Genome-wide association studies of binge-eating behaviour and anorexia nervosa yield insights into the unique and shared biology of eating disorder phenotypes

Jet D Termorshuizen PhD[1,*]; Helena L Davies PhD[2,3,4,*]; Sang-Hyuck Lee MSc[2,5,*]; Jessica K Dennis PhD[6,7]; Christopher Hübel MD, PhD[2,8,9]; Jessica S Johnson MPH, MFA[10]; Yi Lu PhD[1]; Melissa A Munn-Chernoff PhD[11]; Triinu Peters PhD[12,13,14]; Baiyu Qi MPH[15]; Katherine E Schaumberg PhD[16,17]; Rebecca H Signer MS[18]; Karanvir Singh MSc[7]; Abigail R ter Kuile PhD[2,5,19]; Laura M Thornton PhD[10]; Jiayi Xu PhD[20]; Shuyang Yao PhD[1]; Zeynep Yilmaz PhD[1,8,10,21]; Ruyue Zhang PhD[1,22]; Johan Zvrskovec PhD[2,5]; Mohamed Abdulkadir PhD[8]; Ziada Ayorech PhD[23]; Elizabeth C Corfield[24,25,26]; Alexandra Havdahl PhD[23,24,25]; Kristi Krebs PhD[27]; Taralynn M Mack[28]; Maria Niarchou PhD[29]; Teemu Palviainen[30]; Julia M Sealock PhD[31,32]; Jessica H Baker PhD[33]; Andrew W Bergen PhD[34,35]; Andreas Birgegård PhD[1]; Vesna Boraska Perica PhD[36]; Katharina Bühren MD[37]; Roland Burghardt MD[38]; Matteo Cassina MD[39]; Giovanni Castellini PhD, MD[40]; Enrico Collantoni MD, PhD[41]; James J Crowley PhD[22,42]; Unna N Danner PhD[43]; Franziska Degenhardt MD[44]; Janiece E DeSocio PhD, RN, PMHNP-BC[45]; Christian Dina PhD[46]; Monika Dmitrzak-Węglarz PhD[47]; Laramie E Duncan PhD[48]; Karin M Egberts MD, PhD[49,50]; Lenka Foretova MD, PhD[51]; Ina Giegling PhD[52]; Fragiskos Gonidakis MD[53]; Scott D Gordon PhD[54]; Jakob Grove PhD[21,55,56,57]; Sébastien Guillaume MD, PhD[58]; Jerry D Guintivano PhD[10,22]; Annette M Hartmann MD[52]; Konstantinos Hatzikotoulas PhD[59]; Stefan Herms MSc[60,61,62]; Hartmut Imgart MD[63]; Susana Jiménez-Murcia PhD[64,65,66,67,68]; Antonio Julià PhD[69]; Gursharan Kalsi PhD[2]; Deborah Kaminská PhD[70]; Leila J Karhunen PhD[71]; Hannah L. Kennedy PhD[72]; Kirsty M Kiezebrink PhD, FHEA, RNutr[73]; Theresa Kolb MSc[74,75]; Janne T Larsen PhD[8,55]; Dong Li PhD[76,77,78]; Lisa Lilenfeld PhD[79]; Mario Maj MD, PhD[80]; Morten Mattingsdal PhD[81,82]; Paolo Meneguzzo MD; PhD[41,83]; Allison L Miller PGDipSc[84]; Karen S Mitchell PhD[85,86]; Alessio Maria Monteleone MD, PhD[87]; Catherine M Olsen PhD, MPH[88]; Leonid Padyukov MD, PhD[89]; Richard Parker BA(Hons)[54]; Michaela A. Pettie PhD[84]; Dalila Pinto PhD[18,90]; Anu Raevuori MD, PhD[91,92]; Samuli Ripatti PhD[30,31,92]; Marion E Roberts PhD[75,93]; Paolo Santonastaso MD[41]; Androula Savva MSc[42]; Ulrike H Schmidt MD, PhD[75]; Alexandra Schosser MD, PhD, MBA[94]; Jochen Seitz MD[14,44]; Lenka LS Slachtova PhD[95]; Agnieszka Slopien MD, PhD[96]; Sandro Sorbi MD[97]; Peter S Straub MS[29]; Jin P Szatkiewicz PhD[22]; Friederike I Tam MD[74]; Elena Tenconi PsyD, PhD[41,83]; Alfonso Tortorella MD[98]; Artemis Tsitsika MD, PhD[99]; Annemarie A van Elburg MD, PhD[43,100]; Gudrun Wagner Dr, MSc, DPO[101]; Hunna J Watson PhD, MPsychClin, MBiostats[10,102]; Roger AH Adan PhD[43,103,104]; Lars Alfredsson PhD[105,106]; Ole A Andreassen MD, PhD[82,107,108]; Helga Ask PhD[23, 24]; Anders D. Børglum MD, PhD[21,55,56]; Harry A Brandt MD[109,110]; David Collier PhD[2]; Steven Crawford MD[110]; Scott Crow MD[111]; Lea K Davis PhD[18,90,112]; Martina de Zwaan MD[113]; George Dedoussis PhD[114]; Danielle M Dick PhD[115]; Stefan Ehrlich MD, PhD[74]; Xavier Estivill MD, PhD[116]; Angela Favaro MD, PhD[41,83]; Fernando Fernández-Aranda PhD[64,65,66,67]; Krista Fischer PhD[27,117]; Andreas J Forstner MD[62,118,119]; Philip Gorwood MD, PhD[120,121]; Hakon Hakonarson MD, PhD[76,78]; Johannes Hebebrand MD[44]; Beate Herpertz-Dahlmann MD[122]; Anke Hinney PhD[12,13]; James I Hudson MD, ScD[123]; Craig Johnson PhD[124]; Jennifer Jordan PhD[72,125]; Allan S Kaplan MD[126]; Jaakko Kaprio MD, PhD[30]; Andreas FK Karwautz MD[127]; Martien JH Kas PhD[105,128]; Walter H Kaye MD[129]; James L Kennedy MD, FRCP(C)[130,131]; Martin A Kennedy PhD[84]; Anna Keski-Rahkonen MD, PhD, MPH[92]; Youl-Ri Kim MD, PhD[132]; Kelly L Klump PhD[133]; Mikael Landén MD, PhD[1,134]; Stéphanie Le Hellard PhD[135]; Kelli Lehto PhD[27]; Qingqin S. Li PhD[136,137]; Jolanta Lissowska PhD[138]; Jurjen J. Luykx MD,PhD[139,140,141]; Sarah L Maguire

BScPsych(Hons), M.A., DCP, PhD[142]; Nicholas G Martin PhD[54]; Manuel Mattheisen MD[143,144]; Sarah E Medland PhD[145,146,147]; Philip Mehler MD, FACP, FAED, CEDS[124,148]; Nadia Micali MD, PhD[3,4,149]; James E Mitchell MD[150]; Palmiero Monteleone MD[151]; Preben Bo Mortensen MD, DrMedSc[8]; Benedetta Nacmias PhD[97]; Roel A Ophoff PhD[152]; Hana Papezova MD, PhD, FAED[70]; Nancy L Pedersen PhD[1]; Liselotte V Petersen PhD[8,55]; Luisa S Rajcsanyi PhD[12,13]; Nicolas Ramoz PhD[153,154]; Ted Reichborn-Kjennerud MD, PhD[24,155]; Valdo Ricca MD[40]; Stephan Ripke MD, PhD[32,156,157]; Dan Rujescu MD[52]; Filip Rybakowski MD, PhD[158]; Stephen W Scherer PhD, FRSC[159,160]; Margarita CT Slof-Op 't Landt PhD[161,162]; Howard Steiger, PhD[163,164]; Patrick F Sullivan MD, FRANZCP[10,22]; Beata Świątkowska PhD[165]; Eric F van Furth PhD[161]; Tracey D Wade PhD[166]; Thomas Werge PhD[4,167]; David C Whiteman MBBS(Hons), PhD, FAFPHM[88]; D. Blake Woodside MD, MSc, FRCPC[130]; Ya-Ke Wu, PhD, RN[10,168]; Stephan Zipfel MD[169,170]; Eating Disorders Working Group of the Psychiatric Genomics Consortium; Estonian Biobank (EstBB); Cynthia M Bulik PhD[1, 10,171,^]; Laura M Huckins PhD[20,^]; Gerome Breen PhD[2,5,^]; Jonathan RI Coleman PhD[2,5,^]

1 Department of Medical Epidemiology and Biostatistics; Karolinska Institutet; Stockholm; Sweden
2 Institute of Psychiatry, Psychology and Neuroscience, Social, Genetic and Developmental Psychiatry (SGDP) Centre; King's College London; London; United Kingdom
3 Center for Eating and Feeding Disorders Research, Mental Health Center Ballerup; Copenhagen University Hospital - Mental Health Services; Copenhagen; Denmark
4 Institute of Biological Psychiatry; Mental Health Center Sct. Hans; Mental Health Services Copenhagen; Roskilde; Denmark
5 National Institute for Health Research Biomedical Research Centre; King's College London and South London and Maudsley National Health Service Trust; London; United Kingdom
6 Department of Medical Genetics; University of British Columbia; Vancouver; British Columbia; Canada
7 Graduate Program in Bioinformatics; University of British Columbia; Vancouver; British Columbia; Canada
8 National Centre for Register-based Research; Aarhus University; Aarhus; Denmark
9 Clinic for Child and Adolescent Psychiatry, Psychotherapy and Psychosomatics; German Red Cross Hospitals Berlin; Berlin; Germany
10 Department of Psychiatry; University of North Carolina at Chapel Hill; Chapel Hill; North Carolina; United States
11 Department of Community, Family, and Addiction Sciences; Texas Tech University; Lubbock; Texas; United States
12 Section for Molecular Genetics in Mental Disorders; LVR University Clinic Essen, University of Duisburg-Essen; Essen; Northrhine-Westfalia; Germany
13 Institute of Sex and Gender-Sensitive Medicine; University Hospital Essen, University of Duisburg-Essen; Essen; Northrhine-Westfalia; Germany
14 Center for Translational Neuro- and Behavioral Sciences; University Hospital Essen, University of Duisburg-Essen; Essen; Northrhine-Westfalia; Germany
15 Department of Epidemiology; University of North Carolina at Chapel Hill; Chapel Hill; North Carolina; United States
16 Department of Psychiatry; University of Wisconsin; Madison; Wisconsin; United States

17 Department of Psychology; University of Texas; Austin; Texas; United States
18 Department of Genetics and Genomic Sciences; Icahn School of Medicine at Mount Sinai; New York; New York; United States
19 Department of Clinical, Educational, and Health Psychology; University College London; London; United Kingdom
20 Department of Psychiatry; Yale University; New Haven; Connecticut; United States
21 Department of Biomedicine; Aarhus University; Aarhus; Denmark
22 Department of Genetics; University of North Carolina at Chapel Hill; Chapel Hill; North Carolina; United States
23 Department of Psychology; PROMENTA Research Centre; University of Oslo; Oslo; Norway
24 PsychGen Centre for Genetic Epidemiology and Mental Health; Norwegian Institute of Public Health; Oslo; Norway
25 Psychiatric Genetic Epidemiology Group, Research Department; Lovisenberg Diakonale Hospital; Oslo; Norway
26 MRC Integrative Epidemiology Unit, Population Health Sciences; Bristol Medical School; University of Bristol; Bristol; United Kingdom
27 Estonian Genome Centre, Institute of Genomics; University of Tartu; Tartu; Estonia
28 Vanderbilt Genetics Institute; Vanderbilt University; Nashville; Tennessee; United States
29 Department of Genetic Medicine; Vanderbilt University Medical Center; Nashville; Tennessee; United States
30 Institute for Molecular Medicine Finland FIMM, Helsinki Institute of Life Science HiLIFE; University of Helsinki; Helsinki; Finland
31 Analytic and Translational Genetics Unit; Broad Institute of the Massachusetts Institute of Technology and Harvard University; Massachusetts General Hospital; Boston; Massachusetts; United States
32 Stanley Center for Psychiatric Research; Broad Institute of the Massachusetts Institute of Technology and Harvard University; Cambridge; Massachusetts; United States
33 Independent Researcher; Mebane; North Carolina; United States
34 Oregon Research Institute; Springfield; Oregon; United States
35 Department of Medicine; New Jersey Medical School, Rutgers University; Newark; New Jersey; United States
36 Department for Medical Biology; University of Split School of Medicine; Split; Croatia
37 Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy; Ludwig-Maximilians-Universität München; Munich; Germany
38 Department of Child and Adolescent Psychiatry; Oberberg Fachklinik Fasanenkiez Berlin; Berlin; Germany
39 Department of Women's and Children's Health; University of Padova; Padova; Italy
40 Department of Health Sciences; University of Florence; Florence; Italy
41 Department of Neuroscience; University of Padova; Padova; Italy
42 Department of Clinical Neuroscience; Karolinska Institutet; Stockholm; Sweden
43 Altrecht Eating Disorders Rintveld; Altrecht Mental Health Institute; Zeist; Utrecht; The Netherlands
44 Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy; LVR University Clinic Essen, University of Duisburg-Essen; Essen; Northrhine-Westfalia; Germany

45 College of Nursing; Seattle University; Seattle; Washington; United States
46 Nantes Université, CNRS, INSERM, l'institut du thorax, F-44000 Nantes, France
47 Department of Psychiatric Genetics, Medical Biology Center; Poznan University of Medical Sciences; Poznan; Poland
48 Department of Psychiatry and Behavioral Sciences; Stanford University; Stanford; California; United States
49 Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, Center of Mental Health; University Hospital Wuerzburg; Würzburg; Bavaria; Germany
50 Department of Psychiatry; Reinier van Arkel; s-Hertogenbosch; Northern Brabant; The Netherlands
51 Department of Cancer, Epidemiology and Genetics; Masaryk Memorial Cancer Institute; Brno; Czech Republic
52 Department of Psychiatry and Psychotherapy, Comprehensive Center for Clinical Neurosciences and Mental Health (C3NMH); Medical University of Vienna; Vienna; Austria
53 First Department of Psychiatry; National and Kappodistrian University of Athens (NKUA); Athens; Greece
54 Department of Genetics; Queensland Institute of Medical Research QIMR Berghofer Medical Research Institute; Brisbane; Queensland; Australia
55 The Lundbeck Foundation Initiative for Integrative Psychiatric Research (iPSYCH); Aarhus University; Aarhus; Denmark
56 Center for Genomics and Personalized Medicine; Aarhus University; Aarhus; Denmark
57 Bioinformatics Research Centre; Aarhus University; Aarhus; Denmark
58 Department of Emergency and Post-Emergency Psychiatry; CHU, University of Montpellier; Montpellier; France
59 Helmholtz Zentrum München - German Research Centre for Environmental Health; Institute of Translational Genomics; Neuherberg; Germany
60 Human Genomics Research Group, Department of Biomedicine; University of Basel; Basel; Basel-Stadt; Switzerland
61 Department of Genomics, Life & Brain Center; University of Bonn; Bonn; Northrhine-Westfalia; Germany
62 Institute of Human Genetics; University of Bonn, School of Medicine & University Hospital Bonn; Bonn; Northrhine-Westfalia; Germany
63 Eating Disorders Unit; Parkland-Klinik; Bad Wildungen; Germany
64 Department of Clinical Psychology; University Hospital Bellvitge; Hospitalet del Llobregat (Barcelona); Barcelona; Catalonia; Spain
65 Department of Clinical Sciences; School of Medicine and Health Sciences; University of Barcelona; Hospitalet del Llobregat (Barcelona); Barcelona; Catalonia; Spain
66 Ciber Physiopathology of Obesity and Nutrition (CIBERObn); Instituto de Salud Carlos III; Madrid; Spain
67 Psychoneurobiology of Eating and Addictive Behaviors Research Group; Bellvitge Biomedical Research Institute (IDIBELL); Hospitalet del Llobregat (Barcelona); Barcelona; Catalonia; Spain
68 Centre for Psychological Services; University of Barcelona; Barcelona; Catalonia; Spain

69 Rheumatology Research Group; Vall d'Hebron Research Institute; Barcelona; Catalonia; Spain
70 Department of Psychiatry; First Faculty of Medicine; Charles University and General University Hospital; Prague; Czech Republic
71 Institute of Public Health and Clinical Nutrition; University of Eastern Finland; Kuopio; Finland
72 Department of Psychological Medicine; University of Otago; Christchurch; New Zealand
73 Institute of Applied Health Sciences; University of Aberdeen; Aberdeen; Scotland; United Kingdom
74 Division of Psychological and Social Medicine and Developmental Neurosciences, Faculty of Medicine, Technische Universität Dresden, Dresden, Germany
75 Department of Psychological Medicine; Institute of Psychiatry, Psychology and Neuroscience; King's College London; London; United Kingdom
76 Center for Applied Genomics; Children's Hospital of Philadelphia; Philadelphia; Pennsylvania; United States
77 Division of Human Genetics; Children's Hospital of Philadelphia; Philadelphia; Pennsylvania; United States
78 Department of Pediatrics; University of Pennsylvania Perelman School of Medicine; Philadelphia; Pennsylvania; United States
79 Clinical Psychology Program; The Chicago School, Washington DC, College of Clinical Psychology; Washington DC; United States
80 Department of Psychiatry; University of Campania "Luigi Vanvitelli"; Naples; Italy
81 Department of Medical Research; Vestre Viken Hospital Trust, Bærum Hospital; Gjettum; Norway
82 Division of Mental Health and Addiction; NORMENT KG Jebsen Centre; Oslo University Hospital; Oslo; Norway
83 Padova Neuroscience Center; University of Padova; Padova; Italy
84 Department of Pathology and Biomedical Science; University of Otago; Christchurch; New Zealand
85 National Center for PTSD; VA Boston Healthcare System; Boston; Massachusetts; United States
86 Department of Psychiatry; Boston University Chobanian & Avedisian School of Medicine; Boston; Massachusetts; United States
87 Department of Mental and Physical Health and Preventive Medicine; University of Campania "Luigi Vanvitelli"; Naples; Italy
88 Department of Population Health; Queensland Institute of Medical Research QIMR Berghofer Medical Research Institute; Brisbane; Queensland; Australia
89 Department of Medicine Solna; Division of Rheumatology; Karolinska Institutet; Stockholm; Sweden
90 Department of Psychiatry; Division of Psychiatric Genomics; Icahn School of Medicine at Mount Sinai; New York; New York; United States
91 Department of Psychiatry; Helsinki University Hospital; Helsinki; Finland
92 Department of Public Health; University of Helsinki; Helsinki; Finland
93 Department of General Practice & Primary Healthcare, Faculty of Medical & Health Sciences; The University of Auckland; Auckland; New Zealand
94 Faculty of Medicine; Sigmund Freud University; Vienna; Austria

95 Institute of Biology and Medical Genetics; First Faculty of Medicine; Charles University; Prague; Czech Republic
96 Department of Child and Adolescent Psychiatry; Poznan University of Medical Sciences; Poznan; Poland
97 Department of Neuroscience, Psychology, Drug Research and Child Health (NEUROFARBA); University of Florence; Florence; Italy
98 Department of Psychiatry; University of Perugia; Perugia; Italy
99 Adolescent Health Unit, Second Department of Pediatrics, "P. & A. Kyriakou" Children's Hospital; National and Kappodistrian University of Athens (NKUA); Athens; Greece
100 Department of Clinical Psychology, Faculty for Social Sciences; Utrecht University; Utrecht; Utrecht; The Netherlands
101 Eating Disorders Unit, Department of Child and Adolescent Psychiatry; Medical University of Vienna; Vienna; Austria
102 Discipline of Psychology; Curtin University; Perth; Western Australia; Australia
103 Department of Translational Neuroscience; UMC Utrecht Brain Center; University Medical Center Utrecht, Utrecht University; Utrecht; The Netherlands
104 Department of Physiology; Institute of Neuroscience and Physiology; Sahlgrenska Academy at University of Gothenburg; Gothenburg; Sweden
105 Institute of Environmental Medicine; Karolinska Institutet; Stockholm; Sweden
106 Centre for Occupational and Environmental Medicine; Region Stockholm; Stockholm; Sweden
107 Centre for Precision Psychiatry; University of Oslo; Oslo; Norway
108 KG Jebsen Centre for Neurodevelopmental Disorders Research; University of Oslo; Oslo; Norway
109 Eating Recovery Center; Hunt Valley; Maryland; United States
110 Department of Psychiatry; ERC Pathlight; University of Maryland, St. Joseph Medical Center; Baltimore; Maryland; United States
111 Department of Psychiatry; University of Minnesota; Minneapolis; Minnesota; United States
112 The Weindrich Department of AI and Human Health; Icahn School of Medicine at Mount Sinai; New York; New York; United States
113 Department of Psychosomatic Medicine and Psychotherapy; Hannover Medical School; Hannover; Germany
114 Department of Nutrition and Dietetics; Harokopio University; Athens; Greece
115 Department of Psychiatry; Rutgers University; Piscataway; New Jersey; United States
116 Research Department; Quantitative Genomics Laboratories (qGenomics); Barcelona; Catalonia; Spain
117 Institute of Mathematics and Statistics; University of Tartu; Tartu; Estonia
118 Institute of Neuroscience and Medicine (INM-1); Research Center Juelich; Juelich; Germany
119 Centre for Human Genetics; University of Marburg; Marburg; Germany
120 Université Paris Cité, INSERM U1266 (IPNP); Institute of Psychiatry and Neuroscience of Paris; Paris; Ile de France; France
121 Sainte-Anne hospital (CMME); GHU Paris Psychiatrie et Neurosciences; Paris; Ile de France; France

122 Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy; RWTH Aachen University; Aachen; Germany
123 Biological Psychiatry Laboratory; McLean Hospital; Harvard Medical School; Belmont; Massachusetts; United States
124 Eating Recovery Center; Denver; Colorado; United States
125 Specialist Mental Health Clinical Research Unit; Health New Zealand - Canterbury; Christchurch; New Zealand
126 Department of Psychiatry; Centre for Addiction and Mental Health; University of Toronto; Toronto; Ontario; Canada
127 Department of C & A Psychiatry; Medical University of Vienna; Vienna; Austria
128 Groningen Institute for Evolutionary Life Sciences; University of Groningen; Groningen; The Netherlands
129 Department of Psychiatry; University of California San Diego; San Diego; California; United States
130 Department of Psychiatry; University of Toronto; Toronto; Ontario; Canada
131 Tanenbaum Centre; Centre for Addiction and Mental Health; Toronto; Ontario; Canada
132 Department of Psychiatry; Ilsan Paik Hospital, Inje University; Goyang; South Korea
133 Department of Psychology; Michigan State University; East Lansing; Michigan; United States
134 Department of Psychiatry and Neurochemistry; Institute of Neuroscience and Physiology; University of Gothenburg; Gothenburg; Sweden
135 Department of Clinical Science; University of Bergen; Bergen; Norway
136 Department of Neuroscience; Janssen Research & Development, LLC; Titusville; New Jersey; United States
137 Human Genetics & Genomics; CHDI Management, Inc.; Princeton, New Jersey; United States
138 Maria Sklodowska-Curie National research Institute of Oncology; Warsaw; Poland
139 Department of Psychiatry; Amsterdam University Medical Center; Amsterdam; The Netherlands
140 Department of Psychiatry and Neuropsychology; School for Mental Health and Neuroscience, Maastricht University Medical Center; Maastricht; The Netherlands
141 GGZ InGeest; Amsterdam; The Netherlands
142 InsideOut Institute; University of Sydney; Sydney; Australia
143 Department of Community Health and Epidemiology; Dalhousie University; Halifax; Nova Scotia; Canada
144 Institute of Psychiatric Phenomics and Genomics (IPPG); Ludwig-Maximilians-Universität München; Munich; Germany
145 Department of Mental Health and Neuroscience; Queensland Institute of Medical Research QIMR Berghofer Medical Research Institute; Brisbane; Queensland; Australia
146 School of Psychology; University of Queensland; Brisbane; Queensland; Australia
147 School of Psychology and Counselling; Queensland University of Technology; Brisbane; Queensland; Australia
148 University of Colorado School of Medicine; Aurora; Colorado; United States
149 Great Ormond Street Institute of Child Health; University College London; London; United Kingdom

150 Department of Psychiatry and Behavioral Science; University of North Dakota; Fargo; North Dakota; United States
151 Department of Medicine, Surgery and Dentistry "Scuola Medica Salernitana"; University of Salerno; Salerno; Italy
152 Department of Psychiatry and Biobehavioral Sciences; University of California Los Angeles; Los Angeles; California; United States
153 Université Paris Cité; Paris; Ile de France; France
154 INSERM U1266; INSERM U1266; Paris; Ile de France; France
155 Institute of Clinical Medicine; University of Oslo; Oslo; Norway
156 German Center for Mental Health (DZPG); Berlin-Potsdam; Germany
157 Department of Psychiatry and Psychotherapy; Charité - Universitätsmedizin; Berlin; Germany
158 Department of Adult Psychiatry; Poznan University of Medical Sciences; Poznan; Poland
159 The Centre for Applied Genomics, Program in Genetics and Genomic Biology; The Hospital for Sick Children; Toronto; Ontario; Canada
160 McLaughlin Centre and Department of Molecular Genetics; University of Toronto; Toronto; Ontario; Canada
161 GGZ Rivierduinen Eating Disorders Ursula; Leiden; The Netherlands
162 Department of Psychiatry; Leiden University Medical Centre; Leiden; The Netherlands
163 Psychiatry Department; McGill University; Montreal; Quebec; Canada
164 Eating Disorders Continuum; Douglas Mental Health University Institute; Montreal; Quebec; Canada
165 Department of Environmental Epidemiology; Nofer Institute of Occupational Medicine; Lodz; Poland
166 Discipline of Psychology; Flinders Institute for Mental Health and Wellbeing; Adelaide; South Australia; Australia
167 Department of Clinical Medicine; University of Copenhagen; Copenhagen; Denmark
168 School of Nursing; University of North Carolina at Chapel Hill; Chapel Hill; North Carolina; United States
169 Department of Psychosomatic Medicine and Psychotherapy; University Medical Hospital Tuebingen; Tuebingen; Germany
170 German Centre for Mental Health, Tuebingen; University Tuebingen; Tuebingen; Germany
171 Department of Nutrition; University of North Carolina at Chapel Hill; Chapel Hill; North Carolina; United States

* Joint first authors
^ Joint senior authors

Correspondence address:
Dr Jonathan Coleman,
Social, Genetic and Developmental Psychiatry Centre,
Institute of Psychiatry, Psychology and Neuroscience – PO80,
King's College London,
London

SE5 8AF,
United Kingdom
jonathan.coleman@kcl.ac.uk

ABSTRACT

Eating disorders —including anorexia nervosa (AN), bulimia nervosa, and binge eating disorder—are clinically distinct but exhibit symptom overlap and diagnostic crossover. Genomic analyses have mostly examined AN. We conducted the first genomic meta-analysis of binge-eating behaviour (BE; 39,279 cases, 1,227,436 controls), alongside new analyses of AN (24,223 cases, 1,243,971 controls) and its subtypes (all European ancestries). We identified six loci associated with BE, including loci associated with higher body mass index (BMI) and impulse-control behaviours. AN GWAS yielded eight loci, validating six loci. Subsequent polygenic risk score analysis demonstrated an association with AN in two East Asian ancestry cohorts. BE and AN exhibited similar positive genetic correlations with psychiatric disorders, but opposing genetic correlations with anthropometric traits. Most of the genetic signal in BE and AN was not shared with BMI. We have extended eating disorder genomics beyond AN; future work will incorporate multiple diagnoses and global ancestries.

Eating disorders include anorexia nervosa (AN), bulimia nervosa, and binge eating disorder, among others. They are diagnostically distinct, yet show considerable overlap in symptoms, as reflected in diagnostic migration over time [1–3]. AN is characterised by low weight, fear of weight gain, and an inability to recognise the seriousness of the low weight. It has two subtypes, which both exhibit low weight achieved by caloric restriction and increased energy expenditure. In the binge eating/purging subtype (AN-BP), this is coupled with binge eating (BE) and/or purging behaviours, whereas the restricting subtype (AN-R) lacks these features. Bulimia nervosa occurs in individuals at normal or high weight and is characterised by the combination of BE and compensatory behaviours (e.g., fasting, self-induced vomiting, laxative use, diuretic use). Binge eating disorder shares the BE component of bulimia nervosa and also occurs at both normal and high weights, but lacks recurrent compensatory behaviours [1].

Genome-wide association studies (GWASs) of eating disorders have focused primarily on AN [4–7], in part due to its elevated mortality [8], with only a few GWASs of other eating disorder-related phenotypes to date [9,10]. The most recent AN GWAS [6] included 16,992 cases and identified eight genome-wide significant loci. Genetic correlations based on common single nucleotide polymorphisms (SNP-r_gs) showed high correlations with other psychiatric disorders, and suggested that metabolic and anthropometric factors might also underlie AN pathophysiology [6,7]. The metabolic aspect of AN is reflected by a positive SNP-r_g with high-density lipoprotein cholesterol and negative SNP-r_gs with insulin resistance, leptin, and type 2 diabetes. Importantly, these SNP-r_gs were independent of body mass index (BMI), a significant finding given that low BMI is a defining feature of AN.

To fully understand the genetic landscape of eating disorders, it is essential to advance beyond AN. A substantial genetic correlation between AN and bulimia nervosa has been shown in family and twin studies [11,12], suggesting that these phenotypes share genetic risk. This may partially reflect the presence of BE as a transdiagnostic symptom, common both to bulimia nervosa and to AN-BP. Psychiatric diagnoses defined by the DSM and ICD

have been criticised for failing to capture patient experiences or biological realities, given that disorders often exist on a continuum, overlap, co-occur, present heterogeneously, and morph over time [13]. Studying only those who meet diagnostic thresholds can bias research toward the most severe cases and excludes individuals who have significant pathology but do not meet strict diagnostic criteria. Alternative, symptom-based approaches have been proposed to complement diagnostic approaches and advance the field of eating disorders [14]. Here, we present the first GWAS of BE across multiple cohorts. We additionally conducted an AN GWAS meta-analysis with augmented sample size and increased statistical power, and the first GWASs of explicitly defined AN subtypes (AN-R and AN-BP). We identified genetic commonalities and differences between BE and AN.

RESULTS

Summary of phenotypes

We operationalised five phenotypes (**Table 1**): broadly-defined BE (BE-BROAD) and narrowly-defined BE (BE-NARROW), and AN, AN-R, and AN-BP. We primarily report results for BE-BROAD (which captures the common genetic component of BE with greater statistical power than BE-NARROW; Methods) and for AN. Analyses of BE-NARROW, AN-R, and AN-BP are reported in Supplementary Results. We contrast our BE-BROAD results with those of a previous publication that assessed a model-derived BE phenotype in the Million Veteran Program [10] (Supplementary Material).

Table 1 approximately here

GWAS meta-analyses

Our BE-BROAD GWAS included 17 European-ancestry datasets with 39,279 cases and 1,227,436 controls, assessing 6,244,919 common (minor allele frequency≥1%), high-confidence (imputation INFO score>0.6), autosomal single nucleotide polymorphisms

(SNPs; Supplementary Table 1). Conditional and joint analyses confirmed six independently associated loci (Figure 1; Table 2; Supplementary Figures 1-6; Supplementary Table 2). The liability-scale SNP-based heritability for BE-BROAD was 5% (SE=0.4%, assuming population prevalence of 4.5% [15]). The intercept was 1.03, significantly >1 and the attenuation ratio was 0.14 (SE=0.04) (Supplementary Table 3). While an intercept >1 can indicate confounding, this ratio suggests that inflation was primarily due to polygenicity [16].

For AN, we identified eight independently associated loci in 26 European-ancestry datasets with 24,223 cases and 1,243,971 controls, across 6,926,820 common, high-confidence, autosomal SNPs (Figure 1; Table 2; Supplementary Figures 7-14; Supplementary Tables 1 and 2). Six of these loci were detected in a previous AN GWAS [6] and two were newly identified. Three previously significant loci (on chromosome 2 and 3 from Watson et al [6], and on chromosome 12 from Duncan et al [7]) did not reach genome-wide significance (P=2x10^{-7}–6x10^{-7}). The liability-scale SNP-based heritability was 13% (SE=0.7%, assuming population prevalence of 1.5% [17]), with an intercept of 1.02, significantly >1 and an attenuation ratio of 0.07 (SE=0.03), suggesting inflation was largely due to polygenicity (Supplementary Table 3).

We also conducted analyses on chromosome X for a subset of studies with available data (Supplementary Table 4). No genome-wide significant loci were identified for BE-BROAD nor AN, although one genome-wide significant locus was identified for BE-NARROW (Supplementary Results, Supplementary Figure 15, Supplementary Table 5).

Given known sex differences in eating disorders, and mostly female cases in our data (96% in BE-BROAD, 94% in AN), we carried out female-only GWASs as sensitivity analyses. Results were similar to the main analyses, with differences attributable to the reduction in sample size (Supplementary Results, Supplementary Table 6).

Figure 1 approximately here

Table 2 approximately here

Genetic relationship between eating phenotypes and other traits

We assessed the genetic similarity of BE-BROAD and AN through examining their SNP-r_g with each other and with other traits, via univariate and bivariate causal mixture modelling ("MiXeR analysis") and via case-case GWAS [18,19]. The SNP-r_g between BE-BROAD and AN was 0.46 (SE=0.04, P=3.44x10^{-30}), indicating moderate genetic overlap (Supplementary Table 7). MiXeR analysis resulted in a similar estimate of SNP-r_g (0.49, SE=0.01). Bivariate MiXeR indicated an overlap of 2654 causal variants between BE-BROAD and BE-AN (83% of all BE-BROAD causal variants) with highly concordant effects (95%; Figure 2a). However, the bivariate MiXeR model was not distinguishable from the minimal possible overlap model, which requires an overlap >2450 variants given the genetic correlation observed (Supplementary Results, Supplementary Table 8). Case-case GWAS leverages a genetic distance measure representing the average squared difference in allele frequency at causal SNPs [18]. The genetic distance was highest between individuals with AN and controls (0.46), similar to that between case groups (0.40), whereas the genetic distance between individuals with BE-BROAD and controls was smaller (0.25, Figure 2b). We identified case-divergent loci on chromosomes 1, 3, and 5, overlapping with independent loci identified in the AN GWAS (Figure 1), suggesting that some loci differentiating AN from controls also differentiate AN from BE-BROAD.

Figure 2 approximately here

We used LDSC to calculate pairwise SNP-r_g for both BE-BROAD and AN with 225 traits including psychiatric, personality, metabolic, and anthropometric traits (Table 3, Supplementary Table 9). We generally observed positive SNP-r_g of BE-BROAD with psychiatric traits and anthropometric phenotypes, except for persistent thinness and pubertal growth which displayed negative SNP-r_g. In contrast, significant SNP-r_g with metabolic traits were absent for BE-BROAD except for BMI-adjusted fasting insulin. We validated and

extended previously observed SNP-r_g patterns with AN [6] (Table 3, Supplementary Table 9). We found positive SNP-r_g across psychiatric disorders, as well as with neuroticism, educational attainment, and physical activity. We observed negative SNP-r_g across anthropometric traits, and notably, a non-significant SNP-r_g with persistent thinness. Metabolic SNP-r_g mirrored Watson et al., with predominantly negative SNP-r_g, except for total cholesterol in high-density lipoprotein [6].

Next, we tested for significant differences between the SNP-r_g of BE-BROAD and SNP-r_g of AN with other traits (Figure 3, Supplementary Figure 16, Supplementary Table 10). Most psychiatric and behavioural traits and disorders showed similar SNP-r_g with BE-BROAD and AN. However, attention deficit/hyperactivity disorder (ADHD) showed greater SNP-r_g (difference P=1.06x10^{-7}) with BE-BROAD than with AN, and obsessive-compulsive disorder showed greater SNP-r_g (P=4.79x10^{-5}) with AN than with BE-BROAD. BE-BROAD was positively correlated with Alcohol Use Disorder Identification Test problem items, four smoking-related phenotypes, and general risk tolerance, while AN was not (P≤4.83x10^{-7}). AN was negatively correlated with automobile speeding propensity, while BE-BROAD was not (P=1.21x10^{-4}).

BE-BROAD and AN also diverged in their associations with anthropometric and metabolic traits. For example, BE-BROAD was positively correlated with waist-to-hip ratio while AN was negatively correlated (P=2.05x10^{-31}). BE-BROAD showed a stronger pattern of SNP-r_g with certain socio-demographic traits than AN (P<2x10^{-4}), displaying negative SNP-r_gs with age at menarche and age at first birth in females, and positive SNP-r_gs with social deprivation and loneliness. In contrast, AN showed no significant SNP-r_g with these traits but was more strongly positively associated with educational traits such as college/university completion than was BE-BROAD.

Figure 3 approximately here

Table 3 approximately here

The genetic signal in our BE-BROAD GWAS may partly be influenced by AN, given that 18% of our BE-BROAD cases had (known) AN (Supplementary Table 11). As a sensitivity analysis, we conducted an additional BE-BROAD GWAS, excluding cohorts that specifically focused on AN recruitment (Supplementary Results). The SNP-r_g between BE-BROAD and the reduced GWAS did not differ from unity (0.96, SE=0.07), but SNP-r_gs with anthropometric traits were stronger in the reduced GWAS, suggesting that AN cases with BE may mask BE-anthropometric genetic associations (Supplementary Figure 17, Supplementary Table 12).

Role of BMI genetics

The role of BMI in eating disorders is complex, with low BMI being pathognomonic of AN and individuals with binge eating disorder often being overweight [1]. If the genetics of BMI were the only genetic cause of BE and AN (in opposite directions), then our GWASs would just be proxy GWASs for BMI. To assess this, we applied GWAS-by-subtraction to remove the genetic variance of BMI from BE-BROAD and from AN separately [20]. We modelled a factor shared between each eating phenotype and BMI, and a *non-BMI* factor only loaded on by the eating phenotype. The shared factor explained 12% (SE=2.7%) of genetic variance in BE-BROAD, leaving 88% (SE=7.9%) accounted for by the *non-BMI* factor. In AN, the shared factor accounted for 10% (SE=1.4%) of genetic variance, leaving 90% (SE=5.5%) accounted for by the *non-BMI* factor. GWASs of the *non-BMI* factor for BE-BROAD and for AN generally resulted in larger p-values for lead SNPs, but larger effect sizes in the same direction as the original GWAS (Figure 1, Supplementary Table 13). In pairwise SNP-r_g analyses of the *non-BMI* component of BE-BROAD and AN, SNP-r_g with psychiatric disorders typically remained stable or slightly increased relative to the respective full GWAS, whereas SNP-r_g with anthropometric and metabolic traits were typically attenuated (Supplementary Figure 18, Supplementary Table 14).

We also conducted exploratory two-sample Mendelian randomisation (MR) analyses of each eating phenotype with BMI, testing causal effects in both directions. We used SNPs in linkage equilibrium as genetic instruments, with $P<5x10^{-6}$ for BE-BROAD and AN, and $P<5x10^{-9}$ for BMI [21]. For each analysis, we used the full BE-BROAD or AN GWAS, and the GWAS of the respective *non-BMI* component (Supplementary Results, Supplementary Table 15). Significant results were found in both directions between increased BE-BROAD risk and increased risk for higher BMI. BE-BROAD was still associated with increased risk for higher BMI when using just the *non-BMI* component. When examining the effect of BMI on the *non-BMI* component of BE-BROAD, odds ratios (ORs) were consistently >1, but the different methods were not consistently significant. Significant results were found in both directions between increased risk of AN and decreased BMI. However, the *non-BMI* component of AN risk was not associated with BMI, and results were inconsistent across MR methods when examining the effect of BMI on the *non-BMI* component of AN.

Genetically-regulated gene expression

We used S-PrediXcan [22] to identify predicted genetically-regulated gene expression associated with our phenotypes (Supplementary Figure 19; Supplementary Table 16). For BE-BROAD, two gene-tissue associations were significant at the experiment-wide threshold (*PRKAR2A*-Sigmoid colon and *KLHDC8B*-Heart, atrial appendage; $P<8.32x10^{-8}$). In AN, we observed 300 experiment-wide significant gene-tissue associations ($P<8.32x10^{-8}$) with 29 unique genes, predominantly from the gene-dense locus on chromosome 3:47-52Mb. Among the experiment-wide significant associations, 94 were in central nervous system tissues and 41 in gastrointestinal tissues. Within-tissue significant results for both phenotypes are described in the Supplementary Results.

We calculated cross-tissue predicted genetically-regulated gene expression using S-MultiXcan [23] to identify apparent tissue-specific associations that are better interpreted as cross-tissue (Supplementary Results, Supplementary Figure 20, Supplementary Table 17).

Ten genes had significant ($P<2.25x10^{-6}$) cross-tissue expression in BE-BROAD, four of which were identified as tissue-level associations (including *KLHDC8B* but not *PRKAR2A*). Similarly, 43 genes showed significant cross-tissue expression in AN, of which 23 were identified as tissue-level associations.

Gene-level associations

We used MAGMA v1.10 [24] to conduct gene-wise analyses of the aggregate effect of SNPs mapped to protein-coding genes (Supplementary Table 18); gene-set analyses of groups of genes with shared functional, biological, or other characteristics (Supplementary Table 19); and gene-set analyses restricted to genes targeted by medications (Supplementary Table 20; Supplementary Results). We also examined enrichment of signal within drug-sets belonging to the same class of drugs (Supplementary Table 21; Supplementary Results). For BE-BROAD, MAGMA identified 22 genes significant after Bonferroni correction ($P<2.59x10^{-6}$). *FTO* had the strongest association ($P=2.8x10^{-22}$), while 9/22 (43%) of the Bonferroni-significant genes were linked to the gene-dense locus on chromosome 3:47-52Mb. The MAGMA gene set, drug set, and drug-class analysis yielded no significant results for BE-BROAD. For AN, MAGMA identified 76 significant genes ($P<2.58x10^{-6}$). Most (54/76, 71%) were again mapped to chromosome 3:47-52Mb. Gene-set analysis identified enrichment in three biological pathways, related to the binding targets of RBFOX1-3 (RNA binding proteins that regulate neuronal alternative splicing [25]), and to mutation-constrained genes with pLI>0.9. Drug-set analysis revealed no significant drug sets, but antimigraine preparations as a class were significantly associated with AN, consistent with previous associations between AN and migraine polygenic risk scores (PRS) [26].

Genes implicated both through proximity (MAGMA) and through effects on gene expression (S-PrediXcan) are more likely to be functionally relevant than those implicated through proximity alone [27]. Further restricting our MAGMA gene-wise results to genes that

were at least tissue-level significant in S-PrediXcan resulted in seven prioritised genes across four loci in BE-BROAD, and 38 prioritised genes across eight loci in AN (Supplementary Table 22; Supplementary Results).

Tissue and cell-type analyses

We used stratified LDSC [28] to estimate the enrichment of SNP-based heritability for BE-BROAD and AN among genes specifically expressed in GTEx human tissues (Supplementary Figure 21, Supplementary Table 23; Supplementary Results) and in cell types from the Human Brain Atlas (Supplementary Figure 22, Supplementary Table 24) [29,30]. After accounting for multiple testing, no associations were significant.

Polygenic prediction

We tested whether higher BE-BROAD and AN PRS were associated with a higher risk of BE-BROAD and AN, using a leave-one-cohort-out design in well-powered cohorts representative of the ascertainment methods employed in the study (Methods; Supplementary Table 25). Individuals with 1 SD higher BE-BROAD PRS had an average OR of 1.11 for BE-BROAD (average 95% confidence interval [CI] across cohorts: 1.08–1.14; *P* range: 2.38×10^{-15}–0.022; Supplementary Figure 23). The average liability-scale variance explained was 0.32%. Individuals with 1 SD higher AN PRS had an average OR of 1.50 for AN (average 95% CI 1.42–1.59, all $P < 2 \times 10^{-16}$; Supplementary Figure 23) and AN PRS explained 2.32% liability-scale variance on average.

We additionally tested the cross-ancestry prediction of AN in two East Asian ancestry cohorts from Korea and Japan. The European AN PRS was positively associated with AN in the combined Korean and Japanese cohort, with an OR of 1.36 (95% CI 1.09 – 1.70, *P*=0.0066), explaining 1.3% of the variance (assuming population prevalence of AN at 1.5%).

Next, we tested whether genetic risk of BE-BROAD and AN were shared across males and females, using a similar leave-one-cohort-out design. BE-BROAD PRS based on female-only GWAS were positively associated with BE-BROAD risk in males (OR 1.06–1.20), but not all results were significant, possibly due to low case numbers in some cohorts. Similarly, AN PRS calculated based on female-only GWAS were positively associated with AN risk in males (OR 1.07–1.41), but the results were not consistently significant (Supplementary Figure 24, Supplementary Table 26; Supplementary Results).

We further assessed if BE-BROAD and AN PRS differed across BE and AN subgroups, comparing control individuals to (a) those with BE-BROAD only; (b) with both BE-BROAD and AN (BE+AN); and (c) with AN only. Overall, we found both BE-BROAD and AN PRS to be elevated in all subgroups compared to controls ($P \leq 0.019$). Among the subgroups, the BE+AN and BE-BROAD-only groups did not significantly differ on BE-BROAD PRS, and typically had a significantly higher BE-BROAD PRS than the AN-only group (Supplementary Figure 25). The BE+AN and AN-only groups did not significantly differ from each other on AN PRS, while the BE-BROAD-only group had a significantly lower AN PRS than both other groups in one cohort, but not in another (Supplementary Figure 25, Supplementary Table 27; Supplementary Results).

DISCUSSION

In this first genome-wide analysis of BE, we implicate six genomic loci. These have been previously associated with smoking [31], risk-taking behaviour [32], and age at menarche [33]. Overlap between BE and impulse-control behaviours was further observed in positive SNP-r_g with smoking, general risk tolerance, and problematic alcohol use. Loss of control is a key component of BE, and impulse-control behaviours have been associated with binge-type eating disorders clinically [34–36]. Alongside significant polygenic overlap with a range of psychiatric disorders, our findings imply that BE shares genetic underpinnings with psychiatric disorders and impulse-control behaviours.

Loci associated with BE-BROAD have also been implicated in anthropometric traits, including a BMI-related signal near *FTO* [21,37] that was not associated with AN or its subtypes. The *FTO* locus has been studied extensively (summarised in Loos and Yeo [38]), but it has been challenging to determine the causal mechanism that contributes to a high BMI [38]. One study found that *FTO* was related to BE independent of BMI [39], and suggested that BE could mediate the pathway between *FTO* and a high BMI. Together with our results, this implies that the relationship between *FTO* and high BMI could result partly from binge-eating behaviours. Further research should examine whether broader disordered eating behaviours mediate the relationship between *FTO* and high BMI. Genetic correlations of BE with anthropometric traits and impulse-control behaviours mirror clinical observations in individuals with BE, and suggest the relationship of BE to these traits and behaviours is partly due to pleiotropy rather than being purely environmental or a consequence of BE itself.

For AN, we validated six previously identified loci, identified two new loci, and identified one locus for AN-R. The four single-gene loci identified in Watson et al. [6] remained genome-wide significant, suggesting that genes located in these regions—*CADM1, MGMT, FOXP1, PTBP2*—may warrant further investigation in the aetiology of AN [40]. The AN-R-identified locus narrowly missed genome-wide significance in AN (P=5.89x10^{-8}) and has previously been implicated in schizophrenia [41]. The locus contains several genes, but only *DLX1* was indicated by both proximity-based and expression-based gene mapping. *DLX1* is differentially expressed in the brain and may be involved in several processes of neural development [42]. Further studies are needed to confirm that *DLX1* is implicated in AN-R aetiology, given the multigenic nature of the locus. Our gene-level results should generally be viewed cautiously, as greater power is needed to effectively fine-map associated loci and link causal variants to genes.

Despite increasing our effective sample size for AN by 64% since our previous freeze [6], we identified only two new loci, and two previously implicated loci were no longer genome-wide significant. We speculate that this is because our new cohorts were primarily population-based and used more lenient case criteria compared to the clinical diagnoses and targeted AN-specific recruitment previously used [43]. Consistent with this, AN PRS typically captured less variance in AN in new cohorts (Supplementary Results, Supplementary Table 25). Genetic signal also tends to become more heterogeneous as GWAS sample sizes increase [44,45]. This is supported by our data. Although none of the lead SNPs associated with AN showed a heterogeneous effect (Supplementary Figures 7-14), variant heterogeneity statistics (I^2) were higher in our AN GWAS than in the AN GWAS published by Watson et al (Supplementary Results, Supplementary Table 28) [6]. Nonetheless, our AN GWAS measured variants with greater precision than that of Watson et al (average variant SE=0.0199 vs SE=0.0255 in Watson et al), indicating that power increased despite the increase in heterogeneity (Supplementary Results).

We have previously hypothesised that AN is a metabo-psychiatric disorder [6]—this study yields for the first time the ability to investigate shared and distinct metabolic and psychiatric components across multiple eating phenotypes. BE-BROAD has typical genetic features of a psychiatric disorder, including significant SNP-r_g with psychiatric traits akin to previous findings [6,7,26,46]. The SNP-r_g pattern for BE-BROAD is similar to that of AN, with notable exceptions. AN had a positive SNP-r_g with obsessive-compulsive disorder, whereas BE-BROAD showed no significant association. Conversely, BE-BROAD was positively genetically correlated with ADHD, whilst AN showed no significant association. However, there was a significant positive SNP-r_g between ADHD and *non-BMI* AN, suggesting that opposing SNP-r_gs of AN and ADHD with BMI were previously masking this association.

We observed some key differences for SNP-r_gs with non-psychiatric traits. BE-BROAD displayed a significantly stronger, negative SNP-r_g with age at menarche whilst AN

showed no genetic overlap, consistent with previous research that showed early-onset AN was negatively associated with age at menarche, but typical-onset AN was not [47]. Earlier age at menarche has been associated with increased impulse-associated traits like substance use and risky behaviour [48], and with BMI [49]. Both impulsivity and BMI are often higher in those who binge eat [1,34,35]. Observational studies find that later age of menarche is associated with AN [48,50]; however, disentangling this from the effects of starvation and low body weight is difficult. Significant SNP-r_gs between BE-BROAD and anthropometric traits were positive compared with the negative SNP-r_gs observed in AN, consistent with previous research [26,46]. We also validated our previous finding that significant SNP-r_gs with AN are concentrated in metabolic-related traits [6], in contrast to BE-BROAD.

Eating disorders and their component features share genetic factors with anthropometric traits, but these effects act in opposite directions depending on the presentation. To investigate this further, we assessed BE-BROAD and AN after subtracting the genetic component each shares with BMI. The BMI component accounted for 12% of the genetic variance of BE-BROAD and 10% of AN, despite low BMI being a diagnostic requirement for AN. The low variance explained by the BMI component argues that neither our AN nor BE-BROAD GWAS are BMI GWAS by proxy. This is further supported by the *FTO* locus association with BE-BROAD, which is observed in AN-ascertained cohorts where affected individuals are likely to have lower BMI than unaffected individuals. Consistent with previous literature [51], we also found no evidence for a genetic overlap between persistent thinness and AN, indicating that the cognitive-behavioural component of AN distinguishes these low-BMI phenotypes on a genomic level.

However, BMI is a blunt measure of body composition [52], and its relationship with eating disorders is complicated, with evidence that the negative SNP-r_g between AN and BMI is driven by genetic enrichment for AN risk in females with lower BMI, rather than a uniform linear relationship across BMI [53]. Both BE and AN are heterogeneous conditions,

and subtypes may have differing relationships with BMI. Subtracting the BMI component from our AN and BE-BROAD GWASs affords tentative insights into the physiological aspects of these illnesses, but BMI itself has a partly behavioural aetiology [28]. More sophisticated analyses with a wider range of body composition measurements and eating disorder presentations (including atypical AN in the normal or high BMI range) will provide deeper understanding.

We present the first GWAS meta-analysis of BE, accompanied by the largest investigation of AN and AN subtypes to date, with chromosome X analysed for all phenotypes. We have extended eating disorder genetic research beyond AN alone and demonstrated that BE is a psychiatric phenotype with distinctive genetic relationships with external traits. Nonetheless, readers should consider the following limitations. Our analyses only considered individuals of European genetically inferred ancestry, limiting their generalisability. We analysed PRSs in two small East Asian cohorts, but used European prevalence estimates to convert risk to the liability scale, potentially introducing bias. Differences in population structure, genetic architecture, and environmental factors, such as lower average BMI in Korean and Japanese populations [54], could influence the cross-ancestry prediction of AN. More GWAS and PRS studies in East Asian populations are required to improve the accuracy of genetic risk predictions and our understanding of AN genetics in these populations.

This limitation extends to other global ancestries and should be considered from both phenotypic and genotypic perspectives [14]. The historic focus of eating disorder studies on females of European descent has influenced disease characterization, potentially leading to certain diagnostic criteria being over- or under-represented. Prioritizing variants common in European ancestry populations may prevent the genetic architecture of eating disorders being fully characterised. To address this, ongoing research efforts aim to enhance the

generalisability and applicability of findings across a wider range of individuals through global data collection across all populations experiencing eating disorders.

Given the known diagnostic crossover between eating disorders, cohorts that only contributed cross-sectional diagnoses or symptoms are unable to account for later development of a disorder or symptom; for example, an individual with AN-R might go on to develop AN-BP [2]. It is not possible to fully mitigate this limitation. However: (1) many of our cohorts include individuals beyond the typical age of diagnosis, making new diagnoses or diagnostic crossover less likely; (2) hidden diagnostic crossover likely contributes to false negatives and under-estimation of differences between GWAS, rather than introducing false positives [55]; and (3) our previous work has shown that rates of diagnostic contamination (a similar effect to hidden diagnostic crossover) would need to occur at extremely high levels to affect locus discovery [56].

Despite our best efforts at harmonisation, heterogeneity might exist within the phenotypes due to factors such as distinct methods of ascertainment. Ideally, one approach to ascertainment would be used, such as structured diagnostic interviews within speciality clinics. However, using multiple ascertainment approaches can increase sample size and thus power despite the resulting heterogeneity. Fourth, our sample is mostly female, and results may not necessarily generalise to those who are not female. Finally, although strongly associated with BE-BROAD and AN risk, PRS remain very weak predictors of BE-BROAD and AN status. Combining PRS with other risk factors is needed to further improve prediction accuracy of BE-BROAD and AN.

Historically, binge-type eating disorders have been overshadowed by research on AN, despite their higher prevalence. This paper redresses that imbalance. We identified six genetic loci relating to broadly defined BE, validated six loci related to AN, reported two novel loci, and found one locus related to the restricting subtype of AN. We demonstrate that BE is genetically related to several other psychiatric phenotypes, with both shared and distinct patterns compared to AN, providing genetic substantiation of clinically-observed

comorbidity patterns. The number of loci we implicate in BE-BROAD and AN is typical of early-phase GWAS studies [57], motivating GWAS meta-analyses of all major eating disorders (AN, bulimia nervosa, binge eating disorder, avoidant/restrictive food intake disorder) and transdiagnostic behaviours (e.g., BE, restriction) to drive variant discovery and further refine our understanding of shared and unique genetic features that distinguish presentations and inform a genetically guided nosology of eating disorders [13].

Data availability

Individual level genotype data (except in countries where sharing of individual level data is prohibited by national law) and summary statistics used in this study are available to bona fide researchers working in collaboration with a member of the Eating Disorders Working Group of the PGC via secondary analysis proposals (https://pgc.unc.edu/for-researchers/data-access-committee/data-access-information/).

Summary statistics from this work will be made available via Figshare and the PGC website on publication (https://pgc.unc.edu/for-researchers/download-results/).

Code availability

Code underlying this work will be made available on GitHub (https://github.com/psychiatric-genomics-consortium/PGC3_EAD) on publication.

Author contributions

Specific author contributions are listed in Supplementary Table 29. All authors made substantial contributions to the conception or design of the work, or to the acquisition, analysis, or interpretation of data for the work, and critically reviewed the work for important intellectual content, gave final approval of the version to be published, and agreed to be

accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Conflicts of Interest

Susana Jiménez-Murcia and Fernando Fernández-Aranda have received consultancy and speaker honoraria from Novo Nordisk. Ole Andreassen is a consultant to Presicion Health and Cortechs.ai, and has received speaker's honoraria from Lundbeck, Janssen, Lilly, and Otsuka. James Kennedy is a member of the Scientific Advisory Board for Myriad Neuroscience. Mikael Landén has received lecture honoraria from Lundbeck pharmaceuticals. Qingqin Li was an employee of Janssen Research & Development, LLC when the work was completed and holds company equity. Nadia Micali receives an honorarium as associate editor on the European Eating Disorders Review. Dan Rujescu served as consultant for Janssen, received honoraria from Boehringer-Ingelheim, Gerot Lannacher, Janssen and Pharmagenetix, received travel support from Angelini, Janssen and Schwabe, and served on advisory boards of AC Immune, Boehringer-Ingelheim, Roche and Rovi. Patrick Sullivan is a shareholder in Neumora Therapeutics and serves on the advisory board. Cynthia Bulik receives royalties from Pearson Education, Inc. and has served as a consultant for Orbimed. No other authors report conflicts of interest.

Funding and ethics statements

The Psychiatric Genomics Consortium is supported by NIH grant R01 MH124851. Author and cohort-specific funding and ethics statements are provided in the Supplementary Materials.

REFERENCES

1. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders (DSM-5®)*. (American Psychiatric Pub, 2013).

2. Abdulkadir, M. *et al.* Descriptives and genetic correlates of eating disorder diagnostic transitions and presumed remission in the danish registry. *Biol. Psychiatry* (2025) doi:10.1016/j.biopsych.2025.01.008.

3. Solmi, M. *et al.* Outcomes in people with eating disorders: a transdiagnostic and disorder-specific systematic review, meta-analysis and multivariable meta-regression analysis. *World Psychiatry* **23**, 124–138 (2024).

4. Wang, K. *et al.* A genome-wide association study on common SNPs and rare CNVs in anorexia nervosa. *Mol. Psychiatry* **16**, 949–959 (2011).

5. Boraska, V. *et al.* A genome-wide association study of anorexia nervosa. *Mol. Psychiatry* **19**, 1085–1094 (2014).

6. Watson, H. J. *et al.* Genome-wide association study identifies eight risk loci and implicates metabo-psychiatric origins for anorexia nervosa. *Nat. Genet.* **51**, 1207–1214 (2019).

7. Duncan, L. *et al.* Significant Locus and Metabolic Genetic Correlations Revealed in Genome-Wide Association Study of Anorexia Nervosa. *Am. J. Psychiatry* **174**, 850–858 (2017).

8. 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).

9. Wade, T. D. *et al.* Genetic variants associated with disordered eating. *Int. J. Eat. Disord.* **46**, 594–608 (2013).

10. Burstein, D. *et al.* Genome-wide analysis of a model-derived binge eating disorder

phenotype identifies risk loci and implicates iron metabolism. *Nat. Genet.* **55**, 1462–1470 (2023).

11. Bulik, C. M. *et al.* Understanding the relation between anorexia nervosa and bulimia nervosa in a Swedish national twin sample. *Biol. Psychiatry* **67**, 71–77 (2010).

12. Yao, S. *et al.* Genetic and environmental contributions to diagnostic fluctuation in anorexia nervosa and bulimia nervosa. *Psychol. Med.* **51**, 62–69 (2021).

13. Tiego, J. *et al.* Precision behavioral phenotyping as a strategy for uncovering the biological correlates of psychopathology. *Nat. Ment. Health* **1**, 304–315 (2023).

14. Huckins, L. M. *et al.* What next for eating disorder genetics? Replacing myths with facts to sharpen our understanding. *Mol. Psychiatry* **27**, 3929–3938 (2022).

15. 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).

16. Loh, P.-R., Kichaev, G., Gazal, S., Schoech, A. P. & Price, A. L. Mixed-model association for biobank-scale datasets. *Nat. Genet.* **50**, 906–908 (2018).

17. Galmiche, M., Déchelotte, P., Lambert, G. & Tavolacci, M. P. Prevalence of eating disorders over the 2000-2018 period: a systematic literature review. *Am. J. Clin. Nutr.* **109**, 1402–1413 (2019).

18. Peyrot, W. J. & Price, A. L. Identifying loci with different allele frequencies among cases of eight psychiatric disorders using CC-GWAS. *Nat. Genet.* **53**, 445–454 (2021).

19. Frei, O. *et al.* Bivariate causal mixture model quantifies polygenic overlap between complex traits beyond genetic correlation. *Nat. Commun.* **10**, 2417 (2019).

20. Demange, P. *et al.* Investigating the genetic architecture of noncognitive skills using GWAS-by-subtraction. *Nat. Genet.* **53**, 35–44 (2021).

21. Pulit, S. L. *et al.* Meta-analysis of genome-wide association studies for body fat distribution in 694 649 individuals of European ancestry. *Hum. Mol. Genet.* **28**, 166–174 (2019).

22. Barbeira, A. N. *et al.* Exploring the phenotypic consequences of tissue specific gene expression variation inferred from GWAS summary statistics. *Nat. Commun.* **9**, 1825 (2018).

23. Barbeira, A. N. *et al.* Integrating predicted transcriptome from multiple tissues improves association detection. *PLoS Genet.* **15**, e1007889 (2019).

24. 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).

25. Fisher, E. & Feng, J. RNA splicing regulators play critical roles in neurogenesis. *Wiley Interdiscip. Rev. RNA* **13**, e1728 (2022).

26. Hübel, C. *et al.* One size does not fit all. Genomics differentiates among anorexia nervosa, bulimia nervosa, and binge-eating disorder. *Int. J. Eat. Disord.* **54**, 785–793 (2021).

27. Forgetta, V. *et al.* An effector index to predict target genes at GWAS loci. *Hum. Genet.* **141**, 1431–1447 (2022).

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

29. Siletti, K. *et al.* Transcriptomic diversity of cell types across the adult human brain. *Science* **382**, eadd7046 (2023).

30. Yao, S. *et al.* Connecting genomic results for psychiatric disorders to human brain cell types and regions reveals convergence with functional connectivity. *Nat. Commun.* **16**, 395 (2025).

31. Pasman, J. A. *et al.* Genetic risk for smoking: disentangling interplay between genes

and socioeconomic status. *Behav. Genet.* **52**, 92–107 (2022).

32. Baselmans, B. *et al.* The Genetic and Neural Substrates of Externalizing Behavior. *Biol Psychiatry Glob Open Sci* **2**, 389–399 (2022).

33. Pickrell, J. K. *et al.* Detection and interpretation of shared genetic influences on 42 human traits. *Nat. Genet.* **48**, 709–717 (2016).

34. Farstad, S. M. *et al.* The influence of impulsiveness on binge eating and problem gambling: A prospective study of gender differences in Canadian adults. *Psychol. Addict. Behav.* **29**, 805–812 (2015).

35. Meule, A. & Platte, P. Facets of impulsivity interactively predict body fat and binge eating in young women. *Appetite* **87**, 352–357 (2015).

36. Fernández-Aranda, F. *et al.* Impulse control disorders in women with eating disorders. *Psychiatry Res.* **157**, 147–157 (2008).

37. Bradfield, J. P. *et al.* A trans-ancestral meta-analysis of genome-wide association studies reveals loci associated with childhood obesity. *Hum. Mol. Genet.* **28**, 3327–3338 (2019).

38. Loos, R. J. F. & Yeo, G. S. H. The genetics of obesity: from discovery to biology. *Nat. Rev. Genet.* **23**, 120–133 (2022).

39. Micali, N., Field, A. E., Treasure, J. L. & Evans, D. M. Are obesity risk genes associated with binge eating in adolescence? *Obesity (Silver Spring)* **23**, 1729–1736 (2015).

40. Zheng, Y. *et al.* PTBP2 - a gene with relevance for both Anorexia nervosa and body weight regulation. *Transl. Psychiatry* **12**, 241 (2022).

41. Trubetskoy, V. *et al.* Mapping genomic loci implicates genes and synaptic biology in schizophrenia. *Nature* **604**, 502–508 (2022).

42. Letinic, K., Zoncu, R. & Rakic, P. Origin of GABAergic neurons in the human neocortex. *Nature* **417**, 645–649 (2002).

43. Thornton, L. M. *et al.* The Anorexia Nervosa Genetics Initiative (ANGI): Overview and methods. *Contemp. Clin. Trials* **74**, 61–69 (2018).

44. de Vlaming, R. *et al.* Meta-GWAS Accuracy and Power (MetaGAP) Calculator Shows that Hiding Heritability Is Partially Due to Imperfect Genetic Correlations across Studies. *PLoS Genet.* **13**, e1006495 (2017).

45. Wang, X. *et al.* Polygenic risk prediction: why and when out-of-sample prediction R2 can exceed SNP-based heritability. *Am. J. Hum. Genet.* **110**, 1207–1215 (2023).

46. Abdulkadir, M. *et al.* Eating disorder symptoms and their associations with anthropometric and psychiatric polygenic scores. *Eur. Eat. Disord. Rev.* **30**, 221–236 (2022).

47. Watson, H. J. *et al.* Common genetic variation and age of onset of anorexia nervosa. *Biological Psychiatry Global Open Science* **2**, 368–378 (2022).

48. Padrutt, E. R. *et al.* Pubertal timing and adolescent outcomes: investigating explanations for associations with a genetically informed design. *J. Child Psychol. Psychiatry* **64**, 1232–1241 (2023).

49. Bralić, I. *et al.* Association of early menarche age and overweight/obesity. *J. Pediatr. Endocrinol. Metab.* **25**, 57–62 (2012).

50. Klump, K. L. Puberty as a critical risk period for eating disorders: a review of human and animal studies. *Horm. Behav.* **64**, 399–410 (2013).

51. Hübel, C. *et al.* Persistent thinness and anorexia nervosa differ on a genomic level. *Eur. J. Hum. Genet.* **32**, 117–124 (2024).

52. Müller, M. J. From BMI to functional body composition. *Eur. J. Clin. Nutr.* **67**, 1119–1121 (2013).

53. Akingbuwa, W. A. & Nivard, M. G. Detecting Non-linear Dependence through Genome Wide Analysis. *BioRxiv* (2025) doi:10.1101/2025.02.12.637804.

54. NCD Risk Factor Collaboration (NCD-RisC). Rising rural body-mass index is the main driver of the global obesity epidemic in adults. *Nature* **569**, 260–264 (2019).

55. Dueñas, H. R., Seah, C., Johnson, J. S. & Huckins, L. M. Implicit bias of encoded variables: frameworks for addressing structured bias in EHR-GWAS data. *Hum. Mol. Genet.* **29**, R33–R41 (2020).

56. Johnson, J. S. *et al.* Mapping anorexia nervosa genes to clinical phenotypes. *Psychol. Med.* **53**, 2619–2633 (2023).

57. Agrawal, A. *et al.* The Psychiatric Genomics Consortium: discoveries and directions. *Lancet Psychiatry* **12**, 600–610 (2025).

58. Zhou, W. *et al.* Efficiently controlling for case-control imbalance and sample relatedness in large-scale genetic association studies. *Nat. Genet.* **50**, 1335–1341 (2018).

59. Mbatchou, J. *et al.* Computationally efficient whole-genome regression for quantitative and binary traits. *Nat. Genet.* **53**, 1097–1103 (2021).

60. Chen, W. *et al.* Improved analyses of GWAS summary statistics by reducing data heterogeneity and errors. *Nat. Commun.* **12**, 7117 (2021).

61. Lam, M. *et al.* RICOPILI: rapid imputation for consortias pipeline. *Bioinformatics* **36**, 930–933 (2020).

62. Willer, C. J., Li, Y. & Abecasis, G. R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190–2191 (2010).

63. Yang, J. *et al.* Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nat. Genet.* **44**, 369–75, S1 (2012).

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

65. Ekeroth, K., Clinton, D., Norring, C. & Birgegård, A. Clinical characteristics and distinctiveness of DSM-5 eating disorder diagnoses: findings from a large naturalistic clinical database. *J. Eat. Disord.* **1**, 31 (2013).

66. International HapMap 3 Consortium *et al.* Integrating common and rare genetic variation in diverse human populations. *Nature* **467**, 52–58 (2010).

67. Grotzinger, A. D., Fuente, J. de la, 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).

68. 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).

69. Hübel, C. *et al.* Genomics of body fat percentage may contribute to sex bias in anorexia nervosa. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* **180**, 428–438 (2019).

70. 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).

71. R Core Team and Others. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. *Available* (2013).

72. Aschard, H., Vilhjálmsson, B. J., Joshi, A. D., Price, A. L. & Kraft, P. Adjusting for heritable covariates can bias effect estimates in genome-wide association studies. *Am. J. Hum. Genet.* **96**, 329–339 (2015).

73. Hemani, G. *et al.* The MR-Base platform supports systematic causal inference across the human phenome. *eLife* **7**, (2018).

74. Verbanck, M., Chen, C.-Y., Neale, B. & Do, R. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. *Nat. Genet.* **50**, 693–698 (2018).

75. Haycock, P. C. *et al.* Best (but oft-forgotten) practices: the design, analysis, and interpretation of Mendelian randomization studies. *Am. J. Clin. Nutr.* **103**, 965–978 (2016).

76. Bowden, J., Davey Smith, G. & Burgess, S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int. J. Epidemiol.* **44**, 512–525 (2015).

77. Rees, J. M. B., Wood, A. M., Dudbridge, F. & Burgess, S. Robust methods in Mendelian randomization via penalization of heterogeneous causal estimates. *PLoS ONE* **14**, e0222362 (2019).

78. Barbeira, A. N. *et al.* Exploiting the GTEx resources to decipher the mechanisms at GWAS loci. *Genome Biol.* **22**, 49 (2021).

79. Gamazon, E. R. *et al.* A gene-based association method for mapping traits using reference transcriptome data. *Nat. Genet.* **47**, 1091–1098 (2015).

80. Huckins, L. M. *et al.* Gene expression imputation across multiple brain regions provides insights into schizophrenia risk. *Nat. Genet.* **51**, 659–674 (2019).

81. PredictDB Team. CommonMind consortium - Brain Dorsolateral Prefrontal Cortex. (2019).

82. GTEx Consortium. The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science* **369**, 1318–1330 (2020).

83. Gaspar, H. A. & Breen, G. Drug enrichment and discovery from schizophrenia

genome-wide association results: an analysis and visualisation approach. *Sci. Rep.* **7**, 12460 (2017).

84. Durinck, S., Spellman, P. T., Birney, E. & Huber, W. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat. Protoc.* **4**, 1184–1191 (2009).

85. Freshour, S. L. *et al.* Integration of the Drug-Gene Interaction Database (DGIdb 4.0) with open crowdsource efforts. *Nucleic Acids Res.* **49**, D1144–D1151 (2021).

86. Roth, B. L., Lopez, E., Patel, S. & Kroeze, W. K. The multiplicity of serotonin receptors: uselessly diverse molecules or an embarrassment of riches? *Neuroscientist* **6**, 252–262 (2000).

87. Mendez, D. *et al.* ChEMBL: towards direct deposition of bioassay data. *Nucleic Acids Res.* **47**, D930–D940 (2019).

88. Sheils, T. K. *et al.* TCRD and Pharos 2021: mining the human proteome for disease biology. *Nucleic Acids Res.* **49**, D1334–D1346 (2021).

89. Yoo, M. *et al.* DSigDB: drug signatures database for gene set analysis. *Bioinformatics* **31**, 3069–3071 (2015).

90. WHO Collaborating Centre for Drug Statistics Methodology. ATC/DDD Index 2023. (2023).

91. GTEx Consortium *et al.* Genetic effects on gene expression across human tissues. *Nature* **550**, 204–213 (2017).

92. 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).

93. Finucane, H. K. *et al.* Heritability enrichment of specifically expressed genes identifies disease-relevant tissues and cell types. *Nat. Genet.* **50**, 621–629 (2018).

94. Ge, T., Chen, C.-Y., Ni, Y., Feng, Y.-C. A. & Smoller, J. W. Polygenic prediction via Bayesian regression and continuous shrinkage priors. *Nat. Commun.* **10**, 1776 (2019).

95. Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).

96. Choi, S. W., Mak, T. S.-H. & O'Reilly, P. F. Tutorial: a guide to performing polygenic risk score analyses. *Nat. Protoc.* **15**, 2759–2772 (2020).

97. Lee, S. H., Goddard, M. E., Wray, N. R. & Visscher, P. M. A better coefficient of determination for genetic profile analysis. *Genet. Epidemiol.* **36**, 214–224 (2012).

98. Robin, X. *et al.* pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics* **12**, 77 (2011).

99. CRAN: Package pracma. https://doi.org/10.32614/CRAN.package.pracma.

100. Nakai, Y., Nin, K. & Goel, N. J. The changing profile of eating disorders and related sociocultural factors in Japan between 1700 and 2020: A systematic scoping review. *Int. J. Eat. Disord.* **54**, 40–53 (2021).

TABLES

Phenotype	Inclusion criteria: Self-report or questionnaires	Inclusion criteria: Diagnostic codes
BE-BROAD *Binge eating broadly defined* (incl. BE-NARROW)	1. Eating an unusually large amount of food in a short period of time with loss of control. Frequency and duration criteria not required, or; 2. If assessed by a single item like "Have you ever experienced binge eating?"*, or; 3. A self-reported diagnosis of bulimia nervosa or binge eating disorder, confirmation by healthcare professional not required**.	307.51 (ICD-9, DSM-III-R, DSM-IV) F50.2, F50.3 (ICD-10) 307.51 [F50.2] (DSM-5) 307.51 [F50.3] (DSM-5) F50.81 (ICD-10-CM)
BE-NARROW *Binge eating narrowly defined*	1. Eating an unusually large amount of food in a short period of time with loss of control, and these episodes occurred on average at least once a week for at least three months, or; 2. A clinical diagnosis of bulimia nervosa or binge eating disorder via hospital/register records or a structured clinical interview, or; 3. A self-reported diagnosis of bulimia nervosa or binge eating disorder with confirmation by a healthcare professional.	307.51 (ICD-9, DSM-III-R, DSM-IV) F50.2 (ICD-10) F50.81 (ICD-10-CM) 307.51 [F50.2] (DSM-5)
AN *AN - no subtype specified* (incl. AN-R and AN-BP)	1. A clinical diagnosis of AN / AN binge eating/purging subtype / AN restricting subtype via hospital/register records, structured clinical interviews, or online questionnaires based on standardised criteria, or; 2. A self-reported diagnosis of AN / AN binge eating/purging subtype / AN restricting subtype, confirmation by healthcare professional not required. Note: atypical AN (without significantly low body weight) was not specifically excluded from our analyses, but has been historically ill-defined and was not ascertained for in most cohorts. As such, our AN data primarily reflect typical AN.	306.50 (ICD-8) 307.1 (ICD-9, DSM-III-R, DSM-IV) F50.0 (ICD-10) F50.1 (ICD-10) 307.1 [F50.1] (DSM-5)
AN-R *AN - restricting subtype*	1. A clinical diagnosis of AN restricting subtype via hospital/register records, structured clinical interviews, or online questionnaires based on standardised criteria, or; 2. A self-reported diagnosis of AN restricting subtype, confirmation by healthcare professional not required	F50.01 (ICD-10) 307.1 [F50.01] (DSM-5)
AN-BP *AN - binge eating/ purging subtype*	1. A clinical diagnosis of AN binge eating/purging subtype via hospital/register records, structured clinical interviews, or online questionnaires based on standardised criteria, or; 2. A self-reported diagnosis of AN binge eating/purging subtype, confirmation by healthcare professional not required	F50.02 (ICD-10) 307.1 [F50.02] (DSM-5)
Controls	1. No history of any eating disorder and no history of binge eating (broadly or narrowly defined), or; 2. No history of any eating disorder, or; 3. Unscreened (i.e., these individuals may have had an eating disorder or binge eating).	

Table 1: Phenotype definitions. Diagnostic codes are shown with coding system in round brackets, and equivalent codes in square brackets. Control definitions are shown in order of preference. Footnotes: *If assessed with terms such as "psychological overeating", "compulsive eating", etc., these individuals will not be included as BE-BROAD cases. **If Other Specified Feeding or Eating Disorder or Eating Disorder Not Otherwise Specified is diagnosed, these are included only if clearly stated 'subthreshold bulimia nervosa' or 'subthreshold binge eating disorder'.

Phenotype	Locus	CHR	Locus Start	Locus End	Lead SNP	P	REF/ ALT	OR	SE	ALT Freq Cases	ALT Freq Controls	INFO	Previously associated	Genes
BE-BROAD	1	1	74,977,295	75,014,362	rs7541513	2.65×10^{-9}	G/A	1.05	0.009	0.427	0.425	0.952	-	*LRRC53 / TNNI3K*
	2	3	117,420,697	117,953,697	rs1456193	3.25×10^{-8}	T/C	1.06	0.011	0.808	0.802	0.996	-	Intergenic
	3	11	93,171,140	93,231,940	rs2925354	4.91×10^{-8}	G/A	0.93	0.013	0.107	0.112	0.995	-	*DEUP1 / SMCO4*
	4	12	50,169,080	50,285,780	rs7953534	1.40×10^{-8}	G/A	1.05	0.009	0.351	0.333	0.985	-	Multigenic
	5	16	53,797,904	53,845,494	rs11642015	2.97×10^{-17}	C/T	1.07	0.009	0.429	0.417	1.000	-	*FTO*
	6	18	57,730,978	57,914,978	rs66723169	5.02×10^{-9}	C/A	1.06	0.010	0.254	0.209	0.995	-	Intergenic
AN	1	1	96,700,455	97,285,455	rs10747478	1.43×10^{-8}	T/G	0.94	0.011	0.570	0.587	0.998	Yes[20]	*PTBP2*
	2	3	47,501,450	51,741,450	rs113519699	2.86×10^{-18}	A/C	1.18	0.019	0.102	0.087	0.997	Yes[20]	Multigenic
	3	3	70,602,234	71,074,242	rs9310201	3.12×10^{-12}	A/T	1.08	0.011	0.427	0.408	0.982	Yes[20]	*FOXP1*
	4	5	24,868,440	25,300,440	rs6872919	1.07×10^{-9}	A/C	1.07	0.011	0.570	0.532	0.997	Yes[20]	Intergenic
	5	10	75,850,972	76,524,972	rs10740439	1.91×10^{-9}	C/T	0.93	0.012	0.711	0.716	0.994	No	Multigenic
	6	10	131,274,055	131,459,255	rs12762024	3.28×10^{-9}	G/C	1.07	0.011	0.447	0.455	0.995	Yes[20]	*MGMT*
	7	11	114,997,256	115,284,956	rs6589488	9.80×10^{-9}	A/T	0.91	0.016	0.846	0.875	0.990	Yes[20]	*CADM1*
	8	12	17,635,314	17,965,314	rs12817084	8.93×10^{-9}	T/C	1.10	0.011	0.115	0.117	0.991	No	Intergenic

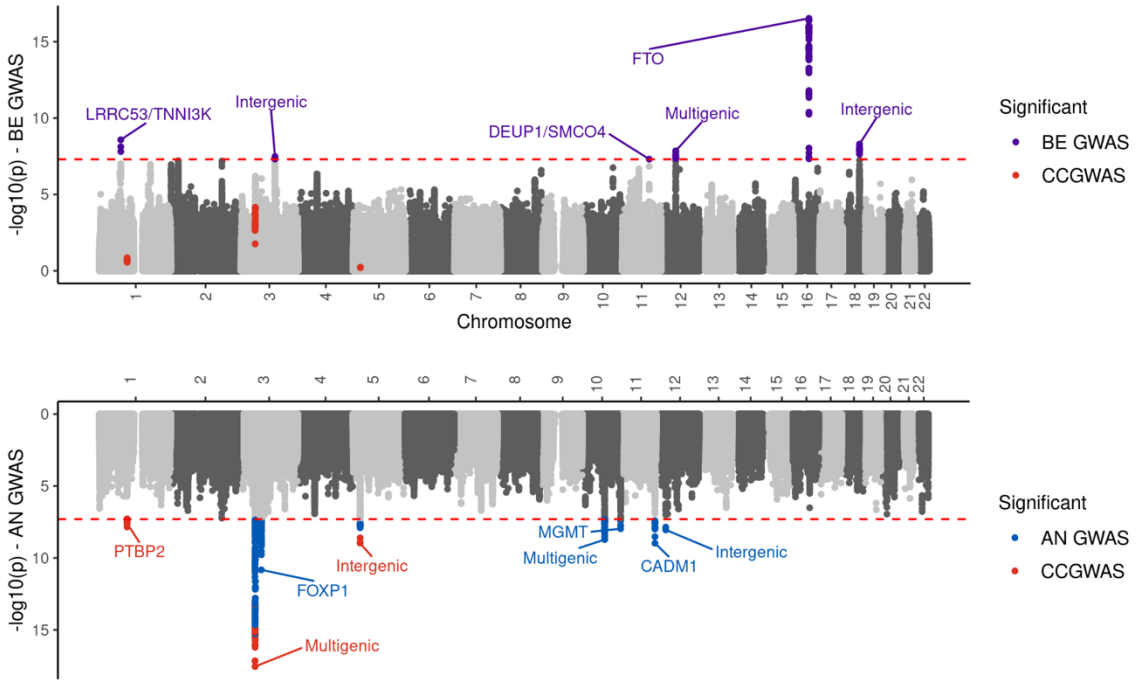
Table 2: Association statistics for genome-wide significant loci for BE-BROAD and AN. Loci are numbered sequentially within-analysis from chromosome 1 to chromosome X. Base pair start and end positions correspond to hg19. Odds ratios (OR) are given relevant to the ALT allele. Genes are listed if at least one transcript in GENCODE version 47 lies at least partially within the locus, with >3 genes defined as multigenic.

Trait	Category	PMID	rg (AN)	SE (AN)	*P* (AN)	rg (BE)	SE (BE)	*P* (BE)
BMI	Anthropometric	30239722	-0.31	0.020	**2.37×10^{-56}**	0.36	0.032	**1.94×10^{-28}**
Body fat %	Anthropometric	30593698	-0.35	0.028	**2.71×10^{-36}**	0.20	0.042	**2.65×10^{-6}**
Fat free mass	Anthropometric	30593698	-0.15	0.022	**2.93×10^{-11}**	0.30	0.034	**6.92×10^{-19}**
Fat mass	Anthropometric	31852892	-0.32	0.023	**3.63×10^{-43}**	0.31	0.035	**1.36×10^{-18}**
Persistent thinness	Anthropometric	30677029	0.11	0.074	0.123	-0.46	0.110	**2.37×10^{-5}**
Pubertal growth (height diff, 8 years-adult)	Anthropometric	23449627	-0.03	0.053	0.606	-0.37	0.079	**3.08×10^{-6}**
Severe early-onset obesity	Anthropometric	30677029	-0.11	0.059	0.0587	0.48	0.080	**1.40×10^{-9}**
Waist-hip ratio	Anthropometric	30239722	-0.25	0.021	**1.72×10^{-34}**	0.19	0.034	**1.37×10^{-8}**
Waist-hip ratio (BMI-adj)	Anthropometric	30239722	-0.09	0.020	**9.33×10^{-6}**	-0.05	0.033	0.119
Fasting insulin (age- & sex-adj)	Metabolism	33402679	-0.30	0.048	**2.79×10^{-10}**	0.15	0.069	0.0292
Fasting insulin (BMI-adj)	Metabolism	34059833	-0.17	0.035	**9.62×10^{-7}**	-0.29	0.047	**6.29×10^{-10}**
HbA1c	Metabolism	28898252	-0.14	0.042	6.00×10^{-4}	-0.03	0.053	0.5242
HbA1c (BMI-adj)	Metabolism	34059833	-0.16	0.034	**5.22×10^{-6}**	0.01	0.044	0.765
Type 2 diabetes	Metabolism	30297969	-0.18	0.024	**1.76×10^{-14}**	0.11	0.040	6.10×10^{-3}
ADHD	Psychiatric	36702997	0.07	0.029	0.0126	0.31	0.042	**2.20×10^{-13}**
Autism	Psychiatric	32747698	0.17	0.044	**1.00×10^{-4}**	0.29	0.057	**3.96×10^{-7}**
Bipolar disorder	Psychiatric	34002096	0.25	0.027	**3.80×10^{-21}**	0.23	0.035	**4.22×10^{-11}**
Insomnia	Psychiatric	30804565	0.10	0.033	2.90×10^{-3}	0.21	0.044	**2.02×10^{-6}**
MDD	Psychiatric	39814019	0.30	0.025	**1.04×10^{-33}**	0.39	0.031	**1.35×10^{-36}**
OCD	Psychiatric	28761083	0.48	0.069	**2.85×10^{-12}**	0.08	0.081	0.322
Probable anxiety diagnosis	Psychiatric	31748690	0.28	0.039	**1.42×10^{-12}**	0.32	0.060	**1.03×10^{-7}**
PTSD	Psychiatric	33510476	0.11	0.049	0.0274	0.24	0.062	**1.00×10^{-4}**
Schizophrenia	Psychiatric	35396580	0.24	0.021	**4.85×10^{-30}**	0.23	0.033	**7.04×10^{-12}**
General risk tolerance	Behavioural	30643258	0.00	0.028	0.991	0.21	0.040	**6.76×10^{-8}**
Loneliness	Behavioural	31518406	0.13	0.032	**4.96×10^{-5}**	0.37	0.044	**4.14×10^{-17}**
Physical activity	Behavioural	36071172	0.19	0.034	**5.87×10^{-8}**	0.09	0.046	0.0633
Educational attainment (years)	Socio-demographic	30038396	0.25	0.022	**1.25×10^{-29}**	0.04	0.030	0.1417
AUDIT-P	Substance use	30336701	0.01	0.040	0.870	0.37	0.057	**7.00×10^{-11}**
Cannabis use disorder	Substance use	33096046	0.02	0.046	0.610	0.30	0.063	**2.78×10^{-6}**
Heavy smoker	Substance use	28166213	-0.04	0.039	0.255	0.30	0.048	**3.82×10^{-10}**
Lifetime cannabis use	Substance use	30150663	0.22	0.038	**8.48×10^{-9}**	0.32	0.051	**6.36×10^{-10}**

Table 3: A representative subset of significant genetic correlations of BE-BROAD and AN with external traits. *Note.* The full set of genetic correlations can be found in Supplementary Table 9. PMID = PubMed ID of GWAS for external trait, BMI = body mass index, adj = adjusted, ADHD = attention deficit/hyperactivity disorder, MDD = major depressive disorder, OCD = obsessive-compulsive disorder, PTSD = post-traumatic stress disorder, AUDIT-P = Alcohol Use Disorder Identification Test Problem items.

Figures

a)



b)

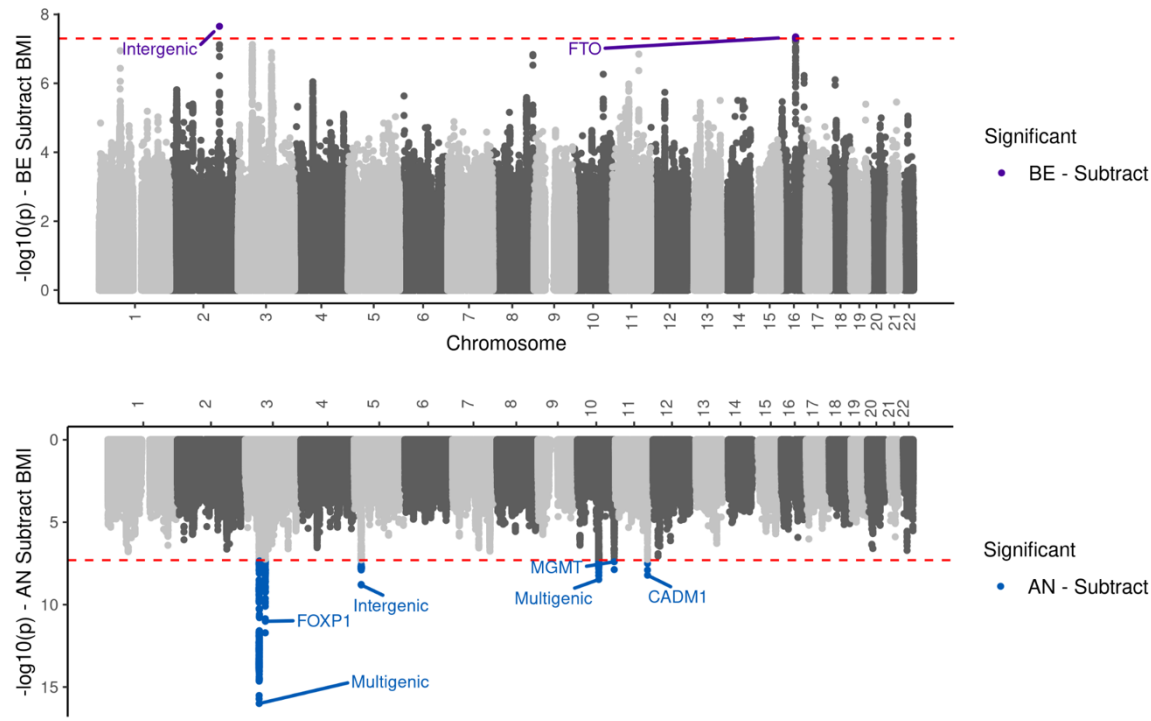
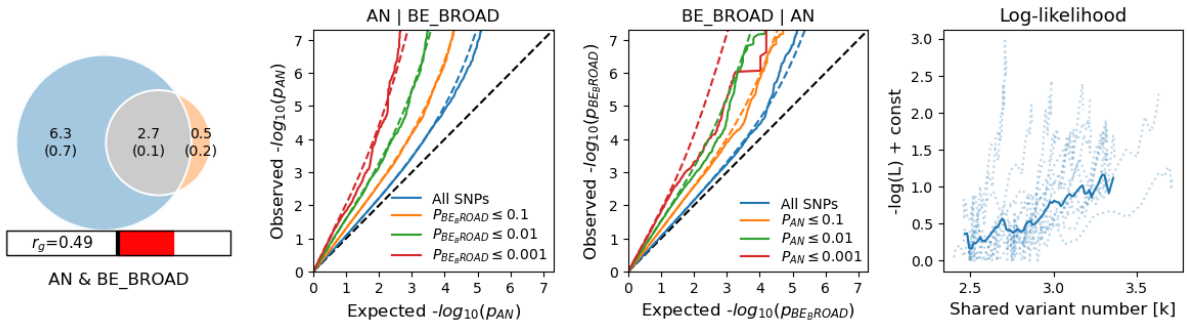


Figure 1: Miami plots showing results from the BE-BROAD (top) and AN (bottom) meta-analyses. The dotted red line is the genome-wide significance threshold ($P \leq 5 \times 10^{-8}$).
a. Main GWAS analyses, with variants reaching genome-wide significance coloured in blue if significant in the anorexia nervosa GWAS and in purple if significant in the binge eating broad GWAS. Variants reaching genome-wide significance in case-case GWAS of BE-BROAD vs AN are coloured in red.
b. GWAS-by-subtraction analyses, showing results from the non-BMI genetic component. Variants reaching genome-wide significance coloured in blue for anorexia nervosa non-BMI component and purple in binge eating non-BMI component.

a)



b)

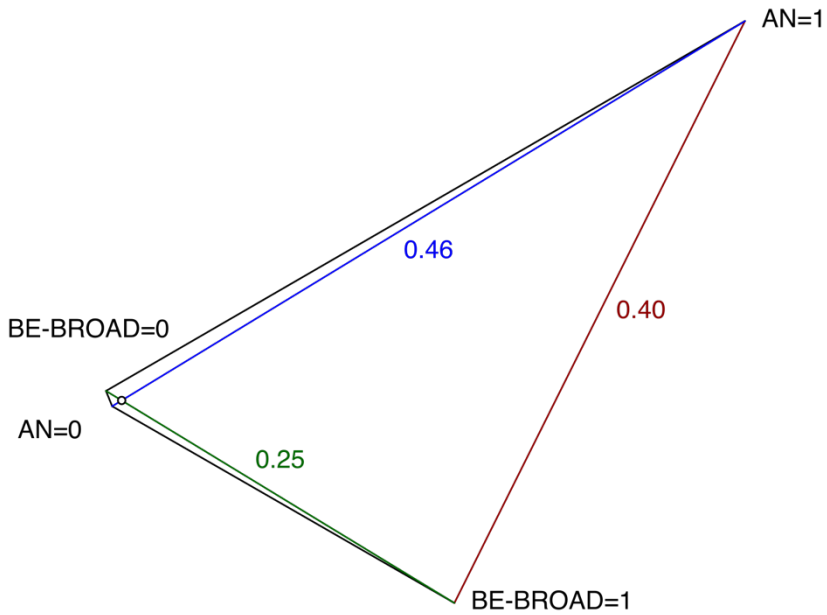


Figure 2: a) Bivariate causal mixture modelling results from MiXeR. (Left) Venn diagram showing overlap of putative causal variants in AN (blue) and BE BROAD (orange), demonstrating high overlap (grey), and moderate genetic correlation (bar beneath). (Middle) Q-Q plots of variant effects for (left) AN conditional on BE BROAD and (right) BE BROAD conditional on AN, demonstrating enrichment for associations with each trait conditional on the other. (Right) Log-likelihood plot of model fit under differing models, from the minimal overlap model (leftmost) to the full overlap model (rightmost), showing best model fit (lowest y-axis value) is not distinguishable from the minimal overlap.
b) Genetic distance between cases and controls of BE-BROAD and AN estimated by case-case GWAS. The genetic distance, $\sqrt{m \times F_{ST,causal}}$ is calculated by taking the square root of the product of *m*, the number of independent causal variants estimated here as 10,000 based on the polygenic nature of BE-BROAD and AN, and $F_{ST,causal}$, the average normalized squared differences in allele frequencies, derived based on the SNP-based heritabilities, genetic correlations, and population prevalences of the two traits.

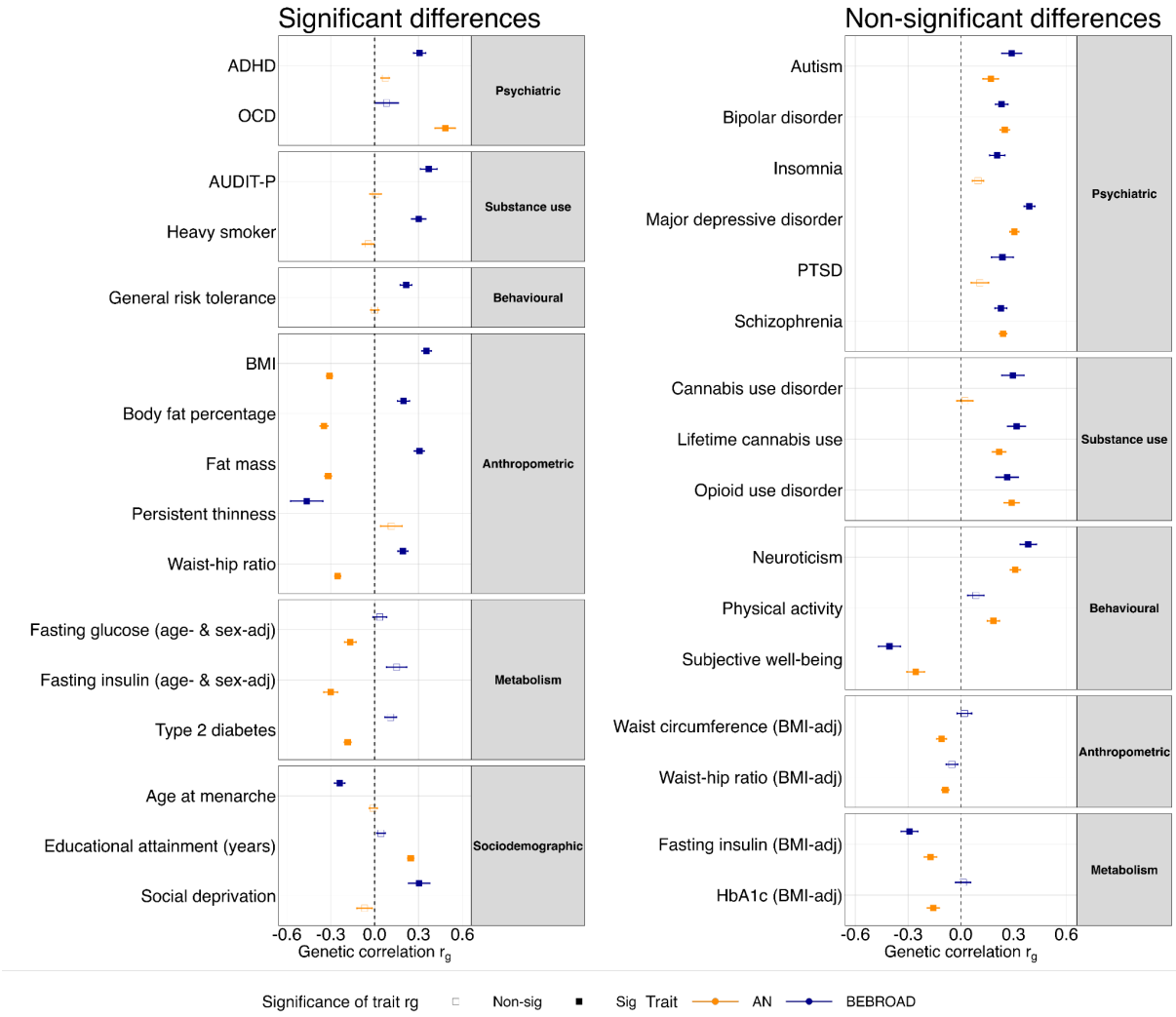


Figure 3: Genetic correlations (r_g) of selected external traits with BE-BROAD and AN, split by those that differ significantly between the eating phenotypes (left) and those that do not (right). The r_gs were computed by Linkage Disequilibrium Score Regression (LDSC). The r_g estimates are indicated by the dots and standard errors are indicated by the lines on either side of each dot. r_g estimates have been corrected for multiple testing via the Bonferroni method. Information about the summary statistics used in our analysis can be found in Supplementary Table 9.

BE-BROAD = binge eating broad definition; AN = anorexia nervosa; MDD = major depressive disorder; PGC = Psychiatric Genomics Consortium; BMI = Body mass index; F = female; M = male; FFM = fat-free mass; AUDIT-P = Alcohol Use Disorder Identification Test problem items

METHODS

Ethics

The individual studies that comprise this investigation were conducted with advance approval by the appropriate Institutional Review Boards or equivalents at the individual study sites. We provide ethical statements for each study site in the Supplementary Note. This work represents a secondary analysis with data from these individual studies.

Summary of cohorts

Detailed descriptions of the ascertainment and definition of cases and controls for each cohort is provided in Supplementary Table 1. Broadly, we identified cases and controls based on clinical diagnoses, diagnostic algorithms, and/or self-report questionnaires [43]. We defined controls as individuals without a history of BE and without a history of an eating disorder, if possible. If this information was unavailable, unscreened controls were included assuming that the large control numbers would outweigh the impact of misclassified individuals in the control groups, given the collective lifetime prevalence of eating disorders is ~5% [17].

We included data from the Eating Disorders Working Group of the Psychiatric Genomics Consortium (PGC-ED; Supplementary Table 1). These data were restricted to individuals of European ancestry due to the limited availability of non-European ancestry samples at the time of analysis—we included two cohorts with individuals of East Asian ancestry for follow-up cross-ancestry polygenic risk score analyses. In total, we combined 27 European ancestry datasets totalling 14 previously analysed [6] and 13 new cohorts. Data from cohorts providing individual-level data (*n*=11) were combined with cohorts that contributed summary statistics (*n*=16; Supplementary Table 1). Detailed descriptions of each of the cohorts are provided in the Supplementary Note. We included data if the total number of cases for any phenotype prior to quality control was >100. If cases for an individual phenotype were <50, we excluded that phenotype from analyses.

We included five different phenotypes: broadly-defined BE (BE-BROAD, 39,279 cases), narrowly-defined BE (BE-NARROW, 15,175 cases), AN (24,145 cases), AN restricting subtype (AN-R, 2524 cases), and AN binge eating/purging subtype (AN-BP, 5245 cases). Detailed phenotypic definitions are provided in Table 1.

Genotype quality control and imputation

Approaches to quality control and imputation differed between cohorts, and are described in detail in the Supplementary Material.

Association analyses

The statistical model used to conduct the GWAS for each cohort (described in Supplementary Table 4) depended on the design of the specific cohort. For unrelated case/control cohorts, we used PLINK2 to conduct logistic regression using an appropriate number of PCs to account for ancestry as necessary within each cohort (Supplementary Table 4). For related or imbalanced case/control cohorts we used SAIGE [58] or REGENIE [59]. Note that for whole genome regression in REGENIE step 1, we used a set of pruned SNPs with MAF>0.01, excluding high LD regions and only including autosomal chromosomes (PLINK command: --indep-pairwise 1500 150 0.2). Further details on cohort-specific aspects of association analysis are provided in the Supplementary Methods. All analyses were two-tailed.

Post-GWAS processing and quality control

We aligned the summary statistics of each GWAS to the TOPMed reference panel in Genome Reference Consortium Build 37 (GRCh37/hg19), using variant positions from ENSEMBL (see URLs). For cohorts that were in GRCh38, we first linked the datasets with the GRCh38 TOPMed reference panel (see URLs) by chromosome/base pair, then selected the SNP rsID labels and linked these labels to the GRCh37 TOPMed reference panel, and

finally extracted the GRCh37 chromosome/base pair information for each SNP. Variants without a rsID label in the GRCh38 TOPMed reference panel were lifted over to GRCh37 using the liftOver tool (see URLs). After alignment, we applied an INFO and MAF filter to include SNPs with an INFO score of>0.3 and MAF>0.01 in cases and controls. We then used DENTIST [60] to remove variants with effects inconsistent with their linkage disequilibrium pattern with other assessed variants, estimating linkage disequilibrium from European ancestry individuals in Phase 3 of the 1000 Genomes project.

Meta-analysis and quality control

We used the post-imputation module of Ricopili [61] to perform meta-analyses in METAL [62] using an inverse-variance weighted fixed-effect model, including assessment of variant heterogeneity as I^2 values (Comparison of variant heterogeneity between our AN GWAS and that of Watson et al is included in the Supplementary Methods [6].) We defined independent significant SNPs with a genome-wide significant p-value ($P<5x10^{-8}$) that were independent ($r^2<0.6$) from each other. We then defined significant genomic loci by merging LD blocks of these independent significant SNPs if they were close to each other (<250 kb). Furthermore, we defined independent lead SNPs if independent significant SNPs were independent of each other at $r^2<0.1$.

We ran a stepwise conditional analysis on our GWAS results to select independently-associated SNPs (conditional on lead SNPs) at each loci using GCTA-COJO [63]. For these analyses we used one of our largest cohorts (*usa2*) as our reference for linkage disequilibrium.

Female-only analyses and female-to-male polygenic risk scoring

We conducted a supplementary female-only GWAS for BE-BROAD and AN and generated female-only PRS using PRS-CS which we applied on male-only datasets with sufficient data (i.e., *n* case and *n* control >100) available. For the BE-BROAD female-only

meta-analysis, all cohorts except *usa1* and *biov* were included, and *alsp, moba,* and *ukd2* were included as male-only target cohorts. For the AN female-only meta-analysis, all cohorts except *itgr*, *spa1, ukd1,* and *net2* were included, and *ipsy, fngn,* and *ukb2* were included as male-only target cohorts.

SNP-based heritability and distinguishing polygenicity from other sources of inflation

We used linkage disequilibrium score regression (LDSC [64]) to estimate SNP-based heritability (h^2_{SNP}). These estimates were transformed to the liability scale, assuming population prevalences as follows: BE-BROAD 4.5% [15], BE-NARROW 3.5% [15], AN 1.5% [17], AN-R 0.8% [17,65], AN-BP 0.7% [17,65]. For all analyses using LDSC, we applied an LD reference panel based on the European subset of the 1000 Genomes Project (1kGP), restricted to SNPs present in the HapMap3 panel [66]. For *N,* we calculated the sum of effective *N* across all cohorts and specified 0.5 for sample prevalence [67]. All LDSC analyses, including tests of the of h^2_{SNP} and SNP-r_g from 0 and from 1, were two-tailed tests of deviation from a chi-square distribution with jack-knifed standard errors.

Test statistics from GWAS of a polygenic trait are expected to be inflated, but inflation may also be due to spurious SNP associations caused by population stratification and cryptic relatedness of study participants. We used statistics from LDSC [64] to determine the source of inflation. Although the LDSC intercept is commonly used to distinguish polygenicity from spuriously inflated statistics, we calculated the attenuation ratio statistic, defined as (LDSC intercept–1)/(mean of association chi-square statistics–1), which may be a less biassed metric compared to the LDSC intercept [16]. We included variants with MAF≥0.01 and INFO≥0.6.

Comparison of the two binge-eating behaviour phenotypes

The two BE phenotype definitions balanced phenotypic certainty with sample size—BE-BROAD had a larger sample size and was potentially a better-powered GWAS than BE-

NARROW but was likely to be more heterogeneous and could lack specificity to BE. To determine which phenotype to carry forward to follow-up analyses, we considered the LDSC intercept and attenuation ratio statistics and calculated the genetic correlation (SNP-r_g) between the two BE phenotypes. Inflation was a more sizable component of the signal in BE-NARROW (attenuation ratio 0.21 ± 0.06) than in BE-BROAD (0.14 ± 0.04). Furthermore, the SNP-r_g between BE-NARROW and BE-BROAD did not differ from unity (1.00 ± 0.03). We therefore concluded that BE-BROAD appropriately captured the common genetic component of BE with greater statistical power than BE-NARROW.

Genetic relationship between traits

We used LDSC to calculate SNP-r_g with several aims. First, we estimated SNP-r_gs between both BE phenotypes, AN, and AN subtypes to assess the genetic relationships among these eating phenotypes. Second, we calculated SNP-r_g between BE-BROAD/AN and 225 traits covering eight categories: 1) psychiatric trait or disorder, 2) substance use, 3) psychological/personality/behavioural, 4) anthropometric, 5) metabolism, 6) blood, 7) sociodemographic, and 8) somatic trait or disease. We selected these traits from an internal catalogue based on their power (h^2_{SNP} Z-score>4) [68]. To assess statistical significance, we tested whether SNP-r_g differed from zero and applied a Bonferroni-corrected p-value threshold of 2.20×10^{-4} based on 225 traits. If, according to this threshold, BE-BROAD and/or AN was significantly genetically correlated with another trait, we used the LDSC block-jackknife procedure (described in more detail in Appendix S1 of Hübel et al [69]) to statistically compare the SNP-r_g between BE-BROAD and AN (Bonferroni-corrected p-value threshold of 2.20×10^{-4}).

We used MiXeR to conduct univariate and bivariate casual mixture modelling of BE-BROAD and AN [19]. MiXeR is an extension of LDSC that uses GWAS summary statistics to model the number of causal variants underlying a trait (polygenicity), as well as inferring the average contribution of each causal variant to heritability (discoverability). Bivariate MiXeR

extends cross-trait LDSC to refine the interpretation of genetic correlation. It enables the shared number of causal variants between traits to be estimated, as well as the extent to which shared causal variants have the same direction of effect. We conducted MiXeR analyses using BE-BROAD and AN summary statistics, limited to autosomal variants with INFO≥0.6 and MAF≥0.01, and with the MHC region (chr6, 26-34Mb) excluded. We followed protocols and used LD reference files provided by the authors of MiXeR (see URLs).

To further investigate the genetic difference between BE-BROAD and AN we used CC-GWAS [18]. CC-GWAS uses GWAS summary statistics to test for allele frequency differences between cases of two phenotypes, as opposed to a traditional GWAS that tests for differences between cases and controls. It generates Z-scores and p-values for allele frequency differences between cases of two phenotypes as a two-tailed test, using both ordinary least squares (OLS) and exact weights. We used BE-BROAD and AN summary statistics including non-ambiguous SNPs from HapMap 3 with INFO≥0.6 and MAF≥0.01. We used the liability-scale h^2_{SNP}, the SNP-r_g and associated covariate intercept between both traits as input. We furthermore set the BE-BROAD population prevalence to 4.5% (range 0.1 - 10%) and AN to 1.5% (range 0.1 - 4.3%) and approximated the number of effective loci to be 10,000–consistent with psychiatric disorder polygenicity [18]. We additionally defined the number of independent CC-GWAS loci with PLINK 1.9 (--clump-p1 5e-8 --clump-p2 5e-8 --clump-r2 0.1 --clump-kb 3000) and defined a genome-wide significant SNP if $P<5\times10^{-8}$ in the CC-GWAS OLS test and $P<10^{-4}$ in the CC-GWAS exact test. Further, CC-GWAS calculates genetic distances between cases and controls of the two traits using $\sqrt{m \times F_{ST,causal}}$ based on liability-scale h^2_{SNP}, SNP-r_g, population prevalence and the number of independent causal variants of the two traits.

Influence of BMI

We used GWAS-by-subtraction [20], an application of genomicSEM [70], in R v4.3.1 [71] to estimate the proportion of variance in BE-BROAD and AN independent of BMI

(Supplementary Methods). Shared genomic covariance across traits is expected due to pleiotropy. Latent Genomic SEM factors explicitly model this covariance, making results less influenced by spurious biases than would conditioning on phenotypic traits in a GWAS [72]. We used GWAS summary statistics from BE-BROAD, AN, and BMI [21].

We specified a structural equation model (SEM) that regressed both sets of summary statistics on a shared variable ("*BMI*") and a non-BMI variable ("*non-BMI*") for each eating phenotypes, respectively (Supplementary Figure 26). Specifically, we specified the two latent variables as a function of BMI and (e.g.) BE-BROAD: "*BMI* =~ *NA* * BEBROAD + start(0.4) * BMI" and *"non-BMI* =~ *NA* * BE-BROAD". In line with the GWAS-by-subtraction specification as it has previously been applied [20], we additionally set the variance of the latent variables to 1 and a covariance of 0, and constrained the model so all (co)variance in BMI and BE-BROAD was captured by *BMI* and *non-BMI*. We used the diagonally-weighted least squares estimator, which is the default setting in Genomic SEM [70]. Additional computational settings are shown in Supplementary Table 30. We then regressed the two latent factors on individual SNPs yielding a GWAS of the latent variables *BMI* and *non-BMI*. We subsequently used LDSC [64] to calculate SNP-r_g of the *non-BMI* factor with all traits identified in initial SNP-r_gs. As a sensitivity analysis, we restricted our BE-BROAD sample in our GWAS to cohorts that were not ascertained for AN, reasoning that this might better capture BE behaviour outside of AN. We applied the same GWAS-by-subtraction model on that selection of cohorts (listed in Supplementary Table 11).

We also conducted exploratory two-sample Mendelian randomisation analyses of BE-BROAD with BMI and AN with BMI, testing causal effects in both directions with two-tailed tests. We used SNPs in linkage equilibrium as genetic instruments, with $P<5x10^{-6}$ for BE-BROAD and AN, and $P<5x10^{-9}$ for BMI. The instrument used for BMI was that suggested by the authors of the BMI GWAS [21]. We repeated analyses using the *non-BMI* factor GWAS of BE-BROAD and of AN. We view these Mendelian randomisation analyses as exploratory because SNPs not passing genome-wide significance were included in the instruments for

the eating phenotypes. To ensure our analyses were robust to potential violations of the assumptions of Mendelian randomisation, we conducted analyses using multiple methods in R 4.3.2, including the packages *TwoSampleMR*, *MendelianRandomisation*, and *MR-PRESSO* (Supplementary Methods) [71,73,74]. These were inverse-variance weighted analysis, MR-Egger, mode-based estimation and median-based estimation. We determined the strength of association of our genetic instruments using the F-statistics [75]. We used Cochran's Q-statistic to test for instrument heterogeneity, with $P<0.05$ indicating heterogeneity [76]. To investigate potential confounding via horizontal pleiotropy, we assessed the deviation of the MR-Egger intercept from 0, and performed a global bias test in MR-PRESSO. We excluded from the analysis SNPs identified by MR PRESSO as pleiotropic. We ran further methods robust to heterogeneity, including the penalised weighted median estimator, the contamination mixture method, and MR-Lasso [77]. We assessed results visually (Supplementary Methods).

Identification of gene-tissue associations with eating phenotypes

We used S-PrediXcan [22] to identify genetically regulated gene expression associated with our phenotypes. We tested the association of gene expression with our eating phenotypes using available GTEx v8 MASHR [22,78,79] and CommonMind DLPFC [80,81] tissue models. MASHR-based PredictDB models use fine-mapping methods for selection of eQTLs included in the predictor models, improving prediction [22,78,79]. We included 45 GTEx v8 MASHR models, removing non-natural tissues (cell lines), tissues with N<100 individuals (kidney cortex), and testis [82]. We performed liftover of our GWAS summary statistics to hg38, harmonisation, and imputation based on recommended preprocessing by Barbeira et. al [78] using *GWAS tools* (see URLs). S-PrediXcan uses GWAS summary statistics to impute the difference in genetically regulated gene expression between levels of the GWAS phenotype (in this instance, cases and controls). We tested whether these imputed differences differed from 0 as a two-tailed Z test. We used two different Bonferroni

significance thresholds: an experiment-wide threshold, where we corrected for 600,382–602,744 tests performed across all tissues ($P<0.05/Tests_{Total} = P<8.32x10^{-8}$), and a tissue-specific threshold, where we corrected for varying numbers of tests performed within each tissue ($P<0.05/Tests_{Tissue\ X}$, Supplementary Table 16). We performed two-tailed exact binomial tests for tissue enrichment using *binom.test()* in R for associations at three different significance thresholds: experiment-wide significant ($P<0.05/Tests_{Total}$), tissue-specific significant ($P<0.05/Tests_{TissueX}$), and nominally significant ($P<0.05$). We reported results with a focus on central nervous system tissues (brain and cervical spinal cord) and gastrointestinal tissues (oesophagus, colon, stomach, and small intestine).

S-MultiXcan is a summary-level method for measuring the joint association of genetically regulated gene expression across tissues with a phenotype of interest, leveraging shared eQTLs across tissues [23]. Using our GTEx v8 MASHR S-PrediXcan results as input, along with MASHR models, and harmonised, imputed GWAS summary statistics, we ran S-MultiXcan on each of our eating phenotypes for all genes (N=22,241). S-MultiXcan gives as output the p-value for association of multi-tissue gene expression with the trait of interest (S-MultiXcan P), along with best single-tissue p-value, both from two-tailed tests. In order to account for potential false positive associations, we removed any significant S-MultiXcan associations where the single best tissue p-value was greater than $1x10^{-4}$ [23]. We set a Bonferroni significance threshold for our results, correcting for the number of genes tested in our S-MultiXcan analysis ($P<0.05/22,241 = 2.25x10^{-6}$).

Gene-wise and gene set analysis, including drug target and drug class analyses

Following a previously published approach [83], we used MAGMA v1.10 [24] to test the association between each phenotype and 1) the aggregate effect of SNPs mapped to protein-coding genes (gene analysis); 2) groups of genes with shared functional, biological, or other characteristics (gene-set analysis); 3) a gene-set analysis restricted to genes targeted by drugs (drug-set analysis) and 4) the enrichment of signal within classes of drug.

We restricted SNPs to those with MAF≥0.01, INFO≥0.6, and which are present in 80% of the total sample and 50% of the cohorts. We mapped SNPs to protein-coding genes, applying a 35 kb upstream and 10 kb downstream window around hg19 gene positions from Ensembl release 75 [84]. We obtained gene-wise p-values with the multi snp-wise model, which combines the lowest and mean p-values of all SNPs mapped to the gene. We tested 19,332-19,418 ENSEMBL genes across the five phenotypes and applied a Bonferroni correction of $P<2.60\times10^{-6}$. We used the 1kGP reference panel for estimating between-SNP LD.

For gene-set analyses, we applied a competitive analysis, which regresses the phenotype on the mean effect of genes within the gene set, with the mean effect of genes outside the gene set as a covariate, in a one-tailed test that the mean effect of genes within the gene set is greater than 0. We defined biological pathways based on gene ontology and canonical pathways from MSigDB v6.1 and psychiatric pathways identified from the literature. We tested 7324-7325 pathways across the five phenotypes and applied a Bonferroni correction of $P<6.83\times10^{-6}$.

For drug-set analyses, we defined drug sets based on drug targets from the Drug-Gene Interaction database DGIdb v4.2.0 [85]; the Psychoactive Drug Screening Database Ki DB [86]; CheMBL v27 [87]; the Target Central Resource Database v6.7.0 [88]; and DSigDB v1.0 [89], all downloaded in October 2020. We applied a competitive analysis and subsequently grouped the results based on the Anatomical Therapeutic Chemical class of the respective drugs [90]. For drug-class analysis, we first ranked all drug-gene sets according to their association in the drug-set analysis. We then generated enrichment curves for specific drug classes, assigning a ‘hit’ if the drug-gene set belonged to the class or a “miss” if it was outside the class. We calculated the area under the curve and determined statistical significance with the Wilcoxon Mann-Whitney test, comparing drug-gene sets within the class to those outside the class as a one-tailed test of whether in-class drugs showed greater association with the disorder than out-of-class drugs. We applied a Bonferroni correction of $P<3.23\times10^{-5}$ (based on 1546-1547 drug sets) for the drug-set analysis and

$P<3.08 \times 10^{-4}$ (based on 162 drug classes) for the drug-class analysis to account for multiple testing.

Tissue and cell-type specific analyses

To identify relevant tissues and cell-types related to the common genetic risk of BE-BROAD and AN, we performed tissue and cell-type heritability (h^2_{SNP}) enrichment analyses. First, we analysed the enrichment of h^2_{SNP} in 27 tissues from the GTEx gene expression data (v8) after excluding tissues with less than 100 donors, non-natural tissues (such as cell lines), and testis tissues (since it was an expression outlier) [91]. Second, we analysed enrichment of h^2_{SNP} in 31 superclusters and 461 cell clusters based on the single-nucleus RNA sequencing data including over three million nuclei from around 100 dissections across the adult human brain [29]. Within each expression dataset, we calculated the specificity of gene expression per tissue or cell type (superclusters and clusters separately), defined as the expression of each gene (counts per million, CPM) in a tissue or cell type (i.e., the superclusters and clusters respectively) divided by the total expression of this gene across all tissues or cell types in the dataset [92]. We then used the genes with the top 10% specificity in each tissue or cell type to perform the heritability enrichment analysis using stratified LDSC [92,93]. Specifically, we compared the per-SNP heritability of SNPs within 100kb flankings of the top 10% specific genes and the per-SNP heritability of other SNPs, using the baseline model that adjusted for 53 baseline annotations [93]. We then used the coefficient *z*-scores to calculate the one-sided p-values. Finally, we accounted for multiple comparisons by calculating FDR per trait for the GTEx dataset (27 tests), the human brain superclusters (31 tests) and clusters (461 tests) respectively.

Polygenic prediction

We used PRS-CS [94] to generate polygenic risk scores (PRS) for BE-BROAD and AN. First, we performed leave-one-cohort-out (LOO) analyses to generate LOO GWAS summary

statistics from all cohorts except the target cohort and used this as the base data to calculate individual-level PRS in each target cohort. We included non-ambiguous SNPs with INFO≥0.6 and MAF≥0.01 in the PRS calculation. We used the 1000 Genomes Project Phase 3 EUR reference as the LD reference panel and provided median sample size per LOO meta-analysis as input for PRS-CS. Posterior SNP effect size estimates from PRS-CS were then combined across chromosomes to calculate individual PRS via PLINK (--score 2 4 6 sum) [95]. We standardised the individual PRS by applying the scale function in R (version 4.3.2) [71]. Using the standardised PRS scores, we first assessed the proportion of variance explained by PRS for each phenotype through calculating the nested Nagelkerke's pseudo-R^2 (that is, the pseudo-R^2 of the full model minus that of the model excluding the PRS) [96]. The inclusion of target cohorts for each phenotype was based on the effective sample size (N_{eff} half>1000), study characteristics, and availability of individual-level data (more detailed information in the Supplementary Methods). Two-tailed logistic regression was performed for BE-BROAD and AN PRS on BE-BROAD and AN, adjusting for cohort-specific PCs. We then converted the variance explained by each PRS (R^2) to the liability scale using population prevalences of BE-BROAD and AN as used for LDSC [97]. In addition, we divided individuals into PRS decile groups and assessed their relative risk of BE-BROAD and AN compared to the group of individuals in the lowest PRS decile. We also examined predictive performance of each PRS regarding their sensitivity, specificity, and precision to predict BE-BROAD and AN status using the *pROC* [98] and *pracma* [99] packages in R by comparing the area under the curve (AUC) of the receiver operating characteristic (ROC) curve and the precision recall (PR) curve of the full model including PRS and PCs as predictors against the null model including PCs only as predictors.

We used female-only GWAS summary statistics as base data in PRS-CS [94] to assess whether female BE-BROAD PRS could be applied to male individuals to assess their BE-BROAD risk in three cohorts (*alsp, moba, ukd2*, $N_{case_male_total}$=1055, $N_{control_male_total}$=38,046).

The same analysis was performed to examine the association of female AN PRS with male AN risk in three cohorts (*ukb2, fngn, ipsy*, $N_{case_male_total}$=1524, $N_{control_male_total}$=388,891).

We assessed whether the levels of BE-BROAD and AN PRS differed across subgroups, including a) those with BE-BROAD only, b) those with both BE-BROAD and AN, and c) those with AN only. Target cohorts for this analysis included cohorts with sufficient data available on some/all subgroups (Supplementary Methods, Supplementary Table 1). Specifically, we selected *aunz* and *sedk* to assess individuals with both BE-BROAD and AN and to assess the AN only group. We selected *ukb2* and *ukd2* for assessing BE-BROAD and AN PRS levels in all subgroups. We assessed differences between subgroups and controls by performing linear regression on each PRS for each subgroup compared to controls, adjusting for cohort-specific genetic PCs. In addition to setting controls as the reference group, we also set different subgroups as the reference group to compare differences in BE-BROAD PRS and AN PRS across different subgroups.

Last, we generated AN PRS from the main meta-analysis – including solely individuals of European genetic ancestry – and applied this to two East Asian genetic ancestry cohorts (N_{case}=77, $N_{control}$=117 in a Japanese cohort, N_{case}=75, $N_{control}$=109 in a Korean cohort). Korean data were merged with Japanese data from GCAN cohorts. The same pre-imputation quality control and imputation method used for the European ancestry cohorts was conducted. We used the 1000 Genomes Project Phase 3 EUR reference as the LD reference panel for PRS, as the reference panel should align with the ancestry in the base GWAS [94]. We used the European AN prevalence estimate for liability scale conversion, as estimates of AN prevalence in East Asia are sparse and what estimates exist are approximately consistent with estimates in countries with primarily European genetic ancestries [100].

Sensitivity analyses using down-sampled BE-BROAD data

A considerable portion (13%) of the cases for the BE-BROAD GWAS includes individuals who were recruited through studies that focused on AN ascertainment. Even though the BE-BROAD phenotype is expected to be heterogeneous given its transdiagnostic nature, we sought to understand the influence of AN on this BE-BROAD phenotype. We therefore conducted an additional BE-BROAD meta-analysis where we excluded individuals with BE-BROAD who were identified in cohorts that were specifically ascertained for AN (e.g., the ANGI cohorts). This resulted in an additional analysis including 87% of the BE-BROAD GWAS, consisting of the following cohorts: *sebe, agds, alsp, biov, esbb, fngn, jans, moba, ukb2, ukd2*. We then calculated SNP-r_g to compare the genetic relationship of the "not-ascertained-for-AN" BE-BROAD GWAS and the original BE-BROAD GWAS with selected traits significantly correlated with the original BE-BROAD GWAS or with the AN GWAS (hypothesising that AN-related effects may mask or drive SNP-r_gs between such traits and the original BE-BROAD GWAS).

URLS

TopMED Hg37 VCF: https://ftp.ensembl.org/pub/grch37/release-113/data_files/homo_sapiens/GRCh37/variation_genotype/TOPMED_GRCh37.vcf.gz

TopMED Hg38 VCF: https://ftp.ensembl.org/pub/grch37/release-113/data_files/homo_sapiens/GRCh38/variation_genotype/TOPMED_GRCh38.vcf.gz

LiftOver: https://ftp.ensembl.org/pub/grch37/release-113/data_files/homo_sapiens/GRCh38/variation_genotype/TOPMED_GRCh38.vcf.gz

MiXeR: https://github.com/precimed/mixer

SPrediXcan best practice: https://github.com/hakyimlab/MetaXcan/wiki/Best-practices-for-integrating-GWAS-and-GTEX-v8-transcriptome-prediction-models

GWAS tools:

https://github.com/hakyimlab/summary-gwas-imputation/wiki/GWAS-Harmonization-And-Imputation