performed a multi-trait-based conditional and joint analysis (GCTA-mtCOJO)[18] using
769 an extension of the GCTA software[36] (Supplementary Table 8). This method uses summary-level
770 data to perform the conditional analysis. We conditioned the results of our anorexia nervosa
771 GWAS on GWAS results for education years[37], type 2 diabetes[38], HDL cholesterol[39], BMI
772 (Hübel, Gaspar, Coleman, Hanscombe, Purves…Breen, unpublished report), schizophrenia[40],
773 and neuroticism[41]. We again used the individual-level genotype data from our largest cohort
774 (*sedk*) to approximate the underlying LD structure. As a first step, the method performs a

775 generalized summary data-based Mendelian randomization (GSMR) analysis to test for causal
776 association between the outcome (i.e., anorexia nervosa) and the risk factor (e.g., schizophrenia).
777 We removed potentially pleiotropic SNPs from this analysis by the heterogeneity in dependent
778 instruments (HEIDI) outlier method[18]. Pleiotropy is the phenomenon when a single locus directly
779 affects several phenotypes. The power of the HEIDI-outlier method is dependent on sample size
780 of the GWAS. Pleiotropic SNPs are defined as the SNPs that show an effect on the outcome that
781 significantly diverges from that expected under a causal model. Second, the GCTA-mtCOJO
782 method calculates the genetic correlation between the exposure and the outcome using linkage
783 disequilibrium score regression (LDSC) to adjust for genetic overlap[19,20]. It also uses the
784 intercept of the bivariate LDSC to account for potential sample overlap[19,20]. As a result, GCTA-
785 mtCOJO calculates conditional betas, conditional standard errors, and conditional *P* values.
786 Subsequently, we clumped the conditional GWAS results using the standard PLINK v1.9[42]
787 algorithm (SNPs with $P < 0.0001$ and $r^2 > 0.1$ within 3 Mb windows) to investigate if any of the
788 genome-wide significant loci showed dependency on genetic variation associated with other
789 phenotypes. As stated in Zhu et al.[18], the GCTA-mtCOJO analysis requires the estimates of b_{xy}
790 of the covariate risk factors on the target risk factor and disease, r_g of the covariate risk factors,
791 heritability (h^2_{snp}) for the covariate risk factors, and the sampling covariance between SNP
792 effects estimated from potentially overlapping samples.
793
794 *eQTL and Hi-C interactions.* Although GWAS findings are informative genome-wide,
795 identifying strong hypotheses about their connections to specific genes is not straightforward.
796 The lack of direct connections to genes constrains subsequent experimental modeling and efforts
797 to develop improved therapeutics. Genomic location is often used to connect significant SNPs to

798 genes, but this is problematic because GWAS loci usually contain many correlated and highly
799 significant SNP associations over hundreds of Kb. Moreover, the three-dimensional (3D)
800 arrangement of chromosomes in cell nuclei enables regulatory interactions between genomic
801 regions located far apart[43]. Chromosome conformation capture methods like Hi-C enable
802 identification of 3D interactions *in vivo*[44,45] and can clarify GWAS findings. For example, an
803 intergenic region associated with multiple cancers was shown to be an enhancer for *MYC* via a
804 long-range chromatin loop[46,47], and intronic *FTO* variants are robustly associated with body mass
805 but influence expression of distal genes via long-range interactions[48]. The *Nature* paper of Won
806 et al.[49] used Hi-C to assess the 3D chromatin interactome in fetal brain, and asserted connections
807 of some schizophrenia associations to specific genes.
808 To gain further understanding of 3D chromatin organization of the brain and to evaluate
809 disease relevance, we applied "easy Hi-C"[50] to postmortem samples ($N = 3$ adult temporal
810 cortex). Library quality and yield from eHi-C are comparable to conventional Hi-C but requires
811 much less starting material. Please refer to the following pre-print for details on methodology,
812 data processing, quality control and statistical models used for these analyses[51]. We generated
813 sufficient reads to enable a kilobase resolution map of the chromatin interactome from adult
814 human brain. To our knowledge, these are the deepest Hi-C data on any human tissue (excluding
815 cell lines) as they generated 22.5X as many *cis*-contacts as for the next largest datasets (DLPFC
816 and hippocampus). We generated tissue RNA-seq, total-stranded RNA-seq, ChIP-seq (H3K27ac,
817 H3K4me3, and CTCF), and open chromatin data (ATAC-seq) for adult brain to help interpret the
818 eHi-C results. We also integrated brain expression and eQTL data from GTEx to aid these
819 analyses. The Hi-C analysis is unbiased in that all chromatin interactions that pass a confidence

820 threshold are considered when evaluating the associations between SNPs and genes (i.e., it is not
821 a capture experiment where only "candidate" SNP-to-gene associations are evaluated).
822 Similar to the work by Won *et al*.[49], we used Hi-C data generated from human adult brain
823 to identify genes implicated by three-dimensional functional interactomics (Supplementary Figs.
824 5 a-h). These Hi-C data (N = 3, anterior temporal cortex) contain more than 103K high-
825 confidence, regulatory chromatin interactions[51]. These interactions capture the physical
826 proximity of two regions of the genome in brain nuclei ("anchors", 10 Kb resolution) although
827 they are separated by 20 Kb to 2 Mb in genomic distance. We focused on the regulatory subset
828 of E-P or P-P (E = enhancer, P = promoter) chromatin interactions (with P defined by location of
829 an open chromatin anchor near the transcription start site of an adult brain-expressed transcript
830 and E defined by overlap with open chromatin in adult brain plus either H3K27ac or H3K4me3
831 histone marks). The presence of a regulatory chromatin interaction from a GWAS locus to a gene
832 provides a strong hypothesis about SNP-to-gene regulatory functional interactions.
833
834 *SNP-based heritability estimation.* LDSC software (https://github.com/bulik/ldsc**)** and method
835 were used to estimate SNP-based heritabilities for each cohort and overall[19,20]. We used
836 precomputed LD scores based on the 1000 Genomes Project European ancestry samples[34]
837 directly downloaded from https://github.com/bulik/ldsc. The liability scale estimate assumed a
838 population prevalence of 0.9%-4% for anorexia nervosa[2,3].
839
840 *Within-trait prediction: polygenic risk scoring.* Polygenic leave-one-dataset-out analysis, using
841 PRSice v2.1.3[52]**,** was conducted in the first instance to identify any extreme outlying datasets. In
842 addition, it enabled the evaluation of the association between anorexia nervosa polygenic risk

score (PRS) and anorexia nervosa risk in an independent cohort as a means of replication of the GWAS results. We derived a PRS for anorexia nervosa from the meta-analysis of all datasets except for the target cohort, then applied the PRS to the target cohort to predict affected status (Supplementary Fig. 16). Logistic regression was performed, including as covariates the first five ancestry components and any other PCs significantly associated with the phenotype in the target cohort, and the target cohort was split into deciles based on anorexia nervosa PRS, with decile 1 comprised of those with the lowest anorexia nervosa PRS serving as the referent.

Anorexia nervosa subtype analysis. PRS analyses were conducted with anorexia nervosa subgroups to investigate prediction of case status across the subtypes. For this, we split the anorexia nervosa cases to two groups based on whether binge eating was present. First, GWAS meta-analyses were conducted for (a) anorexia nervosa with binge eating vs controls (2,381 cases and 10,249 controls; $k = 3$ datasets: *aunz*, *chop*, *usa2*) and (b) anorexia nervosa with no binge eating vs controls (2,262 cases and 10,254 controls; $k = 3$ datasets: *aunz*, *chop*, *usa2*). Controls were randomly split between analyses to maintain independence (Supplementary Fig. 6). Genetic correlation analysis using LDSC[19,20] was conducted to examine the potential genetic overlap of the two anorexia nervosa subtypes (Supplementary Table 9). Second, using PRSice[52], we calculated PRS for each anorexia nervosa subtype separately in the three target cohorts for which anorexia nervosa subtype data were available. Finally, mean PRS scores were estimated for each subtype by cohort after accounting for covariates in R. Subtype phenotyping is described in the Supplementary Methods.

865 *Males.* In order to assess whether sex-specific differences in anorexia nervosa genetic risk load
866 exist, we calculated PRS, using PRSice[52], from a GWAS meta-analysis performed on females
867 only (14,898 cases and 27,545 controls) and applied it to a male-only target cohort (447 cases
868 and 20,347 controls) to predict affected status.
869
870 *Cross-trait analysis: genetic correlations.* Common variant-based genetic correlation (SNP-r_g)
871 measures the extent to which two traits or disorders share common genetic variation. SNP-r_g
872 between anorexia nervosa and 447 traits (422 from an internally curated dataset and 25 from
873 LDHub[53]) were tested using GWAS summary statistics via an analytical extension of LDSC[19,20].
874 The sources of the summary statistics files (PMID, DOI, or unpublished results) used in the
875 LDSC are provided in Supplementary Table 10. When there were multiple summary statistics
876 files available for a trait, significant SNP-r_g reported in the main text were chosen based on the
877 largest sample size and/or matching ancestry with our sample (i.e., European ancestry).
878 Genetic correlations with anorexia nervosa corrected for BMI were carried out to
879 investigate whether the observed genetic correlations between anorexia nervosa and metabolic
880 phenotypes were attributable to BMI or partially independent. We used GCTA-mtCOJO[18] to
881 perform a GWAS analysis for anorexia nervosa conditioning on BMI using BMI summary data
882 from our UK Biobank analysis (described in the next section) to derive anorexia nervosa GWAS
883 summary statistics corrected for the common variants genetic component of BMI
884 (Supplementary Tables 14 and 15).
885
886 *GWAS of related traits in UK Biobank.* Several GWAS analyses were carried out for traits in UK
887 Biobank to allow us to investigate body composition genetics in healthy individuals without a

888 psychiatric disorder, a weight-altering disorder, or who were taking weight-altering medication.
889 We also used UK Biobank to carry out GWAS of physical activity level, anxiety, and
890 neuroticism. For details see the Supplementary Methods.
891
892 *Generalized summary data-based Mendelian randomization (GSMR).* We performed two
893 bidirectional GSMR analyses[18] to test for the causal association between first, BMI and anorexia
894 nervosa, and second, Type 2 diabetes and anorexia nervosa, using an extension of the GCTA
895 software[36] (Supplementary Table 16). We used the individual-level genotype data from our
896 largest cohort (*sedk*) to approximate the underlying LD structure. We removed potentially
897 pleiotropic SNPs from this analysis by the HEIDI outlier method[18]. Pleiotropic SNPs are defined
898 as the SNPs which show an effect on the outcome that significantly diverges from the one
899 expected under a causal model. The method uses the intercept of the bivariate LD score
900 regression to account for potential sample overlap[19,20]. As a rule of thumb GSMR requires
901 GWAS to have at least ten genome-wide significant hits. We lowered the threshold for this
902 requirement to eight SNPs in our analyses of anorexia nervosa as an exposure and BMI or Type
903 2 diabetes as an outcome. Results, therefore, should be interpreted cautiously. We, furthermore,
904 investigated bidirectional conditional effects between BMI or Type 2 diabetes and anorexia
905 nervosa. We used GCTA-mtCOJO to perform a GWAS analysis for anorexia nervosa
906 conditioning on (1) BMI using summary data from our UK Biobank analysis and (2) Type 2
907 diabetes using summary data[38]. Our anorexia nervosa GWAS and the BMI and Type 2 diabetes
908 GWASs are based on independent samples. For BMI, we also re-ran the GSMR analysis using
909 the BMI-adjusted anorexia nervosa GWAS summary data from the GCTA-mtCOJO analysis.
910

Gene-wise analysis. MAGMA v1.06[54] was used to perform a gene-wise test of association with anorexia nervosa based on GWAS summary statistics. MAGMA generates gene-based *P* values by combining SNP-based *P* values within a gene while accounting for LD. In order to include regulatory regions, SNPs are mapped to genes within a 35 kb upstream and 10 kb downstream window, and the gene *P* value is obtained using the "multi=snp-wise" model, which aggregates mean and top SNP association models. We tested 19,846 ENSEMBL genes, including the X chromosome (Supplementary Table 11). As reference panel for the underlying LD structure we used 1000 Genomes European data phase 3[34].

Pathway analysis. MAGMA v1.06[54] was used to perform a competitive pathway analysis, testing whether genes associated with anorexia nervosa were more enriched in a given pathway than all other pathways. The analysis included chrX. Biological pathways were defined using gene ontology pathways and canonical pathways from MSigDB v6.1[55], and psychiatric pathways mined from the literature. A total 7,268 pathways were tested (Supplementary Table 12).

Partitioned heritability. Partitioned heritability was investigated using stratified LDSC[26] which estimates the per-SNP contribution to overall SNP-heritability (SNP-h^2) across various functional annotation categories of the genome (Supplementary Fig. 7). It accounts for linked markers and uses a 'full baseline model' of 24 annotations that are not specific to any cell type. We excluded the MHC region in our analysis. SNP-h^2 can be partitioned in two different ways: a non-cell type-specific and a cell type-specific manner. Partitioned heritability analysis was used to test for cell type-specific enrichment in the GWAS of anorexia nervosa among 10 cell type groups; adrenal and pancreas, cardiovascular, central nervous system (CNS), connective and

bone, gastrointestinal, immune and hematopoietic, kidney, liver, skeletal muscle, and other tissue, which includes adipose tissue (Supplementary Fig. 8).

Gene expression. We conducted a series of gene expression analyses as detailed in the Supplementary Methods.

Reporting summary

Further information on research design is available in the Life Science Reporting Summary linked to this article.

Data availability

The Psychiatric Genomics Consortium's (PGC) policy is to make genome-wide summary results public. Genome-wide summary statistics for the meta-analysis are freely downloadable from PGCs download website (http://www.med.unc.edu/pgc/results-and-downloads). Individual-level data are deposited in dbGaP (http://www.ncbi.nlm.nih.gov/gap) for ANGI-ANZ/SE/US (accession number phs001541.v1.p1) and CHOP/PFCG (accession number phs000679.v1.p1). ANGI-DK individual-level data are not available in dbGaP owing to Danish laws, but are available via collaboration with PIs. GCAN/WTCCC3 individual-level data are deposited in EGA (https://www.ebi.ac.uk/ega) (accession number EGAS00001000913) with the exception of Netherlands and US/Canada, which are available via collaboration with PIs. UK Biobank

individual-level data can be applied for on the UK Biobank website (http://www.ukbiobank.ac.uk/register-apply).

References

33. Moskvina, V., Holmans, P., Schmidt, K.M. & Craddock, N. Design of case-controls studies with unscreened controls. *Ann. Hum. Genet.* **69**, 566-576 (2005).

34. 1000 Genomes Project Consortium *et al.* An integrated map of genetic variation from 1,092 human genomes. *Nature* **491**, 56-65 (2012).

35. Chang, D. *et al.* Accounting for eXentricities: analysis of the X chromosome in GWAS reveals X-linked genes implicated in autoimmune diseases. *PLoS One* **9**, e113684 (2014).

36. 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).

37. Okbay, A. *et al.* Genome-wide association study identifies 74 loci associated with educational attainment. *Nature* **533**, 539-542 (2016).

38. Morris, A.P. *et al.* Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. *Nat. Genet.* **44**, 981-990 (2012).

39. Teslovich, T.M. *et al.* Biological, clinical and population relevance of 95 loci for blood lipids. *Nature* **466**, 707-713 (2010).

40. Schizophrenia Working Group of the Psychiatric Genomics Consortium *et al.* Biological insights from 108 schizophrenia-associated genetic loci. *Nature* **511**, 421-427 (2014).

41. Hübel, C. *et al.* Genomics of body fat percentage may contribute to sex bias in anorexia nervosa. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* **in press**.

42. Chang, C.C. *et al.* Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**, 7 (2015).

43. Dekker, J. Mapping the 3D genome: aiming for consilience. *Nat. Rev. Mol. Cell Biol.* **17**, 741-742 (2016).

44. Dekker, J. Gene regulation in the third dimension. *Science* **319**, 1793-1794 (2008).

45. Ethier, S.D., Miura, H. & Dostie, J. Discovering genome regulation with 3C and 3C-related technologies. *Biochim. Biophys. Acta* **1819**, 401-410 (2012).

46. Pomerantz, M.M. *et al.* The 8q24 cancer risk variant rs6983267 shows long-range interaction with MYC in colorectal cancer. *Nat. Genet.* **41**, 882-884 (2009).

47. Wright, J.B., Brown, S.J. & Cole, M.D. Upregulation of c-MYC in cis through a large chromatin loop linked to a cancer risk-associated single-nucleotide polymorphism in colorectal cancer cells. *Mol. Cell Biol.* **30**, 1411-1420 (2010).

48. Smemo, S. *et al.* Obesity-associated variants within FTO form long-range functional connections with IRX3. *Nature* **507**, 371-375 (2014).

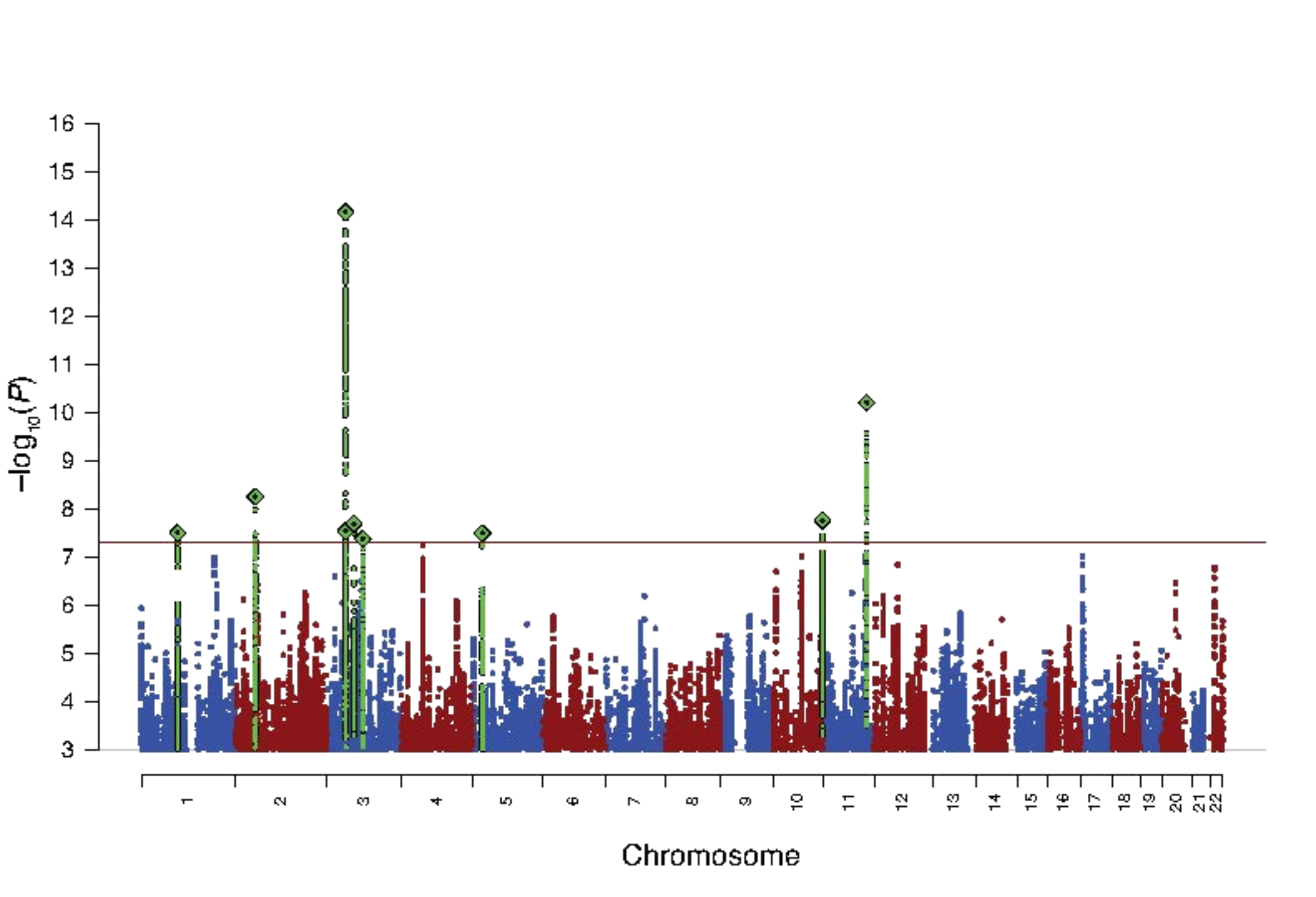
49. Won, H. *et al.* Chromosome conformation elucidates regulatory relationships in developing human brain. *Nature* **538**, 523-527 (2016).

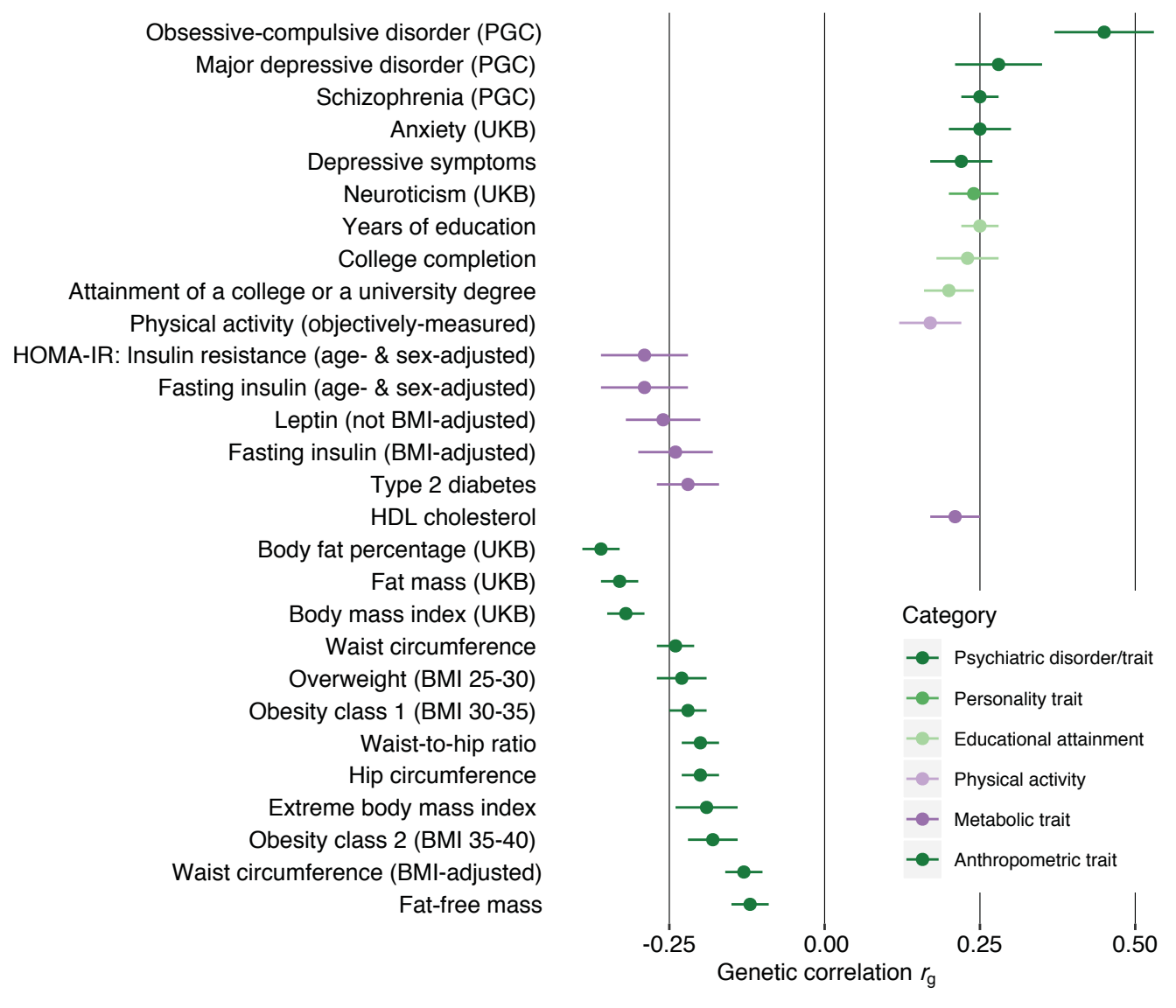
50. Lu, L., Liu, X., Peng, J., Li, Y. & Jin, F. Easy Hi-C: A simple efficient protocol for 3D genome mapping in small cell populations. *bioRxiv*, 245688 (2018).

51. Giusti-Rodriguez, P.M. & Sullivan, P.F. Schizophrenia and a high-resolution map of the three-dimensional chromatin interactome of adult and fetal cortex. *bioRxiv*, 406330 (2018).

52. Euesden, J., Lewis, C.M. & O'Reilly, P.F. PRSice: Polygenic Risk Score software. *Bioinformatics* **31**, 1466-1468 (2015).

53. Zheng, J. *et al.* LD Hub: a centralized database and web interface to perform LD score regression that maximizes the potential of summary level GWAS data for SNP heritability and genetic correlation analysis. *Bioinformatics* **33**, 272-279 (2017).
54. 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).
55. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 15545-15550 (2005).


16
15
14
13
12
11
10
9
8
7
6
5
4
3
$-\log_{10}(P)$
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
Chromosome


Obsessive-compulsive disorder (PGC)
Major depressive disorder (PGC)
Schizophrenia (PGC)
Anxiety (UKB)
Depressive symptoms
Neuroticism (UKB)
Years of education
College completion
Attainment of a college or a university degree
Physical activity (objectively-measured)
HOMA-IR: Insulin resistance (age- & sex-adjusted)
Fasting insulin (age- & sex-adjusted)
Leptin (not BMI-adjusted)
Fasting insulin (BMI-adjusted)
Type 2 diabetes
HDL cholesterol
Body fat percentage (UKB)
Fat mass (UKB)
Body mass index (UKB)
Waist circumference
Overweight (BMI 25-30)
Obesity class 1 (BMI 30-35)
Waist-to-hip ratio
Hip circumference
Extreme body mass index
Obesity class 2 (BMI 35-40)
Waist circumference (BMI-adjusted)
Fat-free mass
Category
Psychiatric disorder/trait
Personality trait
Educational attainment
Physical activity
Metabolic trait
Anthropometric trait
-0.25
0.00
0.25
0.50
Genetic correlation r_g