# Genome-wide association study of borderline personality disorder identifies 11 loci and highlights shared risk with mental and somatic disorders

Fabian Streit 1, 2, 3, 4*, Swapnil Awasthi 5, 6*, Alisha SM Hall 7, 8, Alice Braun 5, 6, Maria Niarchou 9, Eirini Marouli 10, Oladapo Babajide 10, Josef Frank 3, Lea Zillich 3, 4, 11, 12, Carolin M Callies 3, 13, Diana Avetyan 3, Eric Zillich 3, Joonas Naamanka 1, 2, 14, Jean Gonzalez 15, Arvid Harder 16, Yi Lu 16, Zouhair Aherrahrou 17, 18, 19, Zain-Ul-Abideen Ahmad 20, Helga Ask 21, 22, 23, Anthony Batzler 24, Michael E Benros 25, 26, Odette M Brand-de Wilde 27, Søren Brunak 28, Mie T Bruun 29, Lea AN Christoffersen 30, 31, Lucía Colodro-Conde 32, 33, Brandon J Coombes 24, Elizabeth C Corfield 21, 34, Norbert Dahmen 35, Maria Didriksen 36, 37, Khoa M Dinh 37, 38, Srdjan Djurovic 39, 40, Joseph Dowsett 37, Ole Kristian Drange 41, 42, Helene Dukal 3, Susanne Edelmann 43, 44, Christian Erikstrup 7, 38, Mariana K Espinola 3, Eva Fassbinder 45, Annika Faucon 9, Diana S Ferreira de Sá 46, Jerome C Foo 3, 47, 48, 49, Maria Gilles 2, Alfonso Gutiérrez-Zotes 50, 51, Thomas F Hansen 52, 53, Magnus Haraldsson 54, R. Patrick Harper 55, Alexandra Havdahl 21, 23, 34, Urs Heilbronner 56, Stefan Herms 57, 58, Henrik Hjalgrim 26, 59, Christopher Hübel 60, 61, 62, 63, Gitta A Jacob 64, Bitten Aagaard 65, Anders Jorgensen 26, 66, Martin Jungkunz 67, 68, Nikolaus Kleindienst 69, Nora Knoblich 43, Stefanie Koglin 70, Julia Kraft 5, Kristi Krebs 71, Christopher W Lee 72, Yuhao Lin 60, 61, Stefanie Lis 69, 73, Amanda Lisoway 74, 75, Ioannis A Malogiannis 76, Amy Martinsen 77, 78, 79, Tolou Maslahati 70, Katharina Merz 80, Andreas Meyer-Lindenberg 2, Susan Mikkelsen 38, Christina Mikkelsen 37, Arian Mobascher 35, Gerard Muntané 50, 51, 81, Asmundur Oddsson 82, Sisse R Ostrowski 26, 37, Teemu Palviainen 83, Ole BV Pedersen 26, 30, Geir Pedersen 84, 85, Liam Quinn 30, Matthias A Reinhard 80, 86, Florian A Ruths 87, Björn H Schott 88, 89, 90, 91, Michael Schredl 92, Emanuel Schwarz 1, 2, 4, Cornelia E Schwarze 93, Michael Schwinn 37, Tabea Send 2, Engilbert Sigurdsson 54, 94, Katja Simon-Keller 3, Astros T Skuladottir 82, 95, Joaquim Soler 96, 97, 98, Anne Sonley 99, 100, Erik Sørensen 37, Hreinn Stefansson 82, Peter Straub 9, Jaana Suvisaari 101, Martin Tesli 102, 103, Jacob Træholt 37, Henrik Ullum 104, Maja P Völker 3, G Bragi Walters 82, Rujia Wang 60, 61, Christian C Witt 105, 105, Gerhard Zarbock 106, Peter Zill 80, John-Anker Zwart 77, 78, 79, DBDS Genomic Consortium 107, Estonian Biobank Research Team 108, the GLAD Study 109, HUNT All-In Psychiatry 110, Ole A Andreassen 111, 112, 113, Arnoud Arntz 114, Joanna M Biernacka 115, Martin Bohus 69, Gerome Breen 60, 61, Alexander L Chapman 116, 117, Sven Cichon 118, 119, 120, Lea K Davis 9, 121, Michael Deuschle 2, 4, Sebastian Euler 122, Sabine C Herpertz 4, 123, Benjamin Hummelen 84, Andrea Jobst 80, Jaakko Kaprio 83, James L Kennedy 74, 75, 100, Kelli Lehto 71, Klaus Lieb 35, Lourdes Martorell 50, 51, Shelley McMain 100, 124, Richard Musil 80, 125, Vanessa Nieratschker 43, 44, Markus M Nöthen 126, Frank Padberg 80, 86, Aarno Palotie 83, 127, 128, Juan C Pascual 96, 97, 129, Nader Perroud 130, Josep A Ramos-Quiroga 51, 131, 132, 133, Ted Reichborn-Kjennerud 21, 134, Marta Ribases 51, 131, 132, 135, Stefan Roepke 70, 136, Dan Rujescu 137, Sandra Sanchez-Roige 9, 138, 139, Claudia Schilling 92,

Christian Schmahl 4, 69, Kari Stefansson 82, 95, Thorgeir E Thorgeirsson 82, Gustavo Turecki 140, Elisabet Vilella 50, 51, Thomas Werge 26, 31, Bendik S Winsvold 77, 79, 141, Johannes Wrege 142, Marcella Rietschel 3*, Stephan Ripke 5, 6, 143*, Stephanie H Witt 3, 4*

*the authors contributed equally

1, Hector Institute for Artificial Intelligence in Psychiatry, Central Institute of Mental Health, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany
2, Department of Psychiatry and Psychotherapy, Central Institute of Mental Health, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany
3, Department of Genetic Epidemiology in Psychiatry, Central Institute of Mental Health, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany
4, German Center for Mental Health (DZPG), Partner Site Mannheim - Heidelberg - Ulm, Germany
5, Department of Psychiatry and Psychotherapy, Charité Universitätsmedizin Berlin, Berlin, Germany
6, Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge, MA, USA
7, Department of Clinical Medicine, Aarhus University, Aarhus, Denmark
8, Department of Affective Disorders, Aarhus University Hospital - Psychiatry, Aarhus, Denmark
9, Department of Medicine, Division of Genetic Medicine, Vanderbilt University Medical Center, Nashville, TN, USA
10, William Harvey Research Institute, Faculty of Medicine and Dentistry, Queen Mary University London, London, UK
11, Hector Institute for Translational Brain Research, Central Institute of Mental Health, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany
12, Department of Psychiatry and Psychotherapy, Medical Center - University of Freiburg, Faculty of Medicine, University of Freiburg, Freiburg, Germany
13, Health Psychology, School of Social Sciences, University of Mannheim, Mannheim, Germany
14, SleepWell Research Program, Faculty of Medicine, University of Helsinki, Helsinki, Finland
15, Department of Psychiatry, University of California San Diego, La Jolla, USA
16, Department of Medical Epidemiology and Biostatistics, Karolinska Institute, Stockholm, Sweden
17, Institute for Cardiogenetics, University of Lübeck, Lübeck, Germany
18, DZHK (German Research Centre for Cardiovascular Research), partner site Hamburg/Lübeck/Kiel, Lübeck, Germany
19, University Heart Center Lübeck, Lübeck, Germany
20, Social, Genetic and Developmental Psychiatry Centre, Institute of Psychology, Institute of Psychology, Psychiatry and Neuroscience, King's College London, London, UK
21, PsychGen Centre for Genetic Epidemiology and Mental Health, Norwegian Institute of Public Health, Oslo, Norway
22, Department of Child Health and Development, Norwegian Institute of Public Health, Oslo, Norway
23, PROMENTA Research Center, Department of Psychology, University of Oslo, Oslo, Norway

24, Department of Quantitative Health Sciences, Division of Computational Biology, Mayo Clinic, Rochester, MN, USA
25, Copenhagen Research Centre for Biological and Precision Psychiatry, Mental Health Centre Copenhagen, Copenhagen University Hospital, Copenhagen, Denmark
26, Department of Clinical Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark
27, Netherlands Institute for Personality Disorders, PO Box 15933, Halsteren, Netherlands
28, Department of Public Health & Novo Nordisk Foundation Center for Protein Research, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark
29, Clinical Immunology Research Unit, Department of Clinical Immunology, Odense University Hospital, Odense, Denmark
30, Department of Clinical Immunology, Zealand University Hospital, Køge, Denmark
31, Institute of Biological Psychiatry, Mental Health Services, University of Copenhagen, Copenhagen, Denmark
32, Brain and Mental Health Research Program, QIMR Berghofer, Brisbane, QLD, Australia
33, School of Psychology, The University of Queensland, Brisbane, QLD, Australia
34, Nic Waals Institute, Lovisenberg Diaconal Hospital, Oslo, Norway
35, Department of Psychiatry and Psychotherapy, University Medical Center, University of Mainz, Mainz, Germany
36, Department of Neuroscience, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark
37, Department of Clinical Immunology, Copenhagen University Hospital, Rigshospitalet, Copenhagen, Denmark
38, Department of Clinical Immunology, Aarhus University Hospital, Aarhus, Denmark
39, Department of Medical Genetics, Oslo University Hospital, Oslo, Norway
40, Department of Clinical Science, University of Bergen, Bergen, Norway
41, NORMENT Centre, Institute of Clinical Medicine, University of Oslo, Oslo, Oslo, Norway
42, Department of psychiatry, Sørlandet hospital, Kristiansand, Agder, Norway
43, Department of Psychiatry and Psychotherapy, University Hospital Tübingen, Tübingen, Germany
44, German Center for Mental Health (DZPG), partner site Tübingen, Tübingen, Germany
45, Department of Psychiatry and Psychotherapy, Christian-Albrechts-Universität zu Kiel, Kiel, Germany
46, Department of Psychology, Division of Clinical Psychology and Psychotherapy, Saarland University, Saarbrücken, Germany
47, Institute for Psychopharmacology, Central Institute of Mental Health, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany
48, Department of Psychiatry, College of Health Sciences, University of Alberta, Edmonton, AB, Canada
49, Neuroscience and Mental Health Institute, University of Alberta, Edmonton, AB, Canada
50, Hospital Universitari Institut Pere Mata (HUIPM), Institut d'Investigació Sanitària Pere Virgili (IISPV-CERCA), Universitat Rovira i Virgili (URV), Reus, Spain
51, Biomedical Network Research Centre on Mental Health (CIBERSAM), Instituto de Salud Carlos III, Madrid, Spain
52, Neurogenomic, Translational Research Centre, Copenhagen University Hospital, Glostrup, Denmark
53, Danish Headache Center, Copenhagen University Hospital, Rigshospitalet, Copenhagen, Denmark

54, Landspitali University Hospital, Reykjavik, Iceland
55, Bradford District Care NHS Foundation Trust, Bradford, UK
56, Institute of Psychiatric Phenomics and Genomics (IPPG), LMU University Hospital, LMU Munich, Munich, Germany
57, Human Genomics Research Group, Department of Biomedicine, University of Basel, Basel, Switzerland
58, Institute of Human Genetics, University Hospital Bonn & University of Bonn, Bonn, Germany
59, Danish Cancer Institute, Denmark
60, Social, Genetic and Developmental Psychiatry Centre, Institute of Psychology, Psychiatry and Neuroscience, King's College London, London, UK
61, UK National Institute for Health Research (NIHR) Biomedical Research Centre, South London and Maudsley Hospital and King's College London, London, UK
62, National Centre for Register-based Research, Aarhus University, Aarhus, Denmark
63, Clinic for Child and Adolescent Psychiatry, German Red Cross Hospitals Westend, Berlin, Germany
64, Institute for Psychology, Department of Clinical Psychology and Psychotherapy, University of Freiburg, University of Freiburg, Germany
65, Department of Clinical Immunology, Aalborg University Hospital, Aalborg, Denmark
66, Psychiatric Center Copenhagen, Denmark
67, Institute for Medical and Data Ethics, Heidelberg University, Faculty of Medicine, Heidelberg, Germany
68, German Cancer Research Center (DKFZ), Heidelberg, Germany
69, Department of Psychosomatic Medicine and Psychotherapy, Central Institute of Mental Health, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany
70, Department of Psychiatry and Neurosciences, Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt- Universität zu Berlin, Campus Benjamin Franklin, Berlin, Germany
71, Estonian Genome Centre, Institute of Genomics, University of Tartu, Tartu, Estonia
72, Faculty of Health and Medical Sciences, University of Western Australia, Perth, Western Australia, Australia
73, Department of Clinical Psychology, Central Institute of Mental Health, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany
74, Molecular Brain Science, Campbell Family Mental Health Research Institute, Centre for Addiction and Mental Health, Toronto, Ontario, Canada
75, Institute of Medical Science, University of Toronto, Toronto, Canada
76, First Department of Psychiatry, Eginition Hospital Medical School, National and Kapodistrian University of Athens, Athens, Greece
77, Department of Research and Innovation, Division of Clinical Neuroscience, Oslo University Hospital, Oslo, Norway
78, Institute of Clinical Medicine, Faculty of Medicine, University of Oslo, Oslo, Norway
79, HUNT Center for Molecular and Clinical Epidemiology, Department of Public Health and Nursing, Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology, Trondheim, Norway
80, Department of Psychiatry and Psychotherapy, LMU University Hospital Munich, Ludwig Maximilian University Munich, Munich, Germany
81, Institut de Biologia Evolutiva (UPF-CSIC), Department of Medicine and Life Sciences, Universitat Pompeu Fabra, Parc de Recerca Biomèdica de Barcelona, Barcelona, Spain

82, deCODE genetics / AMGEN, Reykjavik, Iceland
83, Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland
84, Department of Research and Innovation, Division of Mental Health and Addiction, Oslo University Hospital, Oslo, Norway
85, Institute of Basic Medical Sciences, Faculty of Medicine, University of Oslo, Oslo, Norway
86, German Center for Mental Health (DZPG), Partner Site Munich-Augsburg, Germany
87, South London and Maudsley NHS Foundation Trust, London, UK
88, Department of Psychiatry and Psychotherapy, University Medical Center Göttingen, Göttingen, Germany
89, German Center for Neurodegenerative Diseases (DZNE), Göttingen, Germany
90, Leibniz Institute for Neurobiology, Magdeburg, Germany
91, Department of Psychiatry and Psychotherapy, Otto von Guericke University, Magdeburg, Germany
92, Department of Psychiatry and Psychotherapy, Sleep Laboratory, Central Institute of Mental Health, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany
93, Department of Clinical Psychology and Psychotherapy, Charlotte Fresenius University, Wiesbaden, Germany
94, Faculty of Medicine, Department of Psychiatry, School of Health Sciences, University of Iceland, Reykjavik, Iceland
95, Faculty of Medicine, University of Iceland, Reykjavik, Iceland
96, Centro de Investigación Biomédica en Red de Salud Mental (CIBERSAM), Institut de Recerca Biomèdica Sant Pau (IIB-Sant Pau), Barcelona, Spain
97, Department of Psychiatry, Hospital de la Santa Creu i Sant Pau, Barcelona, Spain
98, Department of Psychiatry and Forensic Medicine & Institute of Neurosciences, Universitat Autònoma de Barcelona, Bellaterra, Spain
99, Borderline Personality Disorder Clinic, Centre for Addiction and Mental Health, Toronto, Ontario, Canada
100, Department of Psychiatry, University of Toronto, Toronto, Canada
101, Finnish Institute for Health and Welfare (THL), Helsinki, Finland
102, Mental health and suicide, Norwegian Institute of Public Health, Oslo, Norway
103, Division of Mental Health and Addiction, Oslo University Hospital, Oslo, Norway
104, Statens Serum Institut, Copenhagen, Denmark
105, Department of Anaesthesiology and Operative Intensive Care, University Hospital Mannheim, Medical Faculty Mannheim/Heidelberg University, Mannheim, Germany
106, IVAH, Institut für Verhaltenstherapie-Ausbildung Hamburg, gemeinnützige GmbH, Hamburg, Germany
107, Members of the DBDS Genomic Consortium are listed in the Supplementary Methods
108, Members of the Estonian Biobank Research Team are listed in the Supplementary Methods, Estonian Genome Centre, Institute of Genomics, University of Tartu, Tartu, Estonia
109, Members of the the GLAD Study are listed in the Supplementary Methods
110, Members of HUNT All-In Psychiatry are listed in the Supplementary Methods
111, NORMENT Centre, Division of Mental Health and Addiction, Oslo University Hospital, Oslo, Norway
112, KG Jebsen Centre for Neurodevelopmental Disorders, University of Oslo, Oslo, Norway
113, NORMENT Centre, Institute of Clinical Medicine, University of Oslo, Oslo, Norway
114, Department of Clinical Psychology, University of Amsterdam, Amsterdam, Netherlands

115, Department of Quantitative Health Sciences, Department of Psychiatry & Psychology, Division of Computational Biology, Mayo Clinic, Rochester, MN, USA
116, Department of Psychology, Faculty of Arts and Social Sciences, Simon Fraser University, Vancouver, British Columbia, Canada
117, DBT Centre of Vancouver, Vancouver, British Columbia, Canada
118, Department of Biomedicine, University of Basel, Basel, Switzerland
119, Medical Genetics, Institute of Medical Genetics and Pathology, University Hospital Basel, Basel, Switzerland
120, Institute of Neuroscience and Medicine (INM-1), Research Center Juelich, Juelich, Germany
121, Vanderbilt Genetics Institute, Vanderbilt University Medical Center, Nashville, TN, USA
122, Department of Consultation Psychiatry and Psychosomatics, University Hospital Zurich, Zurich, Switzerland
123, Department of General Psychiatry, University Hospital Heidelberg, Heidelberg University, Heidelberg, Germany
124, Borderline Personality Disorder Clinic, General Adult Psychiatry and Health Systems Division, Centre for Addiction and Mental Health, Toronto, Ontario, Canada
125, Oberberg Fachkliniken for Psychiatry, Psychosomatics and Psychotherapy, Bad Tölz, Germany
126, Insitute of Human Genetics, University Hospital Bonn & University of Bonn, Bonn, Germany
127, Analytic and Translational Genetics Unit, Department of Medicine, Department of Neurology and Department of Psychiatry, Massachusetts General Hospital, Boston, MA, USA
128, The Stanley Center for Psychiatric Research and Program in Medical and Population Genetics, The Broad Institute of MIT and Harvard, Cambridge, MA, USA
129, Mental Health and Psychiatry Department, Vic Hospital Consortium, Spain
130, Department of Psychiatric, University Hospitals of Geneva, Geneva, Switzerland
131, Psychiatric Genetics Unit, Group of Psychiatry, Mental Health and Addiction, Vall d'Hebron Research Institute (VHIR), Universitat Autònoma de Barcelona, Barcelona, Spain
132, Department of Psychiatry, Hospital Universitari Vall d'Hebron, Barcelona, Spain
133, Department of Psychiatry and Forensic Medicine, Universitat Autònoma de Barcelona (UAB), Barcelona, Spain
134, Institute of Clinical Medicine, University of Oslo, Oslo, Norway
135, Department of Genetics, Microbiology, and Statistics, Faculty of Biology, Universitat de Barcelona, Barcelona, Spain
136, Department of Psychiatry, Oberberg Fachkliniken for Psychiatry, Psychosomatics and Psychotherapy, Wendisch Rietz, Germany
137, Department of Psychiatry, University of Halle, Halle, Germany
138, Institute for Genomic Medicine, University of California San Diego, La Jolla, CA, USA
139, Department of Psychiatry, University of California San Diego, La Jolla, CA, USA
140, Douglas Institute, Department of Psychiatry, McGill University, Montreal, Canada
141, Department of Neurology, Oslo University Hospital, Oslo, Norway
142, Medical Faculty, University of Basel, Basel, Switzerland
143, Analytic and Translational Genetics Unit, Massachusetts General Hospital, Boston, MA, USA

# Abstract

We conducted the largest genome-wide meta-analysis of borderline personality disorder (BPD) to date, with a discovery sample of 12,339 cases and 1,041,717 controls, and a replication study of 685 cases and 107,750 controls (all participants of European ancestry). We identified 11 independent associated genomic loci, and nine risk genes in the gene-based analysis. We observed a single-nucleotide polymorphism (SNP) heritability of 17.3% and derived polygenic scores (PGS) predicted 4.6% of the phenotypic variance in BPD on the liability scale. BPD showed the strongest positive genetic correlations with GWAS of posttraumatic stress disorder, depression, attention deficit hyperactivity disorder, antisocial behavior, and measures of suicide and self-harm. Phenome-wide association analyses using BPD-PGS confirmed these associations and additionally revealed associations with general medical conditions including obstructive pulmonary disease and diabetes. The present analyses highlight BPD as a polygenic disorder, with the genetic risk showing substantial overlap with psychiatric and physical health conditions.

# Main text

Borderline Personality Disorder (BPD) is a mental disorder characterized by pervasive instability in emotions, interpersonal relationships, and self-image, and by impulsive behavior (symptoms see Supplementary Methods)[1,2]. BPD has a prevalence of 0.92–1.90% in Western countries[3], with symptom onset typically occurring during adolescence. Women are more frequently diagnosed with BPD than men by a ratio of ~3:1, for which a substantial contribution of diagnostic as well as selection bias has been postulated[4,5]. Individuals with BPD display high rates of self-harm, suicidal ideation, and suicide attempts. BPD shows substantial symptom overlap and comorbidity with other mental disorders[5,6], and comorbidity with neurological and somatic health conditions[6,7]. While some psychotherapies are effective in treating BPD[8], no psychopharmacological treatments are FDA□approved specifically for BPD[2,9].

In addition to environmental risk factors such as early interpersonal trauma[10,11], genetic factors substantially contribute to disorder risk. Twin and family studies estimate the heritability of BPD to be 46–69%[12,13], and demonstrate that the genetic risk for BPD is partially shared with other mental disorders, but as well as with continuous traits, e.g. the Big Five personality traits[14,15]. However, a systematic assessment of shared genetic risk with a broad range of disorders and traits is missing.

For many mental disorders, genome-wide association studies (GWAS) meta-analyzing genetic data from tens or hundreds of thousands of cases and controls have successfully identified up to hundreds of genetic risk loci[16]. In contrast, genetic research on BPD lags behind[17]: In the thus far only GWAS of BPD conducted in 998 cases and 1,545 controls[18], no single genome-wide significant variants were identified, but significant genetic correlations ($r_g$) of BPD were observed with bipolar disorder (BIP), schizophrenia (SCZ), and major depressive disorder[18]. While these findings indicate the potential of using genetic approaches to investigate BPD, research based on those results is limited by the large uncertainties in the estimated effect sizes. Additionally, the extent to which sex-specific genetic effects contribute to the observed sex differences in the prevalence and clinical characteristics of BPD remains unclear.

The main aims of the present study were to a) identify novel genetic risk loci for BPD to improve the understanding of the underlying molecular mechanisms, and b) systematically assess the shared genetic risk between BPD and a broad range of related traits and disorders.

# Results

## Genome-wide association study

We performed a discovery GWAS meta-analysis including 6,043,895 genetic markers in 12,339 BPD cases and 1,041,717 controls of European ancestry (overview Fig. 1). In total, 17 studies contributed individual-level data for a total of 2,705 cases meeting DSM-IV criteria for BPD and 4,600 controls, which were combined into five datasets. Additionally, 10 large-scale biobank or cohort studies provided summary statistics from GWAS of BPD ($N_{cases}$=9,634, $N_{controls}$=1,037,117 controls). In these, BPD status was assessed using ICD codes except for the Genetic Links to Anxiety and Depression (GLAD) study, which used a self-reported diagnosis. Details on ascertainment, and inclusion and exclusion criteria are documented in Supplementary Table S1. Sample sizes, basic demographics (sex and age), depression comorbidity, and analysis details are provided in Supplementary Table S2 and the Supplementary Methods. Additionally, we performed meta-analyses stratified by sex (female-only; male-only) as well as a sensitivity analysis excluding subjects with a history of BIP or SCZ—conditions commonly excluded in BPD studies—from the studies providing summary statistics (Supplementary Tables S3–S5). To assess replication, the lead SNPs and PGS derived from the discovery analysis were tested in two independent datasets (total $N_{cases}$=685, total $N_{controls}$=107,750). Additionally, SNPs with $p<1x10^{-6}$ in the discovery GWAS were analyzed in a combined meta-analysis ($N_{cases}$=13,024, $N_{controls}$=1,149,467).

The discovery GWAS had power over 80% to detect effects of 1.1 for variants with an allele frequency between 0.3–0.7[19] (Supplementary Table S6, Supplementary Figure S1).

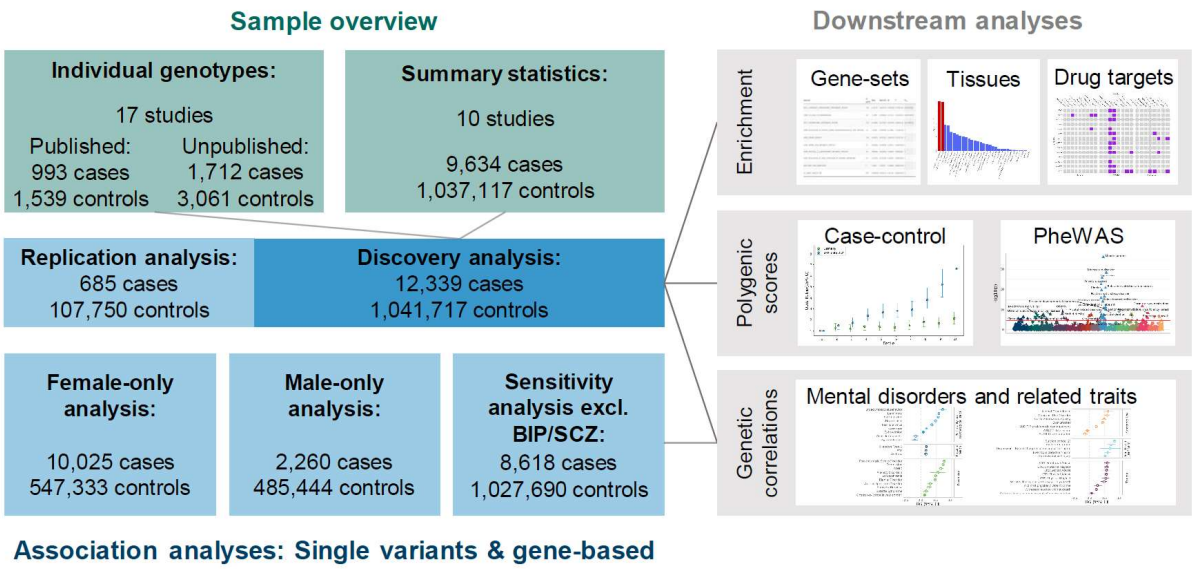



**Figure 1: Schematic overview of the genome-wide association study (GWAS) of borderline personality disorder (BPD).** A total of 17 studies providing individual-level data (analyzed in 5 datasets: see text) and 10 studies providing summary statistics were included in the discovery meta-analysis. Genetic associations were tested at the single-variant and gene level. Gene enrichment

analyses were applied to test for enrichment in gene-sets and tissue-specific gene expression data from 53 human tissues (GTEx), from 29 different ages and 11 general developmental stages (BrainSpan), single-cell based cell types, and drug-target gene set. Polygenic scores were calculated to assess the prediction of BPD case-control status in the analysis datasets with available individual-level data and to test the association of the genetic liability for BPD in phenome-wide association studies (PheWAS) with phecodes in two biobanks (BioVU, UK Biobank). Genetic correlations were calculated between the results of the GWAS of BPD and the GWAS of 50 disorders and traits of interest. BIP=bipolar disorder; SCZ=schizophrenia.

We observed an SNP-heritability of 28.4% (95% confidence interval (CI) [24.9%, 31.8%]), corresponding to an SNP-heritability of 17.2% (95%-CI [15.2%, 19.4%]) on the liability scale (assuming a population prevalence of 1.5%)[3]. The linkage disequilibrium score regression (LDSC)[20] intercept was 1.03 (standard error (SE)=0.0085), with an attenuation ratio of 0.11 (SE=0.031), and we observed a genomic inflation factor ($\lambda_{GC}$) of 1.22 ($\lambda_{1000}$=1.01). There was a high genetic correlation between the studies with individual-level data and those providing summary statistics ($r_g$=0.85, 95%-CI [0.68, 1.03]), which was significantly smaller than 1 ($p$=0.038).

SNP association analysis revealed six independent genome-wide significant loci ($p<5\times10^{-8}$, Fig. 2; Table 1; Supplementary Figs. S2–S14; Supplementary Table S7), and no marker showed significant heterogeneity between studies after multiple testing correction (smallest heterogeneity p-value=0.023). All six genome-wide significant lead SNPs (6/6, $p$=0.016, Tables 1), and 77% of the lead SNPs associated with $p<1\text{x}10^{-6}$ (24/31, $p$=0.0017) showed effects in the same direction in the replication analysis. In the combined analysis, the six loci remained genome-wide significant, and five additional loci with $p<5\text{x}10^{-8}$ were observed (Figure1, Table 1, Supplementary Figs. S2–S24).

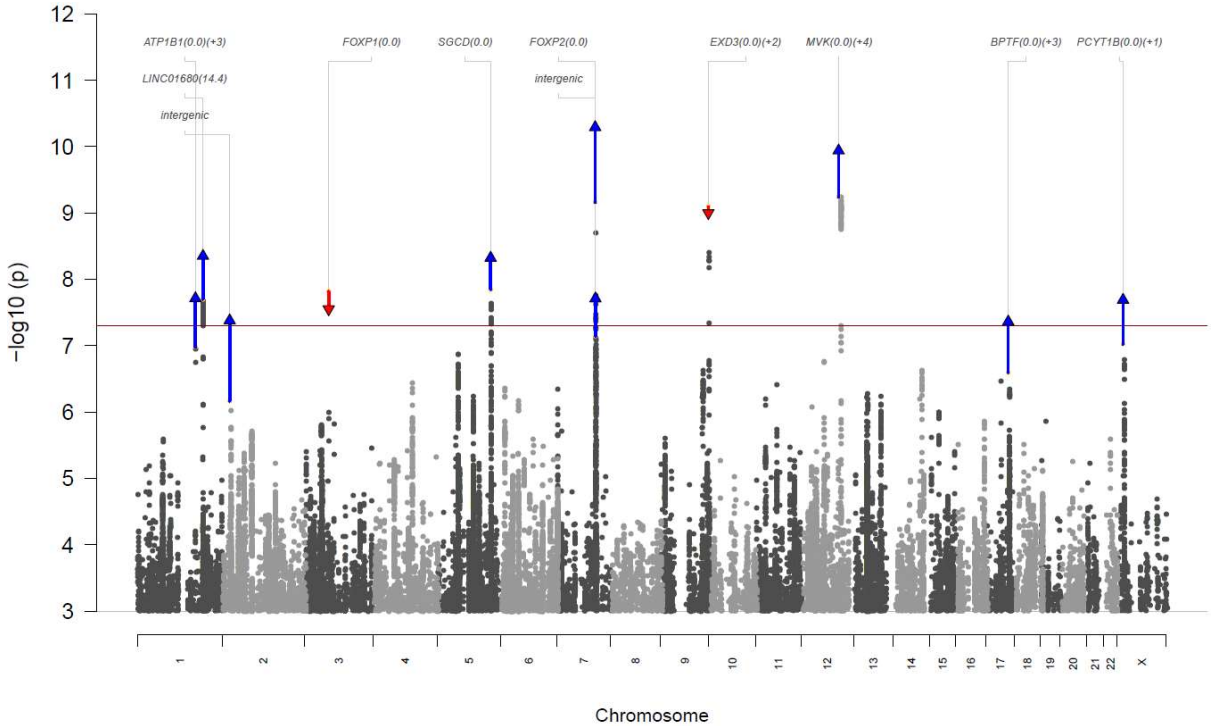



**Figure 2: Manhattan plot of the genome-wide association study (GWAS) of borderline personality disorder (BPD).** The -log10 *p*-value for each single nucleotide polymorphism (SNP) in the discovery GWAS ($N_{cases}$=12,339, $N_{controls}$=1,041,717) is indicated on the y-axis (chromosomal position shown on the x-axis). The red line horizontal indicates genome-wide significance ($p<5\times10^{-8}$). Index SNPs representing independent genome-wide significant associations either in the discovery GWAS or the combined analysis ($N_{cases}$=13,024, $N_{controls}$=1,149,467) are highlighted (upward-pointing blue triangle: increased significance in combined analysis, downward-pointing red triangle: reduced significance in combined analysis). Combined p-values are plotted for lead loci and vertical lines indicate the p-value change from the discovery to the combined analysis.

**Table 1: Lead 11 genome-wide significant SNPs associated ($p<5\times10^{-8}$) with BPD**

| Locus | CHR | SNP | BP | Locus (start..stop) | A1/ A2 | FRQ case | FRQ control | OR | SE | *p* | OR repl. | *p* repl. | OR comb. | *SE comb.* | *p* comb. | Mapped Genes |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| comb. 1 | 1 | rs2143288 | 169084340 | 169082364..169084344 | G/A | 0.858 | 0.868 | 0.89 | 0.022 | $1.06\times10^{-7}$ | 0.86 | 0.035 | 0.89 | 0.021 | **$1.23\times10^{-8}$** | *LINC00970(0.0),LOC101928596(0.0), ATP1B1(0.0),NME7(0.0)* |
| 1 | 1 | rs9970854 | 191125935 | 191005935..191148335 | G/A | 0.595 | 0.597 | 0.92 | 0.015 | **$1.99\times10^{-8}$** | 0.93 | 0.183 | 0.92 | 0.014 | **$7.42\times10^{-9}$** | *LINC01680(14.4)* |
| comb. 2 | 2 | rs12466671 | 22583435 | 22548735..22606235 | A/G | 0.587 | 0.599 | 0.93 | 0.015 | $6.80\times10^{-7}$ | 0.85 | 0.002 | 0.92 | 0.014 | **$2.16\times10^{-8}$** | - |
| 2 | 3 | rs6549383 | 71266885 | 71261185..71267720 | T/C | 0.503 | 0.503 | 1.09 | 0.015 | **$1.53\times10^{-8}$** | 1.02 | 0.664 | 1.08 | 0.014 | **$2.45\times10^{-8}$** | *FOXP1(0.0)* |
| 3 | 5 | rs2135029 | 155856538 | 155743538..155868438 | A/G | 0.594 | 0.599 | 0.92 | 0.014 | **$1.41\times10^{-8}$** | 0.92 | 0.141 | 0.92 | 0.014 | **$4.30\times10^{-9}$** | ***SGCD(0.0)*** |
| 4 | 7 | rs4727799 | 114110568 | 114021868..114224568 | T/C | 0.663 | 0.67 | 1.10 | 0.015 | **$6.92\times10^{-10}$** | 1.15 | 0.015 | 1.10 | 0.015 | **$4.05\times10^{-11}$** | ***FOXP2(0.0)*** |
| comb. 3 | 7 | rs10953781 | 114974695 | 114963595..115112695 | C/T | 0.528 | 0.534 | 1.08 | 0.014 | $7.13\times10^{-8}$ | 1.06 | 0.313 | 1.08 | 0.014 | **$4.38\times10^{-8}$** | - |
| 5 | 9 | rs73581580 | 140251458 | 140242728..140278158 | G/A | 0.836 | 0.839 | 0.89 | 0.019 | **$8.09\times10^{-10}$** | 0.98 | 0.798 | 0.89 | 0.019 | **$1.45\times10^{-9}$** | *NRARP(-4.8)**, EXD3(0.0),** NOXA1(16.4)* |
| 6 | 12 | rs7304862 | 109988891 | 114021868..114224568 | G/T | 0.529 | 0.513 | 1.09 | 0.014 | **$5.71\times10^{-10}$** | 1.09 | 0.114 | 1.09 | 0.014 | **$1.68\times10^{-10}$** | *MYO1H(-52.7), KCTD10 (-23.5), UBE3B(0.0),* ***MMAB(0.0)**, **MVK(0.0)*** |
| comb. 4 | 17 | rs7210027 | 65977053 | 65825053..66098053 | T/C | 0.77 | 0.782 | 0.92 | 0.017 | $2.54\times10^{-7}$ | 0.87 | 0.031 | 0.91 | 0.016 | **$3.12\times10^{-8}$** | *BPTF(0.0),C17orf58(0.0),KPNA2(4.7) ,LINC00674(70.6)* |
| comb. 5 | X | rs5944622 | 24613189 | 24578589..24933189 | T/C | 0.57 | 0.58 | 0.93 | 0.013 | $9.47\times10^{-8}$ | 0.92 | 0.037 | 0.93 | 0.013 | **$2.05\times10^{-8}$** | *PDK3 (0), PCYT1B (0), PCYT1B−AS1 (+5.0), POLA1 (+48.8)* |

Variants are sorted by chromosomal position. CHR=chromosome; Locus (start..stop)=start and end of the boundaries of the region within $r^2$ of 0.6 of the index SNP; A1=allele 1, A2=allele 2, FRQ=frequency of A1 in cases/controls; BP=base pair position; comb. = combined analysis; OR=odds ratio of A1; repl. = replication analysis; SE=standard error; SNP=single nucleotide polymorphism; BPD=borderline personality disorder; Mapped genes=genes within $r^2$ of 0.6 and 50 kb of the index SNP (distance to the 50kb for boundary of the lead SNP in kb is indicated in brackets); Genes in bold were also identified by the gene-based association tests. *P* values in bold indicate genome-wide significance ($p<5x10^{-8}$). OR, SE, and p are reported for the main analysis ($N_{cases}$=12,339, $N_{controls}$=1,041,717), the replication analysis (repl.) $N_{cases}$=685, total $N_{controls}$=107,750), and the combined (comb.) meta-analysis ($N_{cases}$=13,024, $N_{controls}$=1,149,467) of the two datasets.

Gene-based associations using MAGMA (version 1.07)[21,22] indicated nine genes, including *CCDC71* and *DEPDC1B*; in addition to *SGCD, FOXP2, EXD3, MVK, MMAB, PCYT1B*, and *BPTF* already indicated by the genome-wide SNP associations (Extended Data Fig. 1; Supplementary Fig. S25; Supplementary Table S8).

Among the genes implicated by proximity to lead SNPs or by the gene-based analysis, the highest Polygenic Priority Scores (PoPS)[23] were observed for *FOXP2* (score=0.96, rank=1), and *SGCD* (score=0.74, rank=2), with *FOXP1* and *DEPDC1B* also ranking in the top 1% (Supplementary Table S9). Of the genes in proximity to loci that reached significance in the combined analysis, *NME7* and *KPNA2* ranked highest (top 5%).

**Sex-stratified GWAS**

We carried out sex-stratified GWAS meta-analyses for male ($N_{cases}$=2,260, $N_{controls}$=485,444) and female subjects ($N_{cases}$=10,025, $N_{controls}$=547,333) (Supplementary Tables S3&S4).

We used sex-specific population prevalences of 2.25% for females and 0.75% for males to convert heritability estimates to the liability scale[4,5]. For females, we observed an SNP-heritability of 30.3% (95%-CI [26.1%, 34.5%]), corresponding to 20.5% (95%-CI [17.7%, 23,4%]) on the liability scale. For males, the observed SNP-heritability was 19,8% (95%-CI [7.9%, 31.6%]), corresponding to 10.2% (95%-CI [4.1%, 16.3%] on the liability scale, which was significantly lower than in females ($p$=0.00090).

There was a high genetic correlation between the two analyses ($r_g$=0.80, 95%-CI [0.53, 1.07]), which did not differ from 1 ($p$=0.055).

The female-only analysis identified one genome-wide significant risk locus on chromosome 9 (rs73581580; $p=2.83\times10^{-8}$; locus 5 identified in the main analysis (*EXD3*), and a second locus on chromosome 7 (rs10227454, $p=4.99\times10^{-8}$), ~1 Mb from locus 4 (*FOXP2*, $R^2$=0.005, $D'$=0.25; Extended Data Fig. 2; Supplementary Figs. S26–S30, Supplementary Table S10). The gene-based analysis identified genome-wide associations for *DEPDC1B*, *SGCD, MVK*, and *MMAB*, all significant in the main analysis (Extended Data Fig. 3; Supplementary Fig. S31).

The male-only analysis identified two genome-wide significant risk loci that were not genome-wide significant in the main analysis: one on chromosome 2 (rs17757829, $p=1.02\times10^{-8}$) and one on chromosome 20 (rs6032676, $p=9.55\times10^{-9}$) (Extended Data Fig. 4; Supplementary Figs. S32–S36, Supplementary Table S11) and no significant genes (Extended Data Fig. 5; Supplementary Fig. S37).

The six lead SNPs of the main GWAS showed comparable effects and at least nominal significance in the smaller, and therefore lower-powered, sex-stratified analyses (Supplementary Table S12; Extended Data Fig. 6). The nine genes significant in the gene-based analysis of the main GWAS all showed nominal significance in the female-only analysis

(all $p_{female}$<8.15×10$^{-5}$), whereas *CCDC71, DEPDC1B,* and *SGCD* were not significant in the male-only analysis (Supplementary Table S13).

**Sensitivity analysis excluding BIP&SCZ**

In the sensitivity analysis excluding subjects with schizophrenia and bipolar disorder ($N_{cases}$=8,618, $N_{controls}$=1,027,690), locus 4 in *FOXP2* (rs4727799) and locus 5 in *EXD3* (rs73581580) from the main analysis were the only genome-wide significant associations (Extended Data Fig. 7; Supplementary Figs. S38–S42, Supplementary Table S14). All six lead SNPs of the main GWAS showed comparable effects (all $p_{sensitivity}$<3.32×10$^{-5}$; Extended Data Fig. 6; Supplementary Table S12). The nine significant genes in the gene-based analysis of the main analysis, all were nominally significant in the sensitivity analysis ($p_{sensitivity}$<0.0021; Supplementary Table S13) with *FOXP2, EXD3, MMAB, MVK* reaching genome-wide significance, and *ZNF626* being the only additional genome-wide significant gene (Extended Data Fig. 8; Supplementary Fig. S43).

# Enrichment analysis

Enrichment of gene-sets and tissue expression using Genotype-Tissue Expression (GTEx; version 8) data for 53 tissue types[24] and expression data from 29 different ages and 11 general developmental stages (BrainSpan)[25] was tested with MAGMA (version 1.07)[21] as implemented in FUMA[22]: One significant gene-set was identified: the S1P-S1P3 (Sphingosine-1-phosphate - sphingosine-1-phosphate receptor 3) pathway[26] ($p_{adj}$=0.018; Supplementary Table S15). Nominally significant enrichment was observed in several GTEx tissue types, with the most significant enrichment in cerebellar tissue (not significant after correction for multiple testing; all $p_{adj}$>0.32; Supplementary Fig. S44, Supplementary Table S16). The analysis using the two BrainSpan data sets indicated the strongest enrichments 8–21 weeks post conception (not significant after correction; all $p_{adj}$>0.09), and in early to mid prenatal phases respectively ($p_{adj}$<0.025; Supplementary Fig. S45&46, Supplementary Table S17&18).

Analyses based on Human Brain Atlas single-nucleus RNA sequencing data[27], showed SNP-h² enrichment using stratified LDSC[28,29] in 2 of the 31 tested superclusters (*"medium spiny neurons"*;*"LAMP5-LHX6 and Chandelier cells"*; $p_{adj}$<0.0064; Supplementary Fig. S47, Supplementary Table S19).

# Drug target analysis

Of the 16 genes priotized from the discovery GWAS meta-analysis via proximity to the 6 GWAS loci or in the gene-based test, none were highlighted as drug targets by Open Targets, but several showed potential tractability (Supplementary Fig. S48). The Genome for

REPositioning drugs (GREP) pipeline revealed no significant enrichment for drug targets across any ICD or ATC category. The Drug Gene Interaction Database (DGIdb) highlighted drug-gene interactions for *MYO1H* with lithium, *MVP* with ALENDRONATE SODIUM, and *PCYT1B* with the unapproved substances CT-2584 and SPHINGOSINE. In an additional analysis with DRUGSETS[30], based on MAGMA, testing the enrichment of the GWAS associations in 735 drug–gene sets, the most significant drugs included medications for neurological conditions (safinamide, gabapentin, lomerizine, oxcarbazepine), pain (prilocaine, tetracaine), and alcohol-dependence (acamprosate), but also medications for metabolic/somatic conditions (Supplementary Table S20; all $p_{adj}$>0.18).

## Polygenic scoring

BPD-PGS were calculated using PRS-CS[31] in the datasets where individual-level genotype information was available using leave-one-out summary statistics, excluding the respective sample from the discovery GWAS. In addition, we calculated PGS in the two independent replication data sets (Online Methods). For comparison, PGS were additonally calculated based on the first BPD GWAS from 2017[18]. The explained variance was converted to the liability scale (population prevalence = 1.5%[3]).

PGS explained a weighted average of 4.6% (AUC=66.0%) of the phenotypic variance on the liability scale (assuming a lifetime prevalence of 1.5%[3]) in the five datasets included in the discovery meta-analysis (Fig. 3, Supplementary Table S21). In the independent replication samples, PGS explained 3.1% of the variance ($p$=0.0041, AUC=62.4%) in Spain 2, and 2.6% in All of Us ($p$=2.09x$10^{-25}$, AUC=61.4%). The odds ratio for BPD case-control status comparing the highest PGS decile to the lowest decile was OR=6.61 (95%-CI [5.20, 8.41]) for PGS based on the current meta-analysis, compared to OR=2.09 (95%-CI [1.60, 2.72]) for PGS based on the 2017 GWAS. When comparing to the middle 10% of the distribution, we observed an OR of 2.37 (95%-CI [1.94, 2.89])) for the highest decile for PGS based on the current meta-analysis and an OR of 1.44 (95%-CI [1.05, 1.98]) for PGS based on the 2017 GWAS.

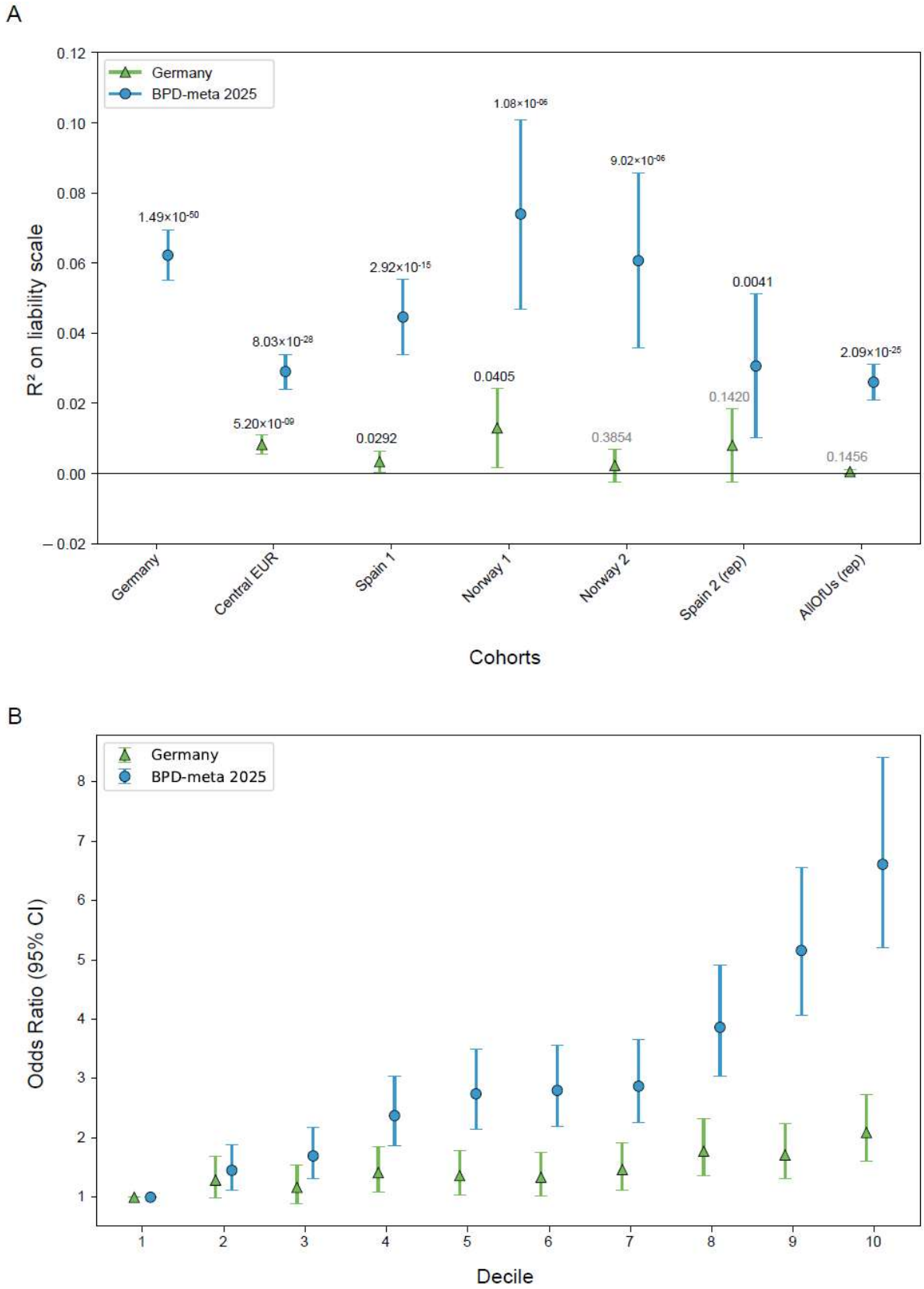



**Figure 3: Polygenic score (PGS) analysis. A) The proportion of variance in case-control status explained by the PGS on the liability scale (y-axis; Liability $R^2$) B) Odds ratio for BPD by PGS deciles, with decile 1 as reference.** Leave-one-out PGS were calculated for the datasets where individual-level genotype data was available, and in an additional independent PGS target sample

(Spain2 (rep.), sample information in Supplementary Tables S1&S2) using PRS-CS. For each prediction, the respective dataset was excluded from the used discovery genome-wide association study (GWAS) meta-analysis. PGS in the independent target dataset were based on the complete current GWAS meta-analysis (blue circles: 2025 meta-GWAS). The explained variance was converted to the liability scale of the population assuming a lifetime disease risk of 1.5%. The analysis presented in Panel B excluded the All of Us dataset, as its individual-level data could not be merged with the other data sets due to data protection regulations. To visualize the increase in variance explained by the PGS, we also calculated PGS based on the first published BPD GWAS, consisting of the Germany sample (green triangles: 2017 GWAS) [18]. Significance: 1*: $p<0.05$; 2*: $p<0.01$; 3*: $p<0.005$; 4*: $p<0.001$; 5*: $p<1\times10^{-4}$; 6*: $p<1\times10^{-8}$; 7*: $p<1\times10^{-12}$, 8*: $p<1\times10^{-20}$, *: $p<1\times10^{-30}$

# Genetic correlations

### Correlations with disorders and traits of interest

In a targeted approach, we calculated genetic correlations with a selection of 50 GWAS of other disorders and traits relevant to BPD, including mental disorders, suicide, self-harm, trauma, substance use, physical health, pain, sleep, personality traits (Big Five GWAS including data from 23andMe, Inc.), and cognition. After Bonferroni correction for multiple testing ($\alpha=0.05/50=0.001$), BPD showed significant genetic correlations with 43 of the 50 tested phenotypes. Among the psychiatric disorders, BPD showed the strongest correlations with post-traumatic stress disorder (PTSD) ($r_g=0.77$), depression ($r_g=0.74$), and attention deficit hyperactivity disorder (ADHD) ($r_g=0.67$). In the other domains, substantial genetic correlations with $|r_g|>0.5$ included material deprivation, chronic pain, broad antisocial behavior, loneliness, externalizing traits, measures of suicide, self-harm, and trauma. Lower but significant genetic correlations ($r_g<0.18$) were observed for the somatic disorders type 2 diabetes and asthma, and body mass index (Fig. 4; Supplementary Table S22).

The sensitivity meta-analysis excluding subjects with SCZ/BIP showed comparable genetic correlations (Supplementary Fig. S49, Supplementary Table S23) with the 50 disorders and traits. Notably, lower but still substantial genetic correlations were observed with BIP ($r_g=0.35$ vs. $r_g=0.47$) and SCZ ($r_g=0.36$ vs. $r_g=0.48$). The genetic correlation of the main meta-analysis and the sensitivity analysis showed a genetic correlation of $r_g=1.00$ (95%-CI [0,98-1,02]).

**Figure 4: Genetic correlations of borderline personality disorder (BPD) with other phenotypes**
Genetic correlations of BPD (total $N_{cases}$=12,339, $N_{controls}$=1,041,717) with 50 disorders and traits: within each group, disorders, and traits are sorted by their genetic correlation. White dot: $p$<0.05; White star: $p$<0.001 (0.05/50 tested correlations). 95% CI: 95% confidence interval; ADHD: attention deficit hyperactivity disorder; AUDIT: Alcohol Use Disorders Identification Test; BMI: body mass index; CTS: Childhood Trauma Screener; PTSD: post-traumatic stress disorder.

## Phenome-wide association studies (PheWAS)

Complementing the targeted genetic correlation analysis, we characterized the genetic signal identified in the BPD GWAS, by performing two phenome-wide association studies (PheWAS). To this end, we tested the association of BPD-PGS (using PRS-CS[31]; excluding each target sample from discovery) with "phecodes", *i.e.*, medical phenotypes based on International Classification of Disease (ICD) diagnoses documented in the electronic health records (EHR), using data from the Vanderbilt University Medical Center Biobank (BioVU)[32] (66,325 subjects; 214 with the phecode 301.20: *"antisocial/borderline personality disorder"*) and Hospital Episode Statistics (HES) in the UK Biobank (UKB) (316,635 subjects; 211 with the phecode 301.20).

**PheWAS in BioVU (EHR data)**

In the PheWAS analysis in BioVU, 69 of the 1,431 tested diagnoses showed an association with BPD-PGS after Bonferroni correction ($p_{adj}$<0.05; Supplementary Table S24). BPD-PGS showed the most statistically significant associations with codes from the mental disorders category (Extended Data Fig. 9), including mood disorders, substance use disorders, anxiety and PTSD, and suicidal ideation or attempt. Associations were also observed in other categories, including neurological (*e.g.*, epilepsy) and somatic disorders (*e.g.*, type 2 diabetes, chronic airway obstruction, hypertension), and pain disorders and symptoms.

**PheWAS in UKB (HES data)**

In the PheWAS analysis in the UKB, 317 of the 1,250 tested diagnoses showed an association with BPD-PGS ($p_{adj}$<0.05; Supplementary Table S25). BPD-PGS showed the most statistically significant associations with codes from the mental disorders category (Extended Data Fig. 10), including mood disorders, tobacco use disorders, and anxiety disorders. As in BioVU, significant associations after Bonferroni correction were observed in other categories, including respiratory (*e.g.*, chronic airway obstruction), digestive (*e.g.*, esophagus, gastroesophageal reflux disease), circulatory, endocrine/metabolic, and neurological diagnoses, and pain disorders and symptoms.

Of the 69 diagnoses that were significant in the BioVU, 56 were also present with sufficient case numbers in the UKB. Of these, 51 showed effects in the same direction and 41 also reached significance in the UKB ($p_{adj}$<0.05). In both samples, the phecode 301.20 (*"antisocial/borderline personality disorder"*) showed the strongest effect size ($OR_{BioVU}$=1.41; $OR_{UKB}$=1.54).

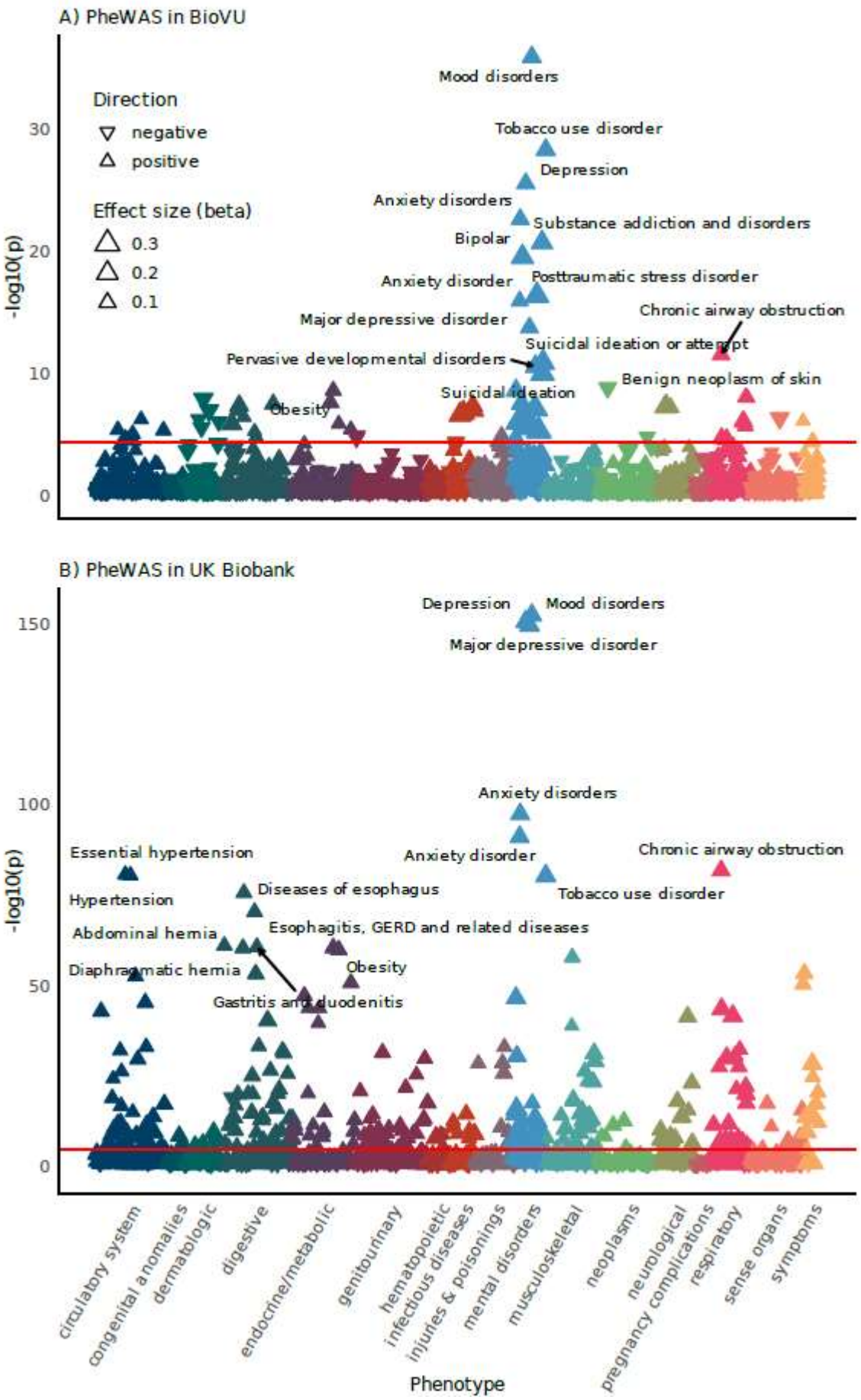

A) PheWAS in BioVU
Direction
negative
positive
Effect size (beta)
0.3
0.2
0.1
-log10(p)
Mood disorders
Tobacco use disorder
Depression
Anxiety disorders
Substance addiction and disorders
Bipolar
Anxiety disorder
Posttraumatic stress disorder
Chronic airway obstruction
Major depressive disorder
Suicidal ideation or attempt
Pervasive developmental disorders
Benign neoplasm of skin
Suicidal ideation
Obesity
B) PheWAS in UK Biobank
Depression
Mood disorders
Major depressive disorder
Anxiety disorders
Essential hypertension
Anxiety disorder
Chronic airway obstruction
Hypertension
Diseases of esophagus
Tobacco use disorder
Abdominal hernia
Esophagitis, GERD and related diseases
Diaphragmatic hernia
Obesity
Gastritis and duodenitis
circulatory system
congenital anomalies
dermatologic
digestive
endocrine/metabolic
genitourinary
hematopoietic
infectious diseases
injuries & poisonings
mental disorders
musculoskeletal
neoplasms
neurological
pregnancy complications
respiratory
sense organs
symptoms
Phenotype

**Figure 5: PheWAS of borderline personality disorder polygenic scores (BPD-PGS) in A) BioVU, and B) UKB: Association of BPD-PGS with 1,431 tested phecodes in BioVU and 1,250 Hospital Episode Statistics (HES) phecodes in UKB.** Statistical significance (-log10(*p*)) is plotted on the y-axis and diagnoses are grouped by category. The red line indicates the significance threshold for Bonferroni correction for the tested phecodes (BioVU: $0.05/1{,}431=3.49\times10^{-5}$; UKB: $0.05/1{,}250=4.00\times10^{-5}$). The top 15 associations of each PheWAS are annotated. Upward-pointing triangles indicate positive associations, and downward-pointing triangles indicate negative associations. The size of the triangles indicates the effect size (beta).

# Discussion

The present GWAS including 12,339 BPD cases and 1,041,717 controls marks a major advance in BPD genetics, with the demonstration of a substantial SNP-heritability, the identification of genome-wide significant risk loci, pathway and tissue enrichment, and a considerably improved predictive value of derived PGS. The study includes the first sex-stratified and X-chromosomal GWAS for BPD, and the first systematic investigation of shared genetic risk of the disorder. The validity and generalizability of the results is supported by the replication of the lead SNP associations and PGS analyses in independent datasets.

The substantial increase in sample size was achieved by using a strategy combining studies with different ascertainment strategies. The approach of our previous study, which included only clinical studies that had explicitly recruited BPD patients[18], was extended by also including biobank and cohort studies that mainly linked diagnoses from electronic health records to genetic data. The high genetic correlation indicates the two subsets capture a largely identical genetic architecture.

A substantial proportion of the heritability observed in family and twin studies (46–69%)[12,13] was explained by common genetic variation (estimated SNP-heritability of 17%). This is consistent with GWAS in other psychiatric disorders, where SNP-heritability often accounts for approximately a third of the estimates from twin and family studies[16]. The substantial contribution of common genetic variation to BPD is also supported by the leave-one-out PGS analyses, predicting a weighted average of 4.6% of the variance on the liability scale. Importantly, the low LDSC intercept (1.02) and the attenuation ratio of 11% indicate the association is largely driven by the polygenic signal and not by confounding.

In the context of identifying genetic risk variants and genes, the present study is the first to identify six genome-wide loci associated with BPD. The genetic risk loci were located in or near genes including *SGCD, EXD3, FOXP1,* and *FOXP2*, and one locus mapped to several genes: *MVK*, *MMAB*, and *UBE3B*. *SGCD* encodes sarcoglycan delta, a component of the sarcoglycan complex, and mutations have been associated with muscular and cardiovascular disorders[33]. Common variation in *SGCD* has also been associated with mental disorders, like SCZ[34,35] and PTSD[36], substance use phenotypes[37–39], and measures of quality

of life[40]. Converging findings in rodents and humans further suggest that the sarcoglycan complex is expressed in the CNS in both neurons and astrocytes[41,42] and may be involved in GABAergic neurotransmission [41] and cerebrovascular function[43].

Among the genes mapped to genes by proximity to the lead SNPs or in the MAGMA gene-based analysis, the highest PoPS scores were observed for *SGCD, FOXP2, FOXP1, DEPDC1B. FOXP1* and *FOXP2* are both members of the Forkhead-box (FOX) transcription factor family that are expressed in the brain and are both associated with speech and language development[44]. The *FOXP2* gene has shown significant associations with externalizing behavior[45,46] and related traits or disorders such as substance use disorders[47–50], broad antisocial behavior[51], and ADHD[52] in several recent GWAS. Observed associations with childhood maltreatment[53] and PTSD[36,54,55] further suggest a trauma-related pathway for BPD, while recent GWAS of the Big Five personality traits showed an association of *FOXP2* with agreeableness[56,57], conscientiousness[57], and neuroticism[57]. Of the genes indicated by the combined analysis, *NME7* (NME/NM23 family member 7) and *KPNA2* (Karyopherin Subunit Alpha 2) had the highest PoPS scores: NME7 is a component of the gamma-tubulin ring complex which plays a role in microtubule organization[58]. It has been associated with antidepressant treatment response in a gene expression study[59] and a GWAS[60]. KPNA2, a key component of the nucleocytoplasmic transport system, has mainly been studied with respect to cancer development, but was recently found with significant brain GWAS/QTL associations in a transcriptomic analysis of MDD[61].

We present the first GWAS to include both sex-stratified and X-chromosome analyses — an initial step towards investigating the sex-specific genetic architecture of the disorder, which is relevant concerning the sex differences in BPD prevalence and presentation. We observed a substantially higher SNP-heritability in females, which has also been described for the two mental disorders showing the strongest genetic correlation with BPD–PTSD and depression[62]. The lead SNPs of the main GWAS, however, showed the same direction and similar effect sizes in the stratified analyses in both sexes, and the genetic correlation of the sex-stratified GWAS was high and did not differ from 1. However, it must be noted that the sex-stratified analyses had limited statistical power, particularly the male-only analysis, which was also indicated by a z-score of the SNP-heritability under 4[63]. Nevertheless, we consider the sex-stratified analysis an important first step and provide it as a resource. Sex-stratified analyses in larger samples may help to elucidate sex-specific risk factors of BPD[64]. Notably, for the first time, we comprehensively analyzed the genetic variation on the X chromosome regarding personality disorders and identified the gene encoding choline-phosphate cytidylyltransferase B (*PCYT1B*) as significantly associated with BPD in the gene-based analysis, as well as an association locus in the gene with genome-wide significance in the combined analysis. *PCYT1B* is expressed in the brain[65,66], but has not been previously

reported to be associated with mental health phenotypes. However, many previous GWAS analyses have not reported X chromosome data[64].

This study is the first to systematically use genetic association data to investigate expression enrichment and potential drug repurposing for BPD, offering initial insights: The results suggest the importance of genes expressed in the brain and during early prenatal development. Single-cell analyses revealed significant SNP-heritability enrichment in genes expressed in the clusters *"medium spiny neurons"* as well as *"LAMP5-LHX6 and Chandelier cells"*. Both clusters have previously been demonstrated to be enriched in mental health phenotypes in a systematic analysis, namely in schizophrenia, IQ, educational attainment, and neuroticism[29], with *"medium spiny neurons"* additionally showing enrichment in MDD[29], and the *"Lamp5-LHX6 and Chandelier cell"* cluster being significantly enriched in the most recent bipolar disorder GWAS[67]. Together, these findings suggest that specific neuronal cell types with established relevance for psychiatric disorders may also play a role in BPD.

In addition, our analyses identified potential drug targets: Drug target analysis of the 16 prioritized genes highlighted *MYO1H,* which has been associated with the response to treatment with lithium, a mood stabilizer also suggested to reduce the risk of suicide attempts[68]. However, it should be noted that the gene was indicated by proximity only in the lithium GWAS[69]. The drug target analysis also suggested *PCYT1B* as a target of sphingosines. The sphingosine-1-phosphate receptor 3 (S1PR3) regulates various cellular processes when bound to its ligand Sphingosine-1-phosphate (S1P), has been implicated in stress resilience in animal studies, and was found to be downregulated in PTSD[70]. Notably, the S1P-S1P3 pathway was the only significant gene set in the present analysis. The genome-wide DRUGSETS analysis indicated drugs with high biological plausibility, including medication for neurological disorders and pain, albeit our results were not statistically significant after correction for multiple testing and should be interpreted with caution.

Through genetic correlation and PheWAS analyses, we address a critical gap in the literature by systematically and comprehensively assessing the shared genetic architecture of BPD across a broad spectrum of other disorders and traits. We confirmed the genetic correlations observed in the previous BPD GWAS[18] with depression, BIP, SCZ[18], neuroticism, openness to experience[71], and loneliness[72]. With respect to the Big Five personality dimensions, the results highlight the enhanced power of the present BPD GWAS: We now additionally observe negative genetic correlations with agreeableness and conscientiousness, which is consistent with the profiles of the Big Five personality dimensions observed in cases with BPD[73–75], as well as with data from twin models[76,77]. This suggests that genetic factors underlying variation

in personality traits in the general population contribute to the risk for BPD. Among mental disorders, BPD showed the strongest genetic correlations with PTSD, depression, ADHD, and anxiety. Notably, these disorders are frequently observed as preceding or comorbid conditions and share clinical features with BPD[5,7,78]. The associations with suicide and self-harm are consistent with the clinical presentation of BPD[79], the correlations with trauma phenotypes highlight the role these experiences play in BPD[10], and the strong correlation of BPD with disorders from the internalizing/externalizing spectrum disorders suggests that BPD risk is influenced by the liability for both dimensions[79,80].

In both PheWAS samples, and consistent with the genetic correlation results, the BPD-PGS were significantly associated with mood disorders, anxiety disorders, PTSD, substance use disorders, as well as suicidal ideation or suicide attempts. The strongest effect was observed for the phecode *"antisocial/borderline personality disorder"* supporting the specificity of the GWAS signal for BPD. BPD-PGS were also associated with a range of physical health phenotypes, including chronic airway obstruction, type 2 diabetes, obesity, and hypertension, for which an increased risk in BPD patients has been previously described[7]. The relatively small number of BPD cases in the PheWAS target samples makes it likely that these associations are driven by shared genetic risk, and not solely by comorbidities of BPD patients in the target samples. These results provide a promising starting point for the investigation of the shared genetic risks of BPD and somatic health, which is highly relevant as a large share of the reduced life expectancy of people with BPD is related to physical health problems[81]. Further research is warranted to understand the mechanisms through which inter-individual differences in genetic liability for BPD influence the risk for different disorders.

The present analyses cannot dissect to what degree comorbidities, the overlap of symptoms, and potential diagnostic misclassifications might have influenced our results[82]. While we addressed the potential overlap with BIP and SCZ, more fine-grained analyses are needed. In BPD, comorbidity with other mental disorders is the rule rather than the exception[1,5], for example, approximately 70% of the investigated BPD patients had a history of depression. Restricting GWAS cases to those without psychiatric comorbidity would reduce the number of available subjects and limit the analysis to a less representative (and drastically smaller) subset. Therefore, we consider it a strength of the present analysis that the biobanks and cohorts providing summary statistics performed the main analysis without excluding cases with psychiatric comorbidity. To assess the impact of this strategy, the sensitivity analysis excluded subjects with either a diagnosis of BIP or SCZ, which are common exclusion criteria for dedicated BPD studies, accounting for 30% of the cases. At the level of the single-variant associations, we observed comparable effect sizes in the sensitivity analysis. We, therefore, consider it unlikely that SCZ or BIP comorbidity substantially biased the GWAS hits. Comparing the genetic correlation of the two iterations of the BPD GWAS with the GWAS of

BIP and SCZ, we observe an attenuated but still substantial genetic correlation in the sensitivity analysis, suggesting shared genetic risk beyond the co-occurrence of the disorders.

The present study represents substantial progress in the identification of the genetic factors underlying BPD. However, the study has several limitations. First, the sample size is still relatively small compared to those of other mental disorders[16], and the power analysis indicated that the statistical power was limited, especially for variants with a lower frequency. The inclusion of additional samples will likely lead to a substantial increase in the number of loci identified[83,84]. Furthermore, expanding the analyses to non-European ancestries is necessary to improve the understanding of the underlying genetic architecture and to facilitate the generalizability of the results[85,86]. Additionally, in this study, we investigated BPD as a categorical diagnosis, as it is also commonly used in the clinical context. However, the diagnosis can be heterogeneous and might differ between cohorts. Future studies should consider the heterogeneity of the disorder, and examine clinical BPD at a more fine-grained level, *e.g.*, by examining individual symptoms or symptom clusters[87,88]. In addition, functional dimensions such as RDoC[89] and classification systems such as HiTop[90,91] as well as borderline personality traits or symptom dimensions[2,17,92] should be incorporated.

In summary, the present GWAS meta-analysis represents a major step forward towards an understanding of the genetic etiology of BPD. It is the first GWAS of a personality disorder to identify genome-wide significant risk loci and genes, and demonstrate that BPD, like other mental disorders, is a complex polygenic disorder.

# Online methods

## Sample description

An overview of the analyses performed can be found in Fig. 1. We conducted a GWAS meta-analysis, including 1,054,056 participants of European ancestry ($N_{cases}$=12,339 $N_{controls}$=1,041,717, $N_{eff}$=2×21,617). In total, 17 studies contributed individual-level data ($N_{cases}$=2,705, $N_{controls}$=4,600 controls). This included data from the prior BPD GWAS[18], and data from an additional 1,712 cases and 3,061 controls (meeting DSM-IV criteria for BPD). Furthermore, we included data from 10 large-scale biobank or cohort studies that provided summary statistics from GWAS of BPD ($N_{cases}$=9,634, $N_{controls}$=1,037,117 controls). Of those, nine cohorts identified BPD status following ICD codes (ICD-10: F60.3, ICD-9: 3018D/301.83, and ICD-8: 3013), while in the GLAD study, a self-reported diagnosis was used.

Details on ascertainment, inclusion and exclusion criteria for cases and controls, and country of origin for each sample are documented in Supplementary Table S1. A detailed description of the study design, ascertainment of cases and controls, and genotyping array for

each study, as well as quality control and imputation specific to studies sharing summary statistics are provided in Supplementary Table S2 and the Supplementary Methods. All subjects gave informed consent, and the studies were approved by the respective ethical committees.

## Genotyping, quality control, and imputation

Samples providing genotype data at the individual level were grouped into datasets based on array and ancestry, resulting in five datasets (Supplementary Table S2). Quality control and imputation were carried out using the RICOPILI pipeline [93]. A detailed description can be found in the Supplementary Methods.

## Genome-wide association study (GWAS)

For the studies providing genotype data at the individual level, GWAS were carried out within each dataset using an additive logistic regression model using PLINK (v1.9)[94] for imputed genetic dosage data with the relevant ancestry principal components (PCs; details in Supplementary Methods) included as covariates. Details on the association analyses for the cohorts providing summary statistics can be found in Supplementary Table S2 and Supplementary Methods. In addition to the main analysis, we performed sex-stratified analyses. Dedicated studies of BPD often exclude subjects with schizophrenia and bipolar disorder. To explore the influence of comorbidity with these severe mental disorders on our results, as a sensitivity analysis, association analyses were performed excluding subjects with a history of BIP or SCZ from the studies providing summary statistics (sensitivity analysis: $N_{cases}$=8,618, $N_{controls}$=1,027,690; Supplementary Tables S3–S5). In the case of the cohorts, where the control population was not filtered for BIP and SCZ in the main analysis (Copenhagen Hospital Biobank and Danish Blood Donor Study; deCODE; Mayo Clinic Biobank; FinnGen), subjects with BIP or SCZ were also excluded from the controls for the sensitivity analysis.

## Meta-analysis

Meta-analysis of all samples providing either individual-level data or summary statistics was conducted using the inverse variance weighted fixed effects model in METAL[95] as implemented in RICOPILI[93] with a genome-wide significance threshold of $5\times10^{-8}$. To evaluate consistency of effect sizes, we computed Cochran's Q-statistics and the corresponding heterogeneity p-values and $I^2$ statistics, which quantifies the percentage of total variation due to heterogeneity. This procedure excluded single-nucleotide polymorphisms (SNPs) with minor allele frequency (MAF)<1% or an imputation quality score (INFO)<0.60. SNPs were translated to the HG19 genomic build (Human build GRCh37) when necessary, and both SNPs and SNP alleles were aligned to the haplotype reference consortium (HRC) reference genome to ensure standardization for the meta-analysis. To ensure a robust analysis with highly credible SNP sets, we excluded SNPs with highly significant heterogeneity ($p$ <0.001) and SNPs with an effective sample size of less than 85%. X chromosome markers were included in the meta-analysis for all samples except the Norwegian Mother, Father and Child Cohort Study (MoBa), where data were not available at the time of analysis. To assess the similarity between the five datasets of studies providing individual-level data and the 10 studies providing summary statistics, separate meta-analyses were calculated for the two subsets.
The statistical power of the main analysis to detect genome-wide significant associations ($p<5\times10^{-8}$), was calculated using the Genetic Power Calculator[19]. For a range of allele frequencies (0.01–0.99), and an effect size of 1.1, we calculated the power of the main analysis ($N_{eff}$ = 2×21,617, calculated as described in [96]), and the required sample size for a power over 80% (details see Supplementary Table S6, Supplementary Figure S1).

SNP-heritability, based on the autosomal markers, was estimated using LDSC[20], both as the observed SNP-heritability and as the SNP-heritability converted to the liability scale, using the $N_{eff}$ (2×21,617=43,234)[96], a corresponding sample prevalence of 0.5, and assuming a population prevalence of 1.5%[3] for the main analysis. For the sex-stratified GWAS, we used sex-specific population prevalences of 2.25% in females and 0.75% in males for the conversion to the liability scale, corresponding to the previously described female:male ratio of 3:1[4,5]. To assess the contribution of confounding to the observed signal, we calculated the attenuation ratio ($\frac{(Intercept-1)}{(mean\ \chi 2\ \ -1)}$).

To test whether the SNP-heritability estimates of the sex-stratified GWAS differed, and to test whether the genetic correlations between GWAS subsets (males vs. females; individual-level vs summary cohorts) differed from 1, we used the block jackknife extension of LDSC[20], which has previously been described[97].

To assess replication of the observed SNP associations, lead SNPs for the thresholds $p<1x10^{-6}$ and $p<5x10^{-8}$ were tested for replication in a meta-analysis of two independent datasets (Spain 2 & All Of Us; total $N_{cases}$ = 685, total $N_{controls}$ = 107,750). Additionally, for SNPs with $p<1x10^{-6}$ in the discovery GWAS analysis, discovery and replication datasets were analyzed in a combined GWAS meta-analysis, and loci with $p<5x10^{-8}$ in the combined analysis are reported.

## Gene-based and gene-set tests

Gene-based, gene-set, and tissue enrichment tests were carried out using MAGMA (version 1.07) [21] as implemented in FUMA (version 1.5.2) [22]. Gene-based tests were performed using the summary statistics of the GWAS meta-analysis for a total of 19,843 genes, with SNPs assigned to genes based on their physical position. SNPs were included in the analysis using boundaries of 35 kilobases (kb) upstream and 10kb downstream of the genes. Gene-sets were tested by analyzing a total of 9,237 gene-sets (MSigDB v2023.1Hs; limited to gene sets with at least 20 mapped genes). For both gene-based and gene-set tests, a Bonferroni *p*-value threshold, corrected for the number of respective tests, was applied (gene-based $\alpha=0.05/19,843=2.5\times10^{-6}$; gene-set $\alpha=0.05/9,237=5.4\times10^{-6}$).

Additionally, Polygenic Priority Scores (PoPS) [23] were calculated, which prioritize genes at GWAS loci using MAGMA gene-level association tests and over 57,000 gene features such as cell-type specific expression gene expression, biological pathways, and protein-protein interactions. PoPS scores were available for 17,702 autosomal genes, and PoPS scores and ranks are reported.

Tissue enrichment expression was carried out using Genotype-Tissue Expression (GTEx; version 8) data for 53 tissue types[24], and data from BrainSpan, representing 29 different ages and 11 general developmental stages[25] as implemented in FUMA (version 1.5.2)[22]. In addition, we utilized the Human Brain Atlas dataset containing single-nucleus RNA sequencing data from approximately 3.369 million nuclei derived from adult post-mortem brain samples, covering 106 anatomical sections[27]. We used the cell-types at the level of 31 “superclusters”, and maintained the existing nomenclature for cell-types. We calculated expression proportion values for each cell-type (“gene specificity”, described in detail in[29] and[98]), resulting in ~1300 genes per cell-type. We subsequently tested the enrichment of heritability in these gene sets using stratified LDSC [28], adjusting for the 53 baseline annotations. To adjust for multiple hypothesis testing in the tissue expression analyses, we applied a Bonferroni p-value threshold adjusted of the respective number of tested tissues.

## Drug target analysis

For the drug target analysis, we prioritized genes that either mapped to the genome-wide risk loci (Table 1) or were significant in the gene-based analysis. To identify drug targets among these genes, we extracted data from the Open Targets platform using the GraphSGL API. The information provided by Open Targets is based on the ChEMBL database [99]. Additionally, we performed a tractability analysis (*i.e.*, the potential to be modulated by a drug) to assess small molecule binding, the presence of accessible epitopes for antibody-based therapy, relevant data for using Proteolysis Targeting Chimeras (PROTACs), and the presence of compounds in clinical trials with modalities other than small molecules or antibodies. Databases were queried on 2024-06-24. Furthermore, we used the Genome for REPositioning drugs (GREP; https://github.com/saorisakaue/GREP) pipeline to test for drug target enrichment across clinical ATC or ICD categories [100]. Finally, we queried the Drug Gene Interaction Database (DGIdb) [101].

In addition to these approaches focussing on the prioritized 16 genes, we explored potential drug targets with the DRUGSETS approach[30], which integrates the genome-wide association signal. Using this approach we tested drug–gene sets, compiled from the Clue Repurposing Hub and the Drug–Gene Interaction Database, each comprising the genes whose protein products are known to be targeted by or to interact with a given drug. We then conducted a competitive gene-set analysis in MAGMA v1.08, conditioning on a background set of all 2,281 drug-target genes, to identify drug-gene sets significantly associated with BPD. To account for multiple testing across the 735 gene sets analyzed, we applied a Bonferroni threshold of $p<0.05/735 = 6.80\times10^{-5}$.

## Polygenic scores

We applied polygenic scoring (PGS) to predict case-control status in the datasets where individual-level genotype information was available. Polygenic scores were calculated using PRS-CS, a method that uses Bayesian regression to calculate updated (posterior) effect sizes by applying continuous shrinkage to the initial (prior) effect sizes from the discovery dataset using linkage disequilibrium (LD) information[31]. The posterior effect sizes account for the LD between SNPs using external LD reference panels constructed from the 1000 Genomes Project Phase 3 European samples. PGS were calculated based on both the present meta-analysis and also the prior BPD GWAS from 2017 (consisting of the Germany dataset) [18] for comparison. Using leave-one-out results of the present meta-analysis (2024 meta-GWAS), where the respective target dataset was always left out of the meta-analysis, we calculated PGS for all datasets with individual-level genotype information (Germany, Central Europe, Spain 1, Norway 1, and Norway 2 as well as the two replication cohorts (Spain 2 & All Of Us;

total $N_{cases}$=685, total $N_{controls}$=107,750; Supplementary Tables S1&S2)). Similarly, we used the results of the prior BPD GWAS (2017 GWAS) to calculate PGS for the datasets not included in that analysis (Central Europe, Spain 1, Norway 1, Norway 2, and Spain 2, All Of Us).

As an effect measure of the association between BPD-PGS and case-control status, *Nagelkerke-pseudo-$R^2$* (*$NkR^2$*) was calculated comparing the $R^2$ of the full model including PGS and covariates (ancestry principal components) as predictors to the reduced (null) model including the covariates only. The resulting $NkR^2$ was then converted to the liability scale [102] of the population assuming a lifetime disease risk of 1.5%[3]. Additionally, the area under the receiver operator characteristic curve (AUC) was calculated for each target dataset. Average $NkR^2$ and AUC were calculated across the five target datasets and weighted by the effective sample size of the respective target samples.

To calculate the odds ratio of the deciles of the PGS distribution, PRS-CS were normalized to have a mean of 0 and unit variance before combining them into two sets of scores for comparison: one based on Witt 2017 ($N_{controls}$=3,496; $N_{cases}$=1,758) and the other on BPD-meta 2024 ($N_{controls}$=5,035; $N_{cases}$=2,751). This part of the PGS analysis excluded the All of Us dataset, as its individual-level data could not be merged due to data protection regulations. Using the normalized PRS-CS, we categorized the data into ten deciles through quantile binning, assigning each observation a decile number ranging from 1 (lowest PRS-CS) to 10 (highest PRS-CS). We then created dummy variables by coding observations within each decile as cases, while those outside that decile were coded as controls. The dummy variable ranged from deciles 2 (Q2) to 10 (Q10), with decile 1 (Q1) serving as the reference category.

To assess the association between the deciles based on PRS-CS and the actual case and control status, we conducted logistic regression analyses using the decile-based dummy variables for Q2 to Q10. The odds ratios (OR) for each decile were calculated by exponentiating the coefficients from the logistic regression model, controlling for PCs (1, 2, 3, 4, 5, 6, 8, 10, 12, 14, 15, and 20). For the reference comparison (Q1/Q1), an odds ratio of 1 was assigned, indicating no effect. The odds ratios for each of the remaining deciles (Q2 to Q10) were computed accordingly $OR_{Q1/Qj}=\exp(\beta j)$ for j=2,3,…,10. Additionally, 95%-CI for the odds ratios were calculated using the standard errors of the coefficients ($CI_{lower}=\exp(\beta j-1.96\times SEj)$; $CI_{upper}=\exp(\beta j+1.96\times SEj)$). As an additional measure of effect, the highest decile (Q10) was compared to the middle 10% of the distribution as reference, and OR and CI were calculated as described above.

## Genetic correlations

Genetic correlations between subsets and with other disorders and traits were calculated using LD score regression[20]. Calculations were carried out with a free intercept and the 1000 Genomes dataset (EUR) as a reference LD structure panel [103].

In a targeted approach, genetic correlations were calculated with 50 GWAS of a range of other disorders and traits relevant to BPD, selected based on reported phenotypic and genetic associations in the literature, theoretical consideration, and data availability. Those included GWAS of mental disorders, suicide, self-harm, trauma, substance use, physical health, pain, sleep, personality traits (Big Five GWAS including data from 23andMe, Inc.), and cognition (Supplementary Table S22). A Bonferroni p-value threshold, corrected for the number of respective tests was applied ($\alpha=0.05/50=1\times10^{-3}$).

## Phenome-wide association studies (PheWAS)

The targeted genetic correlation analyses were complemented by PheWAS, systematically exploring associations of the polygenic predisposition for BPD with over 1,000 diagnosis-based phenotypes across the diagnostic spectrum.

**PheWAS in BioVU (EHR data)**

Data from the Vanderbilt University Medical Center Biobank (BioVU) [32] were used to test the association of BPD-PGS with medical phenotypes based on two or more International Classification of Disease (ICD) diagnoses ("phecodes"; Phecode Map 1.2) documented in the electronic health records (EHR). For each phecode, a case was defined as having ≥1 ICD code in the phecode group on ≥2 distinct dates, and control was defined as having 0 codes in that phecode group. Individuals with only 1 single instance of a code in that phecode group were excluded from the case/control definition for that phecode. PGS were calculated in 66,325 unrelated subjects of European genetic ancestry (214 individuals with phecode 301.20 (*"antisocial/borderline personality disorder"*) based on summary statistics excluding the BioVU samples from the discovery meta-analysis (12,024 cases; 1,038,567 controls). PGS were computed using a continuous shrinkage prior (CS) to SNP effect sizes using PRS-CS[31] and standardized to have a mean of 0 and a standard deviation of 1. Phecodes were tested for association with BPD-PGS when at least 100 cases were available using logistic regression models with sex (defined as sex reported at birth from the EHR), age, and the first 10 genetic PCs as covariates. 1,431 phecodes were included, and a Bonferroni-corrected threshold of $\alpha=0.05/1{,}431=3.49\times10^{-5}$ tests was applied.

**PheWAS in UKB**

An additional PheWAS analysis was carried out using data from the UK Biobank (UKB). Here, PGS for BPD were calculated using PRS-CS based on summary statistics excluding the UKB

samples from the discovery meta-analysis (12,157 cases; 1,039,897 controls) generated by excluding the UKB. The PheWAS was calculated using phecodes (Phecode Map 1.2) based on diagnoses recorded in Hospital Episode Statistics (HES) in a maximum sample of 316,635 subjects (211 individuals with phecode 301.20). For each phecode, a case was defined as having ≥1 ICD code in the phecode group, and control was defined as having 0 codes in that phecode group. A total of 1,250 phecodes with at least 100 cases were tested for association using a Bonferroni-corrected threshold of $\alpha=0.05/1{,}250=4.00\times10^{-5}$ tests.

# Data availability statement

Individual-level data are not publicly available due to ethical restrictions. The results of the meta-analysis data will be publicly available for download upon publication of the manuscript via (via figshare.com; https://pgc.unc.edu/for-researchers/download-results/).

GWAS summary statistics of other traits and disorders used for analyses in this study are publicly available (sources listed in Supplementary Table S22). There exception are the GWAS summary statistics for self-harm[104] which were provided by the respective corresponding authors and the GWAS used for the Big Five personality traits (except Neuroticism) which include data from 23andMe, and can be made available to qualified investigators, if they enter into an agreement with 23andMe that protects participant confidentiality. This study used data from the All of Us Research Program's Controlled Tier Dataset v7, available to authorized users on the Researcher Workbench.

# Code availability statement

Quality control and imputation were carried out using RICOPILI (https://sites.google.com/a/broadinstitute.org/ricopili/; https://github.com/Ripkelab/ricopili). GWAS in datasets with individual data were carried out within each dataset using PLINK 1.9 (http://pngu.mgh.harvard.edu/purcell/plink/). Prephasing/imputation was carried out using EAGLE 2.4.1 (https://alkesgroup.broadinstitute.org/Eagle/) and MINIMAC3 (http://genome.sph.umich.edu/wiki/minimac3). Meta-analyses of GWAS were performed using METAL: (http://www.sph.umich.edu/csg/abecasis/metal/). Genetic correlations and heritability estimates were calculated using LDSC (https://github.com/bulik/ldsc).
FUMA 1.5.2 was used for downstream analyses including gene-based and gene-set analyses (https://fuma.ctglab.nl/), based on MAGMA 1.07 (http://ctglab.nl/software/magma).
Additionally, genes were prioritized using Polygenic Priority Score (PoPS) v0.2: (https://github.com/FinucaneLab/pops).

Enrichment for drug targets was carried out using GREP (https://github.com/saorisakaue/GREP), and DRUGSETS (https://github.com/nybell/drugsets). Weights for polygenic scores were generated using PRS-CS v.1.1.0 (https://github.com/getian107/PRScs).

# Acknowledgments

We thank all research participants and all researchers and clinicians who collected, generated, or processed the data used in this study. We thank all research participants and employees of 23andMe for making this work possible. We gratefully acknowledge *All of Us* participants for their contributions, without whom this research would not have been possible. We also thank the National Institutes of Health's *All of Us* Research Program for making available the participant data examined in this study. We thank FinnGen for providing access to the meta-analysis results. The Genotype-Tissue Expression (GTEx) Project was supported by the Common Fund of the Office of the Director of the National Institutes of Health, and by NCI, NHGRI, NHLBI, NIDA, NIMH, and NINDS.

Fabian Streit is supported by a 2023 NARSAD Young Investigator Grant (#31537) from the Brain & Behavior Research Foundation with support from the Families for Borderline Personality Disorder Research. This research was supported by the Hector foundation II. Björn H Schott receives funding from the German Research Foundation (DFG CRC 1436, SP A05). Alice Braun was supported by the European Union's Horizon program grant 101057454 - "PsychSTRATA", by the German Research Foundation grant 402170461 - "TRR265", and the Berlin Institute of Health at Charité. Elizabeth C Corfield was supported by the RCN (#274611) and the South-Eastern Norway Regional Health Authority (Helse Sør-Øst; #2021045). Ted Reichborn-Kjennerud was supported by the RCN (#274611). AH is supported by the Research Council of Norway (#336085), the South-Eastern Norway Regional Health Authority (#2020022, #2922083, #2022039, #2019097), and the European Union's Horizon Europe Research and Innovation programme (FAMILY #101057529; HOMME #101142786; Marie Skłodowska-Curie grant ESSGN #101073237). Urs Heilbronner was supported by the European Union's Horizon 2020 Research and Innovation Program (PSY-PGx, grant agreement No 945151) and the DFG (project number 514201724). Søren Brunak received funding from the Novo Nordisk Foundation (grants NNF17OC0027594 and NNF14CC0001). Ole Kristian Drange was supported by funding from the Research Council of Norway (RCN #223273, #324499). Kelli Lehto was supported by the Estonian Research Council (grant PSG615), Estonian Centre of Excellence for Well-Being Sciences, funded by the Estonian Ministry of Education and Research (grant TK218) and the European Union's Horizon program

grant 847776. Jean Gonzalez received funding from the National Institutes of Health (grant T32GM139790).
VN was supported by the DFG (DFG, NI 1332/16-1), and the German Federal Ministry of Education and Research (BMBF) through support of the German Center for Mental Health (DZPG, 01EE2306B)]
This research was supported by the NHMRC grants 1172917 and 2025674.

# Conflicts of Interest

OAA is a consultant to Cortechs.ai and Precision Health and has received speaker's honoraria from Lundbeck, Janssen, Sunovion, and Otsuka.
SB has ownerships in Hoba Therapeutics Aps, Novo Nordisk A/S, Lundbeck A/S, and Eli Lilly and Co.
JLK is a Scientific Advisory Board member of Myriad Neuroscience.
RM has received financial research support from Böhringer-Ingelheim and Otsuka Pharmaceuticals. He has received speakers' honoraria from Otsuka Pharmaceuticals and Lundbeck and is a member of the advisory board of Böhringer-Ingelheim.
FP is a member of Scientific Advisory Boards of Brainsway Inc., Jerusalem, Israel, and of Sooma, Helsinki, Finland. He has received speaker's honoraria from Mag&More GmbH, and the neuroCare Group, Munich, Germany.
JARQ was on the speakers' bureau and/or acted as a consultant for Biogen, Idorsia, Casen-Recordati, Janssen-Cilag, Novartis, Takeda, Bial, Sincrolab, Neuraxpharm, Novartis, BMS, Medice, Rubió, Uriach, Technofarma and Raffo in the last 3 years. He also received travel awards (air tickets + hotel) for taking part in psychiatric meetings from Idorsia, Janssen-Cilag, Rubió, Takeda, Bial, and Medice. The Department of Psychiatry chaired by him received unrestricted educational and research support from the following companies in the last 3 years: Exeltis, Idorsia, Janssen-Cilag, Neuraxpharm, Oryzon, Roche, Probitas, and Rubió.
CS received lecturer fees from Idorsia Pharmaceuticals GmbH (2024).
MMN has received fees for membership in an advisory board from HMG Systems Engineering GmbH (Fürth, Germany), for membership in the Medical-Scientific Editorial Office of the Deutsches Ärzteblatt, for review activities from the European Research Council (ERC), and for serving as a consultant for EVERIS Belgique SPRL in a project of the European Commission (REFORM/SC2020/029). MMN receives salary payments from Life & Brain GmbH and holds shares in Life & Brain GmbH.
The other authors declare no competing interests.

# References


1. Bohus, M. *et al.* Borderline personality disorder. *The Lancet* vol. 398 1528–1540 Preprint at https://doi.org/10.1016/s0140-6736(21)00476-1 (2021).
2. Leichsenring, F. *et al.* Borderline personality disorder: a comprehensive review of diagnosis and clinical presentation, etiology, treatment, and current controversies. *World Psychiatry* **23**, 4–25 (2024).
3. Volkert, J., Gablonski, T.-C. & Rabung, S. Prevalence of personality disorders in the general adult population in Western countries: systematic review and meta-analysis. *Br. J. Psychiatry* **213**, 709–715 (2018).
4. Skodol, A. E. & Bender, D. S. Why are women diagnosed borderline more than men? *Psychiatr. Q.* **74**, 349–360 (2003).
5. Grant, B. F. *et al.* Prevalence, correlates, disability, and comorbidity of DSM-IV borderline personality disorder: results from the Wave 2 National Epidemiologic Survey on Alcohol and Related Conditions. *J. Clin. Psychiatry* **69**, 533–545 (2008).
6. Tate, A. E. *et al.* Borderline personality disorder: associations with psychiatric disorders, somatic illnesses, trauma, and adverse behaviors. *Mol. Psychiatry* **27**, 2514–2521 (2022).
7. Hastrup, L. H., Jennum, P., Ibsen, R., Kjellberg, J. & Simonsen, E. Borderline personality disorder and the diagnostic co-occurrence of mental health disorders and somatic diseases: A controlled prospective national register-based study. *Acta Psychiatr. Scand.* **149**, 124–132 (2024).
8. Storebø, O. J. *et al.* Psychological therapies for people with borderline personality disorder. *Cochrane Database Syst. Rev.* **5**, CD012955 (2020).
9. Stoffers-Winterling, J. M. *et al.* Pharmacological interventions for people with borderline personality disorder. *Cochrane Database Syst. Rev.* **11**, CD012956 (2022).
10. Kleindienst, N., Vonderlin, R., Bohus, M. & Lis, S. Childhood adversity and borderline personality disorder. Analyses complementing the meta-analysis by Porter et al. (2020). *Acta Psychiatr. Scand.* **143**, 183–184 (2021).
11. Porter, C. *et al.* Childhood adversity and borderline personality disorder: a meta-analysis. *Acta Psychiatr. Scand.* **141**, 6–20 (2020).
12. Skoglund, C. *et al.* Familial risk and heritability of diagnosed borderline personality disorder: a register study of the Swedish population. *Mol. Psychiatry* **26**, 999–1008 (2021).
13. Torgersen, S. *et al.* A twin study of personality disorders. *Compr. Psychiatry* **41**, 416–425 (2000).
14. Kendler, K. S., Myers, J. & Reichborn-Kjennerud, T. Borderline personality disorder traits and their relationship with dimensions of normative personality: a web-based cohort and twin study. *Acta Psychiatr. Scand.* **123**, 349–359 (2011).
15. Distel, M. A. *et al.* The five-factor model of personality and borderline personality disorder: a genetic analysis of comorbidity. *Biol. Psychiatry* **66**, 1131–1138 (2009).
16. Andreassen, O. A., Hindley, G. F. L., Frei, O. & Smeland, O. B. New insights from the last decade of research in psychiatric genetics: discoveries, challenges and clinical implications. *World Psychiatry* **22**, 4–24 (2023).
17. Streit, F., Colodro-Conde, L., Hall, A. S. M. & Witt, S. H. Genomics of borderline personality disorder. in *Personalized Psychiatry* 227–237 (Elsevier, 2020). doi:10.1016/b978-0-12-813176-3.00019-5.

18. Witt, S. H. *et al.* Genome-wide association study of borderline personality disorder reveals genetic overlap with bipolar disorder, major depression and schizophrenia. *Transl. Psychiatry* **7**, e1155 (2017).
19. Purcell, S., Cherny, S. S. & Sham, P. C. Genetic Power Calculator: design of linkage and association genetic mapping studies of complex traits. *Bioinformatics* **19**, 149–150 (2003).
20. Bulik-Sullivan, B. K. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* **47**, 291–295 (2015).
21. 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).
22. Watanabe, K., Taskesen, E., van Bochoven, A. & Posthuma, D. Functional mapping and annotation of genetic associations with FUMA. *Nat. Commun.* **8**, 1826 (2017).
23. Weeks, E. M. *et al.* Leveraging polygenic enrichments of gene features to predict genes underlying complex traits and diseases. *Nat. Genet.* **55**, 1267–1276 (2023).
24. GTEx Consortium. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science* **348**, 648–660 (2015).
25. Miller, J. A. *et al.* Transcriptional landscape of the prenatal human brain. *Nature* **508**, 199–206 (2014).
26. Schaefer, C. F. *et al.* PID: the Pathway Interaction Database. *Nucleic Acids Res.* **37**, D674–9 (2009).
27. Siletti, K. *et al.* Transcriptomic diversity of cell types across the adult human brain. *Science* **382**, eadd7046 (2023).
28. Finucane, H. K. *et al.* Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nat. Genet.* **47**, 1228–1235 (2015).
29. Yao, S. *et al.* Connecting genomic results for psychiatric disorders to human brain cell types and regions reveals convergence with functional connectivity. *medRxiv* (2024) doi:10.1101/2024.01.18.24301478.
30. Bell, N., Uffelmann, E., van Walree, E., de Leeuw, C. & Posthuma, D. Using genome-wide association results to identify drug repurposing candidates. *Genetic and Genomic Medicine* (2022).
31. 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).
32. Roden, D. M. *et al.* Development of a large-scale de-identified DNA biobank to enable personalized medicine. *Clin. Pharmacol. Ther.* **84**, 362–369 (2008).
33. Vainzof, M., Souza, L. S., Gurgel-Giannetti, J. & Zatz, M. Sarcoglycanopathies: an update. *Neuromuscul. Disord.* **31**, 1021–1027 (2021).
34. Goes, F. S. *et al.* Genome-wide association study of schizophrenia in Ashkenazi Jews. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* **168**, 649–659 (2015).
35. Trubetskoy, V. *et al.* Mapping genomic loci implicates genes and synaptic biology in schizophrenia. *Nature* **604**, 502–508 (2022).
36. Maihofer, A. X. *et al.* Enhancing Discovery of Genetic Variants for Posttraumatic Stress Disorder Through Integration of Quantitative Phenotypes and Trauma Exposure Information. *Biol. Psychiatry* **91**, 626–636 (2022).
37. Liu, M. *et al.* Association studies of up to 1.2 million individuals yield new insights into the genetic etiology of tobacco and alcohol use. *Nat. Genet.* **51**, 237–244 (2019).
38. Zhou, H. *et al.* Genome-wide meta-analysis of problematic alcohol use in 435,563 individuals yields insights into biology and relationships with other traits. *Nat. Neurosci.* **23**, 809–818 (2020).

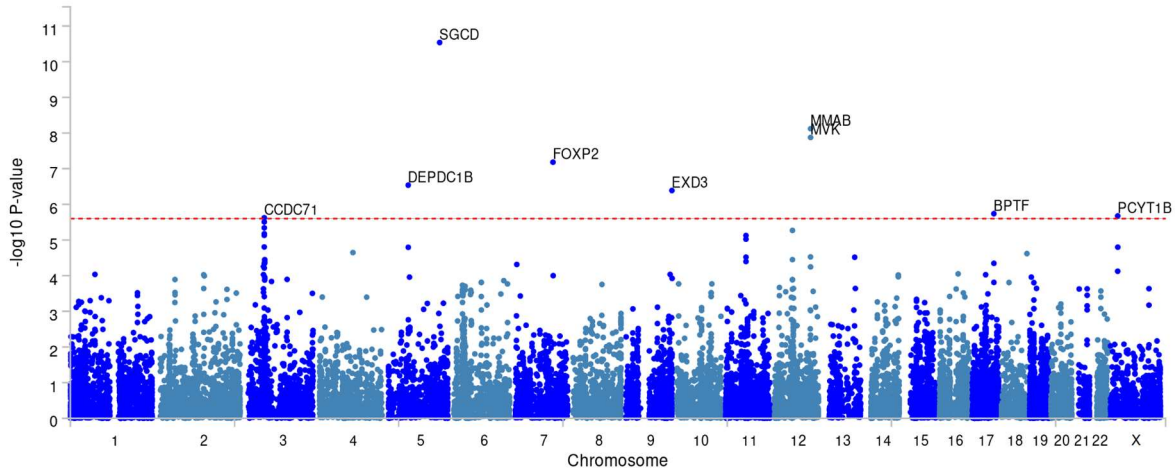
39. Saunders, G. R. B. *et al.* Genetic diversity fuels gene discovery for tobacco and alcohol use. *Nature* **612**, 720–724 (2022).
40. Adjei, A. A. *et al.* Genetic Variations and Health-Related Quality of Life (HRQOL): A Genome-Wide Study Approach. *Cancers* **13**, (2021).
41. Cutroneo, G. *et al.* Sarcoglycans and gaba(a) receptors in rat central nervous system: an immunohistochemical study. *Ital. J. Anat. Embryol.* **120**, 105–116 (2015).
42. Anastasi, G. *et al.* Expression of sarcoglycans in the human cerebral cortex: an immunohistochemical and molecular study. *Cells Tissues Organs* **196**, 470–480 (2012).
43. Boulay, A.-C. *et al.* The Sarcoglycan complex is expressed in the cerebrovascular system and is specifically regulated by astroglial Cx30 channels. *Front. Cell. Neurosci.* **9**, 9 (2015).
44. den Hoed, J., Devaraju, K. & Fisher, S. E. Molecular networks of the FOXP2 transcription factor in the brain. *EMBO Rep.* **22**, e52803 (2021).
45. Baselmans, B. *et al.* The Genetic and Neural Substrates of Externalizing Behavior. *Biol Psychiatry Glob Open Sci* **2**, 389–399 (2022).
46. Karlsson Linnér, R. *et al.* Multivariate analysis of 1.5 million people identifies genetic associations with traits related to self-regulation and addiction. *Nat. Neurosci.* **24**, 1367–1376 (2021).
47. Johnson, E. C. *et al.* A large-scale genome-wide association study meta-analysis of cannabis use disorder. *Lancet Psychiatry* **7**, 1032–1045 (2020).
48. Levey, D. F. *et al.* Multi-ancestry genome-wide association study of cannabis use disorder yields insight into disease biology and public health implications. *Nat. Genet.* **55**, 2094–2103 (2023).
49. Xu, H. *et al.* Identifying genetic loci and phenomic associations of substance use traits: A multi-trait analysis of GWAS (MTAG) study. *Addiction* **118**, 1942–1952 (2023).
50. Deak, J. D. *et al.* Genome-wide association study in individuals of European and African ancestry and multi-trait analysis of opioid use disorder identifies 19 independent genome-wide significant risk loci. *Mol. Psychiatry* **27**, 3970–3979 (2022).
51. Tielbeek, J. J. *et al.* Uncovering the genetic architecture of broad antisocial behavior through a genome-wide association study meta-analysis. *Mol. Psychiatry* **27**, 4453–4463 (2022).
52. Rajagopal, V. M. *et al.* Differences in the genetic architecture of common and rare variants in childhood, persistent and late-diagnosed attention-deficit hyperactivity disorder. *Nat. Genet.* **54**, 1117–1124 (2022).
53. Dalvie, S. *et al.* Genomic influences on self-reported childhood maltreatment. *Transl. Psychiatry* **10**, 38 (2020).
54. Wendt, F. R. *et al.* Using phenotype risk scores to enhance gene discovery for generalized anxiety disorder and posttraumatic stress disorder. *Mol. Psychiatry* **27**, 2206–2215 (2022).
55. Stein, M. B. *et al.* Genome-wide association analyses of post-traumatic stress disorder and its symptom subdomains in the Million Veteran Program. *Nat. Genet.* **53**, 174–184 (2021).
56. Gupta, P. *et al.* A genome-wide investigation into the underlying genetic architecture of personality traits and overlap with psychopathology. *Nat. Hum. Behav.* 1–15 (2024) doi:10.1038/s41562-024-01951-3.
57. Schwaba, T. *et al.* Robust inference and widespread genetic correlates from a large-scale genetic association study of human personality. *bioRxiv* (2025) doi:10.1101/2025.05.16.648988.

58. Liu, P., Choi, Y.-K. & Qi, R. Z. NME7 is a functional component of the γ-tubulin ring complex. *Mol. Biol. Cell* **25**, 2017–2025 (2014).
59. Fiori, L. M. *et al.* Treatment-emergent and trajectory-based peripheral gene expression markers of antidepressant response. *Transl. Psychiatry* **11**, 439 (2021).
60. Li, Q. S., Wajs, E., Ochs-Ross, R., Singh, J. & Drevets, W. C. Genome-wide association study and polygenic risk score analysis of esketamine treatment response. *Sci. Rep.* **10**, 12649 (2020).
61. Goes, F. S. *et al.* Large-scale transcriptomic analyses of major depressive disorder reveal convergent dysregulation of synaptic pathways in excitatory neurons. *Nat. Commun.* **16**, 3981 (2025).
62. Martin, J. *et al.* Examining sex-differentiated genetic effects across neuropsychiatric and behavioral traits. *Biol. Psychiatry* **89**, 1127–1137 (2021).
63. 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).
64. Merikangas, A. K. & Almasy, L. Using the tools of genetic epidemiology to understand sex differences in neuropsychiatric disorders. *Genes Brain Behav.* **19**, e12660 (2020).
65. Fagerberg, L. *et al.* Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. *Mol. Cell. Proteomics* **13**, 397–406 (2014).
66. GTEx Consortium. The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science* **369**, 1318–1330 (2020).
67. O'Connell, K. S. *et al.* Genomics yields biological and phenotypic insights into bipolar disorder. *Nature* **639**, 968–975 (2025).
68. Del Matto, L. *et al.* Lithium and suicide prevention in mood disorders and in the general population: A systematic review. *Neurosci. Biobehav. Rev.* **116**, 142–153 (2020).
69. International Consortium on Lithium Genetics (ConLi+Gen) *et al.* Association of Polygenic Score for Schizophrenia and HLA Antigen and Inflammation Genes With Response to Lithium in Bipolar Affective Disorder: A Genome-Wide Association Study. *JAMA Psychiatry* **75**, 65–74 (2018).
70. Corbett, B. F. *et al.* Sphingosine-1-phosphate receptor 3 in the medial prefrontal cortex promotes stress resilience by reducing inflammatory processes. *Nat. Commun.* **10**, 3146 (2019).
71. Streit, F. *et al.* Borderline Personality Disorder and the Big Five: Molecular genetic analyses indicate shared genetic architecture with Neuroticism and Openness. *Research Square* (2021) doi:10.21203/rs.3.rs-958220/v1.
72. Schulze, A. *et al.* Evidence for a shared genetic contribution to loneliness and Borderline Personality Disorder. *medRxiv* (2023) doi:10.1101/2023.03.16.23286984.
73. Lynam, D. R. & Widiger, T. A. Using the five-factor model to represent the DSM-IV personality disorders: An expert consensus approach. *J. Abnorm. Psychol.* **110**, 401–412 (2001).
74. Samuel, D. B. & Widiger, T. A. A meta-analytic review of the relationships between the five-factor model and DSM-IV-TR personality disorders: a facet level analysis. *Clin. Psychol. Rev.* **28**, 1326–1342 (2008).
75. Trull, T. J. The Five-Factor Model of personality disorder and DSM-5. *J. Pers.* **80**, 1697–1720 (2012).
76. Czajkowski, N. *et al.* A Twin Study of Normative Personality and DSM-IV Personality Disorder Criterion Counts: Evidence for Separate Genetic Influences. *Am. J. Psychiatry*

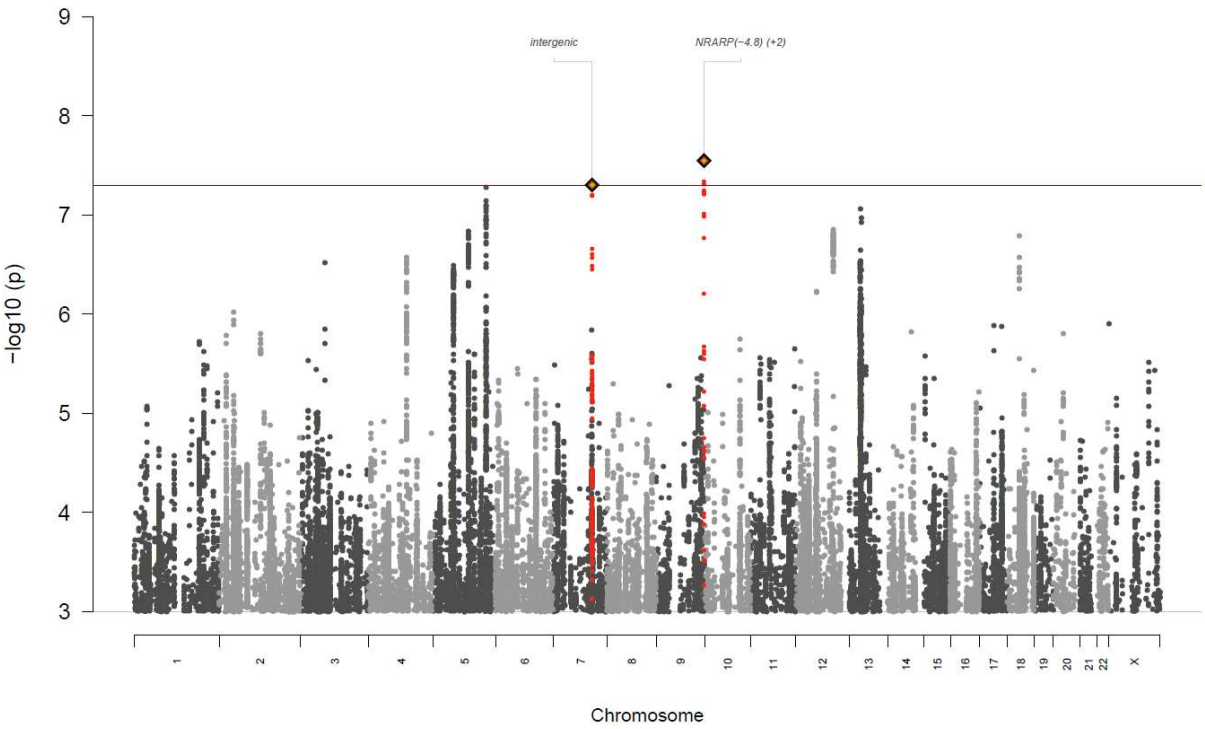
**175**, 649–656 (2018).
77. Kendler, K. S., Myers, J. & Reichborn-Kjennerud, T. Borderline personality disorder traits and their relationship with dimensions of normative personality: a web-based cohort and twin study. *Acta Psychiatr. Scand.* **123**, 349–359 (2011).
78. Garcia-Argibay, M. *et al.* Attention-Deficit/Hyperactivity Disorder and Major Depressive Disorder: Evidence From Multiple Genetically Informed Designs. *Biol. Psychiatry* (2023) doi:10.1016/j.biopsych.2023.07.017.
79. Bozzatello, P., Bellino, S., Bosia, M. & Rocca, P. Early Detection and Outcome in Borderline Personality Disorder. *Front. Psychiatry* **10**, 710 (2019).
80. Eaton, N. R. *et al.* Borderline personality disorder co-morbidity: relationship to the internalizing-externalizing structure of common mental disorders. *Psychol. Med.* **41**, 1041–1050 (2011).
81. Temes, C. M., Frankenburg, F. R., Fitzmaurice, G. M. & Zanarini, M. C. Deaths by Suicide and Other Causes Among Patients With Borderline Personality Disorder and Personality-Disordered Comparison Subjects Over 24 Years of Prospective Follow-Up. *J. Clin. Psychiatry* **80**, (2019).
82. Martin, J., Taylor, M. J. & Lichtenstein, P. Assessing the evidence for shared genetic risks across psychiatric disorders and traits. *Psychol. Med.* **48**, 1759–1774 (2018).
83. Kendall, K. M. *et al.* The genetic basis of major depression. *Psychol. Med.* **51**, 2217–2230 (2021).
84. Visscher, P. M. *et al.* 10 Years of GWAS Discovery: Biology, Function, and Translation. *Am. J. Hum. Genet.* **101**, 5–22 (2017).
85. Peterson, R. E. *et al.* Genome-wide Association Studies in Ancestrally Diverse Populations: Opportunities, Methods, Pitfalls, and Recommendations. *Cell* **179**, 589–603 (2019).
86. Fatumo, S. *et al.* A roadmap to increase diversity in genomic studies. *Nat. Med.* **28**, 243–250 (2022).
87. Wright, A. G. C. *et al.* Clarifying interpersonal heterogeneity in borderline personality disorder using latent mixture modeling. *J. Pers. Disord.* **27**, 125–143 (2013).
88. Bornovalova, M. A., Levy, R., Gratz, K. L. & Lejuez, C. W. Understanding the heterogeneity of BPD symptoms through latent class analysis: initial results and clinical correlates among inner-city substance users. *Psychol. Assess.* **22**, 233–245 (2010).
89. Koudys, J. W., Traynor, J. M., Rodrigo, A. H., Carcone, D. & Ruocco, A. C. The NIMH Research Domain Criteria (RDoC) Initiative and Its Implications for Research on Personality Disorder. *Curr. Psychiatry Rep.* **21**, 37 (2019).
90. Michelini, G., Palumbo, I. M., DeYoung, C. G., Latzman, R. D. & Kotov, R. Linking RDoC and HiTOP: A new interface for advancing psychiatric nosology and neuroscience. *Clin. Psychol. Rev.* **86**, 102025 (2021).
91. Shields, A. N., Giljen, M., España, R. A. & Tackett, J. L. The p factor and dimensional structural models of youth personality pathology and psychopathology. *Curr Opin Psychol* **37**, 21–25 (2021).
92. Lubke, G. H. *et al.* Genome-wide analyses of borderline personality features. *Mol. Psychiatry* **19**, 923–929 (2014).
93. Lam, M. *et al.* RICOPILI: Rapid Imputation for COnsortias PIpeLIne. *Bioinformatics* (2019) doi:10.1093/bioinformatics/btz633.
94. Chang, C. C. *et al.* Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**, 7 (2015).
95. Willer, C. J., Li, Y. & Abecasis, G. R. METAL: fast and efficient meta-analysis of

genomewide association scans. *Bioinformatics* **26**, 2190–2191 (2010).
96. 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).
97. Hübel, C. *et al.* Genetic correlations of psychiatric traits with body composition and glycemic traits are sex- and age-dependent. *Nat. Commun.* **10**, 5765 (2019).
98. 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).
99. Gaulton, A. *et al.* ChEMBL: a large-scale bioactivity database for drug discovery. *Nucleic Acids Res.* **40**, D1100–7 (2012).
100. Sakaue, S. & Okada, Y. GREP: genome for REPositioning drugs. *Bioinformatics* **35**, 3821–3823 (2019).
101. 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).
102. 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).
103. 1000 Genomes Project Consortium *et al.* A global reference for human genetic variation. *Nature* **526**, 68–74 (2015).
104. Campos, A. I. *et al.* Genetic aetiology of self-harm ideation and behaviour. *Sci. Rep.* **10**, 9713 (2020).

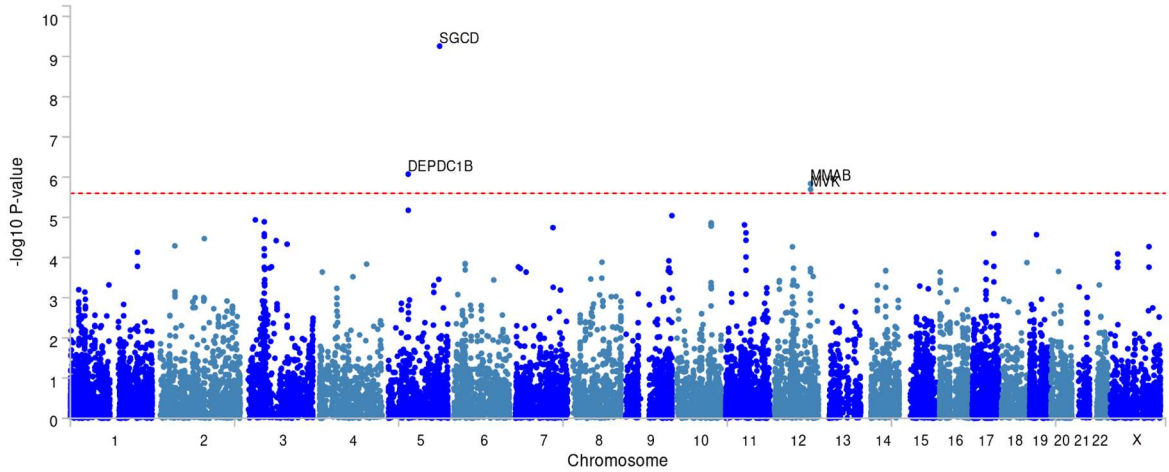
# Extended Data Figures



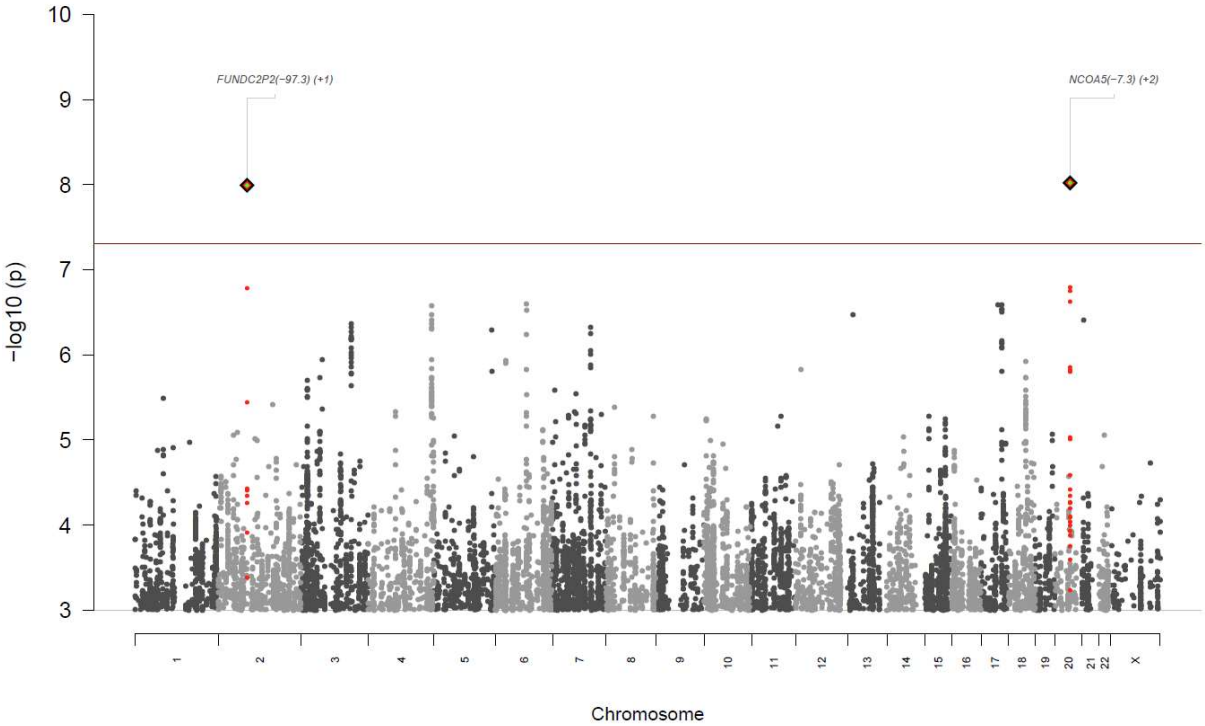

**Extended Data Fig. 1: Manhattan plot of the gene-based analysis GWAS meta-analysis of BPD ($N_{cases}$ = 12,339, $N_{controls}$ = 1,041,717).** The –log10 *p-value* for each gene is indicated on the y-axis (chromosomal position shown on the x-axis). The red line indicates significance after correction for multiple testing ($p$<0.05/19,843 = $2.5\times10^{-6}$). Gene names are given for significant genes.



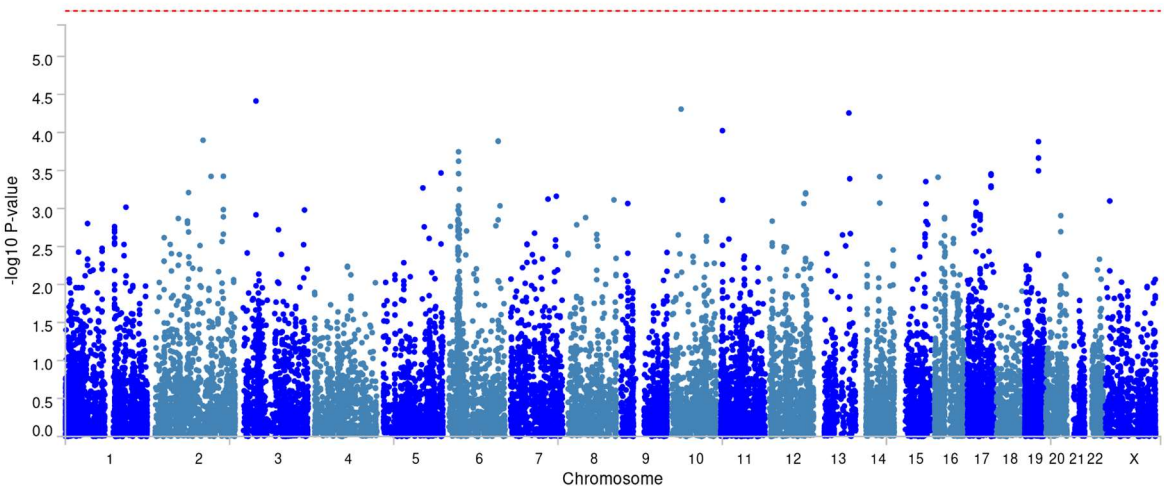

**Extended Data Fig. 2: Female subset - Manhattan plot of the GWAS meta-analysis of BPD ($N_{cases}$ = 10,025, $N_{controls}$ = 547,333).** The –log10 *p-value* for each single-nucleotide polymorphism (SNP) is indicated on the y-axis (chromosomal position shown on the x-axis). The red line indicates genome-wide significance ($p$<$5\times10^{-8}$). Index SNPs representing independent genome-wide significant associations are highlighted in diamonds, and SNPs in linkage disequilibrium with the index SNPs are highlighted in green.



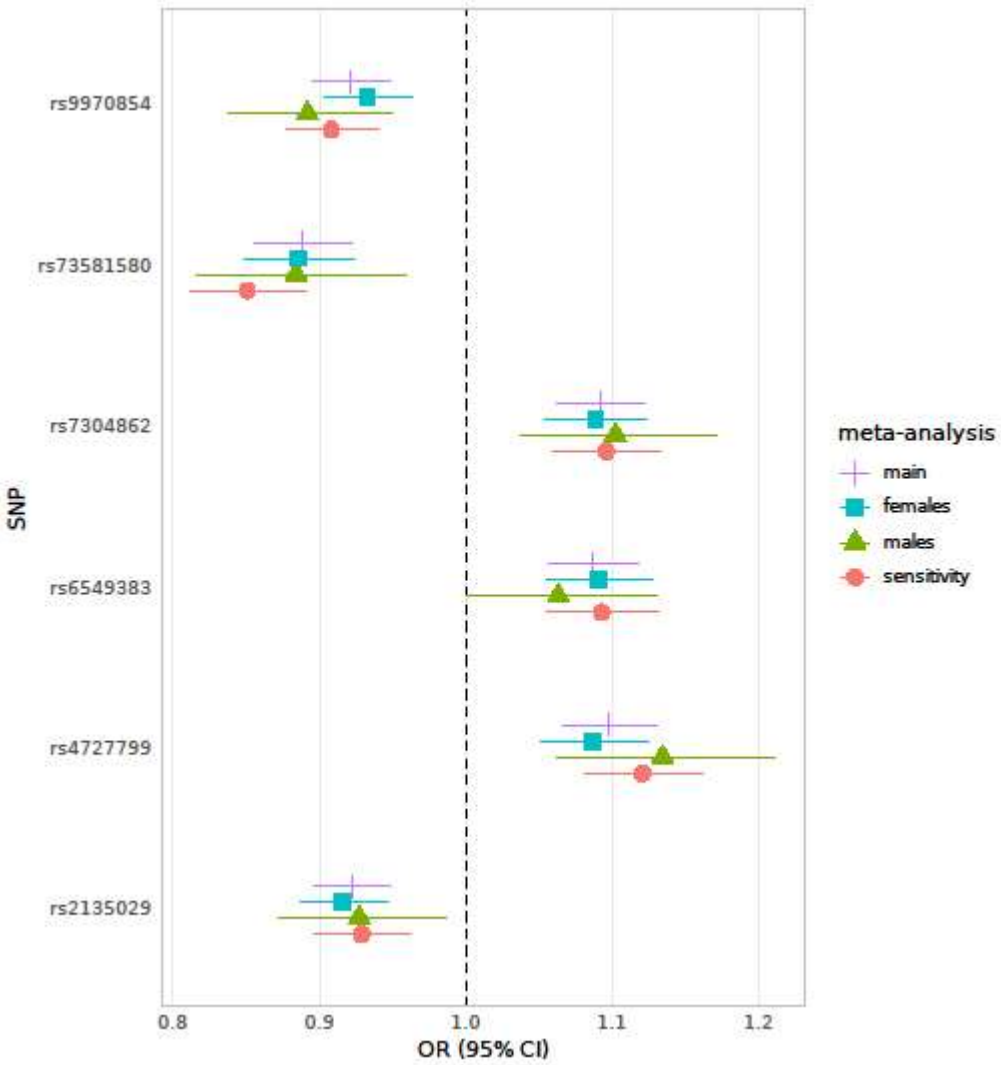

**Extended Data Fig. 3: Female subset - Manhattan plot of the gene-based analysis female GWAS meta-analysis of BPD ($N_{cases}$ = 10,025, $N_{controls}$ = 547,333).** The –log10 *p-value* for each gene is indicated on the y-axis (chromosomal position shown on the x-axis). The red line indicates significance after correction for multiple testing ($p$<0.05/19,844=2.5×10$^{-6}$). Gene names are given for significant genes.



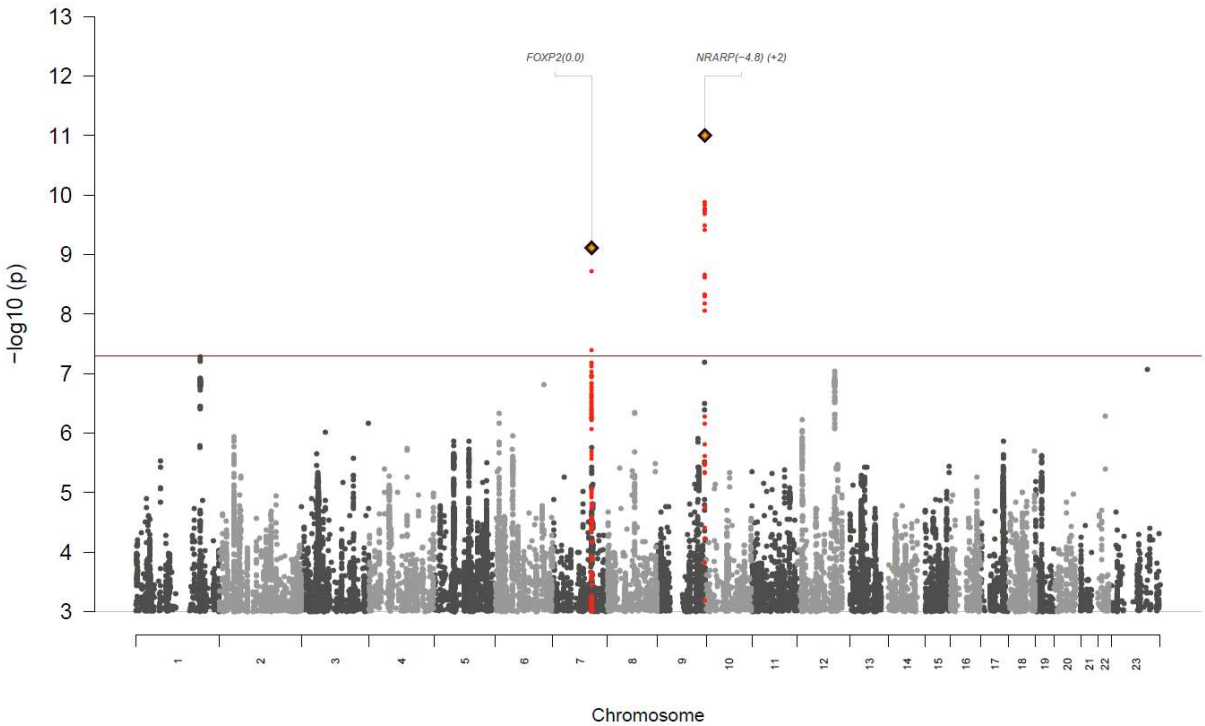

**Extended Data Fig. 4: Male subset - Manhattan plot of the GWAS meta-analysis of BPD ($N_{cases}$ = 2260, $N_{controls}$ = 485,444).** The –log10 *P-value* for each single-nucleotide polymorphism (SNP) is indicated on the y-axis (chromosomal position shown on the x-axis). The red line indicates genome-wide significance ($p$<5×10$^{-8}$). Index SNPs representing independent genome-wide significant associations are highlighted in diamonds, and SNPs in linkage disequilibrium with the index SNPs are highlighted in red.



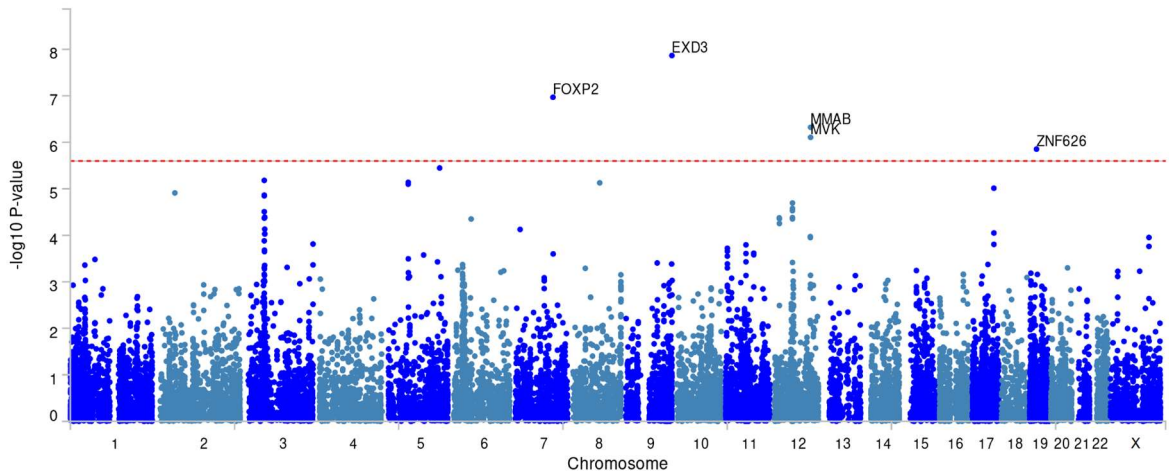

**Extended Data Fig. 5: Male subset - Manhattan plot of the gene-based analysis ($N_{cases}$ = 2260, $N_{controls}$ = 485,444).** The –log10 *p-value* for each gene is indicated on the y-axis (chromosomal position shown on the x-axis). The red line indicates significance after correction for multiple testing ($p$<0.05/19,856=2.5×$10^{-6}$). Gene names are given for significant genes.



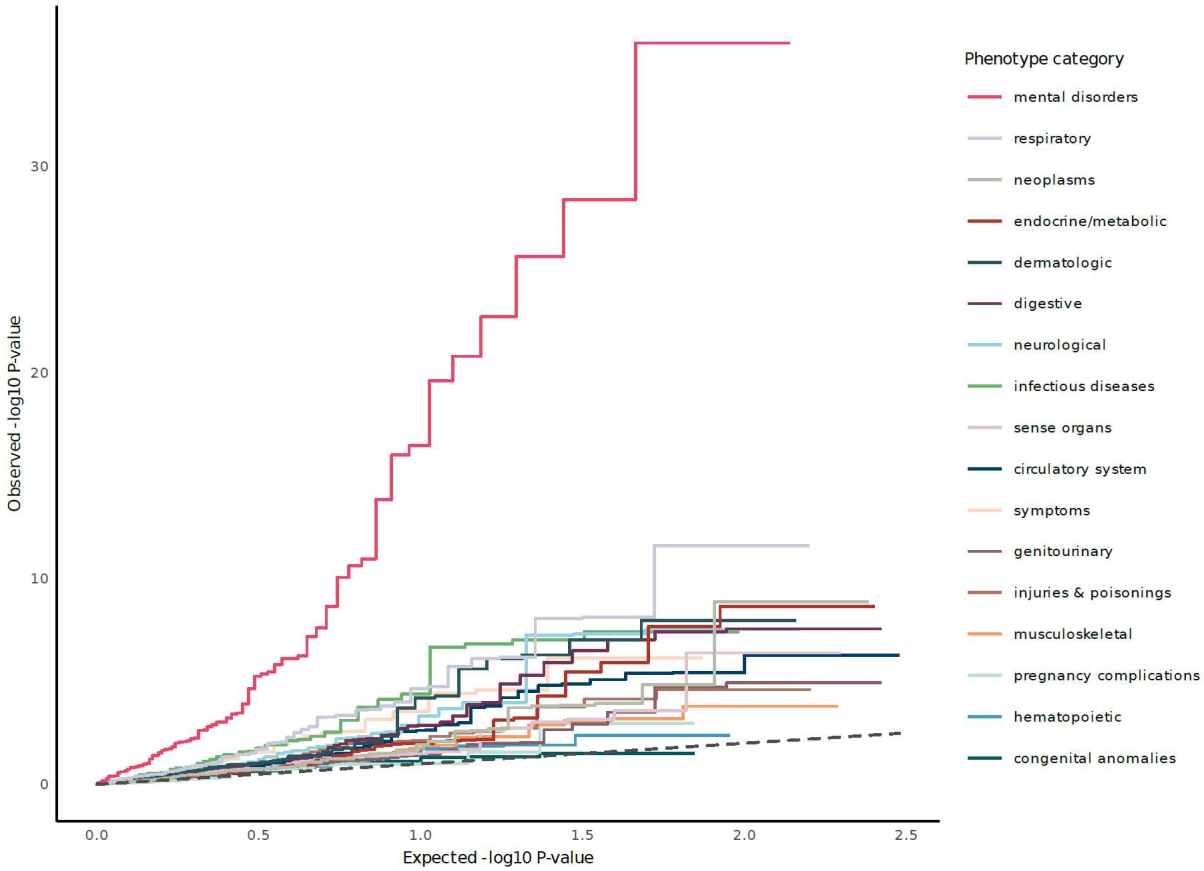

**Extended Data Fig. 6: Comparison of the effect sizes of the six GWAS lead SNPs.** Odds ratios (OR) are plotted for the main meta-analysis, the sex-stratified analyses, and the sensitivity analysis excluding cases with comorbidity of schizophrenia or bipolar disorder from the cohorts providing summary statistics. OR = odds ratio; 95% CI = 95% confidence interval; SNP = single nucleotide polymorphism



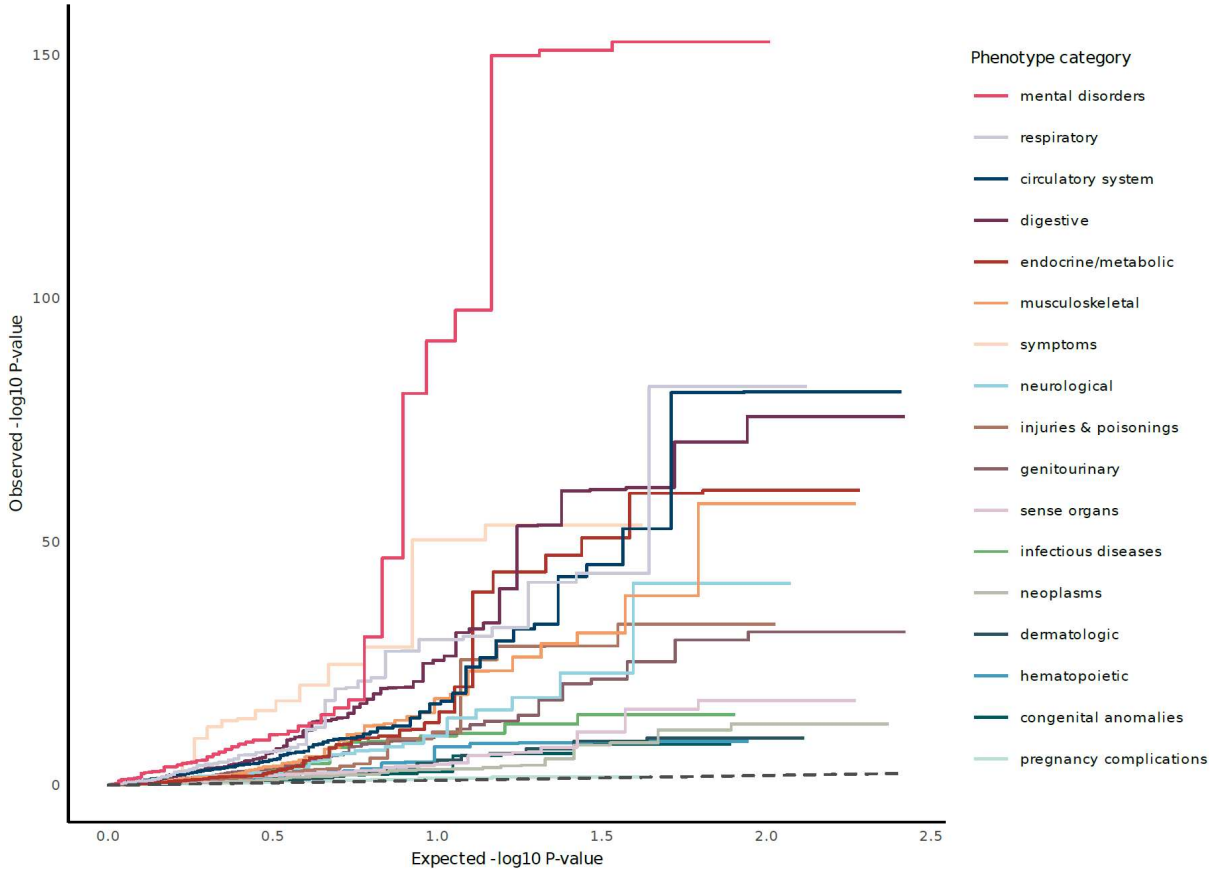

**Extended Data Fig. 7: Sensitivity analysis - Manhattan plot of the GWAS meta-analysis of BPD excluding subjects with schizophrenia or bipolar disorder ($N_{cases}$ = 8,618, $N_{controls}$ = 1,027,690).** The –log10 *p-value* for each single-nucleotide polymorphism (SNP) is indicated on the y-axis (chromosomal position shown on the x-axis). The red line indicates genome-wide significance ($p<5\times10^{-8}$). Index SNPs representing independent genome-wide significant associations are highlighted in diamonds, and SNPs in linkage disequilibrium with the index SNPs are highlighted in red.




**Extended Data Fig. 8: Sensitivity analysis - Manhattan plot of the gene-based analysis in the subset excluding subjects with schizophrenia or bipolar disorder - ($N_{cases}$ = 8,618, $N_{controls}$ = 1,027,690).** The –log10 *p-value* for each gene is indicated on the y-axis (chromosomal position shown on the x-axis). The red line indicates significance after correction for multiple testing ($p<0.05/19{,}842=2.5\times10^{-6}$). Gene names are given for significant genes.




**Extended Data Fig. 9: Quantile–Quantile plot of the PheWAS analysis in BioVU, stratified by phecode category (1,431 tested phecodes).** Observed –log10 *p-values* are shown on the y-axis, and expected –log10 *p-values* are shown on the x-axis.




**Extended Data Fig. 10: Quantile–Quantile plot of the PheWAS analysis in UKB, stratified by phecode category (1,250 tested phecodes).** Observed –log10 *p-values* are shown on the y-axis, and expected –log10 *p-values* are shown on the x-axis.