nature genetics

Article

https://doi.org/10.1038/s41588-026-02654-3

Genome-wide association analyses of borderline personality disorder identify 11 loci and highlight shared risk with mental and somatic disorders

Received: 21 November 2024

Accepted: 1 June 2026

Published online: 20 July 2026

Check for updates

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

Borderline personality disorder (BPD) is a severe mental health condition influenced by environmental risk factors (for example, interpersonal trauma) and genetic factors. We conducted the largest genome-wide association study (GWAS) meta-analysis of BPD so far, 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 9 risk genes in gene-based analyses. We observed a single-nucleotide polymorphism heritability of 17.3% and derived polygenic scores (PGS) that predicted 4.6% of the phenotypic variance in BPD on the liability scale. BPD showed the strongest positive genetic correlations with GWAS of post-traumatic stress disorder, depression, attention deficit hyperactivity disorder, antisocial behavior, and measures of suicide and self-harm. Phenome-wide analyses in Vanderbilt University Medical Center Biobank and UK Biobank using BPD-PGS confirmed these associations and also identified associations with other medical conditions, including obstructive pulmonary disease and diabetes. These analyses highlight BPD as a polygenic disorder, with the genetic risk showing substantial overlap with psychiatric and physical health conditions.

BPD is a mental disorder characterized by pervasive instability in emotions, interpersonal relationships and self-image as well as impulsive behavior[1,2] (for symptoms, see Supplementary Note). 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]. Although some psychotherapies are effective in treating BPD[8], no psychopharmacological treatments have been approved by the US Food and Drug Administration specifically for BPD[2,9].

In addition to environmental risk factors such as early interpersonal trauma[10,11], genetic factors contribute substantially 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 also with continuous traits; for example, 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, GWAS meta-analyses of genetic data from tens or hundreds of thousands of cases and controls have successfully identified hundreds of genetic risk loci[16]. By contrast, genetic research on BPD lags behind[17]. In the only GWAS of BPD conducted so far, encompassing 998 cases and 1,545 controls[18], no single genome-wide significant variants were identified, but significant

e-mail: fabian.streit@zi-mannheim.de; stephanie.witt@zi-mannheim.de

https://doi.org/10.1038/s41588-026-02654-3

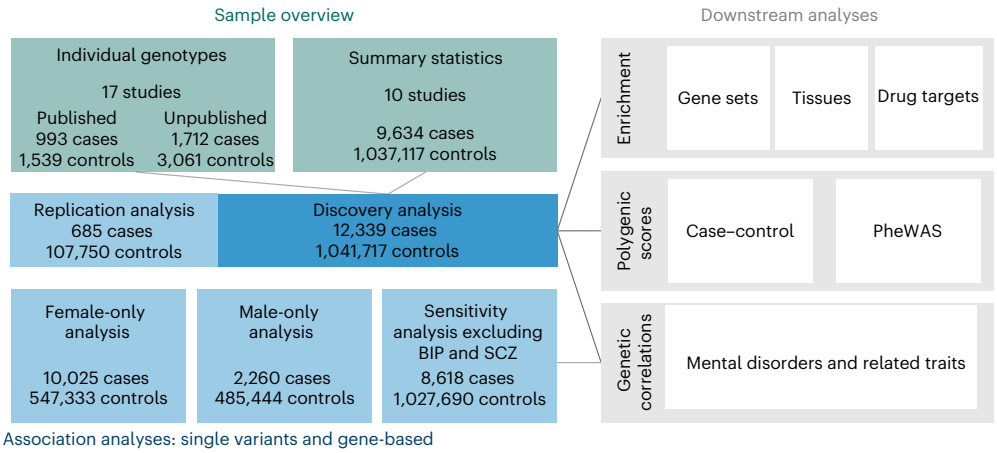


Fig. 1 | Schematic overview of the GWAS of BPD. A total of 17 studies providing individual-level data (analyzed in five datasets) and ten studies providing summary statistics were included in the discovery meta-analysis. Genetic associations were tested at the single-variant and gene level. Gene enrichment analyses were applied to test for enrichment in gene sets and tissue-specific gene expression data from 54 human tissues (GTEx), comprising 29 different ages and 11 general developmental stages (BrainSpan), single-cell-based cell types and drug-target gene sets. PGS were calculated to assess the prediction of BPD case–control status in the analyzed datasets with available individual-level data and to test the association of the genetic liability for BPD in phenome-wide association studies with phecodes in two biobanks (BioVU, UKB). Genetic correlations were calculated between the results of the GWAS of BPD and the GWAS of 50 disorders and traits of interest.

genetic correlations (r_g) of BPD were observed with bipolar disorder (BIP), schizophrenia (SCZ) and major depressive disorder[18]. Although these findings indicate the potential of using genetic approaches to investigate BPD, research based on those results is limited by 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 identify novel genetic risk loci for BPD to improve understanding of the underlying molecular mechanisms and to systematically assess the shared genetic risk between BPD and a broad range of related traits and disorders.

Results

GWAS

We performed a discovery GWAS meta-analysis, including 6,043,895 genetic markers in 12,339 BPD cases and 1,041,717 controls for individuals of European ancestry (Fig. 1). In total, 17 studies contributed individual-level data for a total of 2,705 cases meeting the fourth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (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). In these, BPD status was assessed using International Classification of Disease (ICD) codes, except for the Genetic Links to Anxiety and Depression (GLAD) study, which used a self-reported diagnosis. Details on ascertainment, along with inclusion and exclusion criteria, are documented in Supplementary Table 1. Sample sizes, basic demographics (sex and age), depression comorbidity and analysis details are provided in Supplementary Table 2 and Supplementary Note. Additionally, we performed meta-analyses stratified by sex (female only, male only) and a sensitivity analysis excluding individuals with a history of BIP or SCZ—conditions commonly excluded in BPD studies—based on summary statistics available from the studies providing summary statistics (Supplementary Tables 3–5). To assess replication, the lead single-nucleotide polymorphisms (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 $P < 1 \times 10^{-6}$ in the discovery GWAS were analyzed in a combined meta-analysis (n_{cases} = 13,024; $n_{controls}$ = 1,149,467). The discovery GWAS had over 80% power to detect effects of 1.1 for variants with an allele frequency between 0.3 and 0.7 (ref. 19; Supplementary Table 6 and Supplementary Fig. 1).

We observed a SNP heritability of 28.4% (95% confidence interval (CI), [24.9%, 31.8%]), corresponding to a SNP heritability of 17.3% (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 (s.e. = 0.0085), with an attenuation ratio of 0.11 (s.e. = 0.031), and we observed a genomic inflation factor (λ_{GC}) of 1.22 (λ_{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]; $P = 3.5 \times 10^{-21}$), which was significantly smaller than 1 (z = −1.76; P = 0.038).

SNP association analysis revealed six independent genome-wide significant loci ($P < 5 \times 10^{-8}$; Fig. 2, Table 1, Supplementary Figs. 2–14 and Supplementary Table 7), and no marker showed significant heterogeneity between studies after multiple testing correction (smallest heterogeneity P = 0.023). All 6 genome-wide significant lead SNPs (6 out of 6, P = 0.016; Table 1), and 77% of the lead SNPs associated with $P < 1 \times 10^{-6}$ (24 out of 31, P = 0.0017) showed effects in the same direction in the replication analysis. In the combined analysis, the 6 loci remained genome-wide significant, and 5 additional loci with $P < 5 \times 10^{-8}$ were observed (Fig. 2, Table 1 and Supplementary Figs. 2–24).

Gene-based association analyses using MAGMA (v.1.07)[21,22] implicated nine genes, including *CCDC71* and *DEPDC1B*, in addition to *SGCD*, *FOXP2*, *EXD3*, *MVK*, *MMAB*, *PCYT1B* and *BPTF* already implicated by the genome-wide SNP associations (Extended Data Fig. 1, Supplementary Fig. 25 and Supplementary Table 8).

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 9). 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 study participants

https://doi.org/10.1038/s41588-026-02654-3

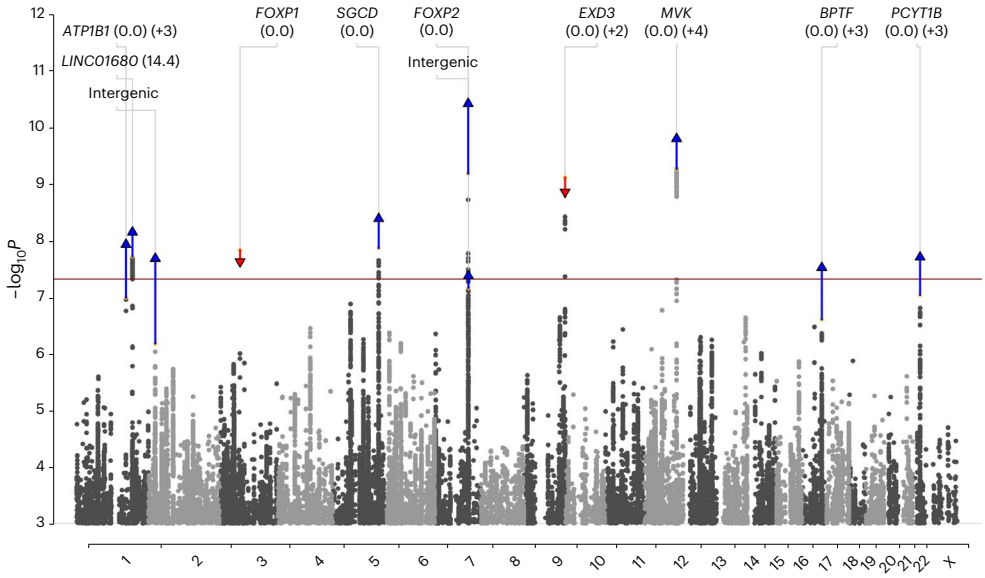


Fig. 2 | Manhattan plot of the GWAS of BPD. The two-sided $-\log_{10}P$ value for each SNP in the discovery GWAS (inverse-variance-weighted meta-analysis; $n_{cases} = 12{,}339$; $n_{controls} = 1{,}041{,}717$) is indicated on the *y* axis (chromosomal position shown on the *x* axis). The red horizontal line indicates genome-wide significance ($P < 5 \times 10^{-8}$). 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.

($n_{cases} = 10{,}025$; $n_{controls} = 547{,}333$) (Supplementary Tables 3 and 4). 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 a 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 ($z = -3.32$; $P = 0.00090$). There was a high genetic correlation between the two analyses ($r_g = 0.80$; 95% CI, [0.53, 1.07]; $P = 4.7 \times 10^{-9}$), which did not differ from 1 ($z = -1.60$; $P = 0.055$).

The female-only analysis identified one genome-wide significant risk locus on chromosome 9 (rs73581580; odds ratio (OR), 0.89; 95% CI, [0.85, 0.92]; $P = 2.83 \times 10^{-8}$; locus 5 identified in the main analysis (*EXD3*) and a second locus on chromosome 7 (rs10227454; OR, 0.87; 95% CI, [0.83, 0.92]; $P = 4.99 \times 10^{-8}$), ~1 Mb from locus 4 (*FOXP2*, $r^2 = 0.005$, $D' = 0.25$; Extended Data Fig. 2, Supplementary Figs. 26–30 and Supplementary Table 10). The gene-based analysis identified genome-wide associations for *DEPDC1B*, *SGCD*, *MVK* and *MMAB*, all significant in the main analysis (Extended Data Fig. 3 and Supplementary Fig. 31).

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; OR, 0.68; 95% CI, [0.59, 0.77]; $P = 1.02 \times 10^{-8}$) and one on chromosome 20 (rs6032676; OR, 0.83; 95% CI, [0.77, 0.88]; $P = 9.55 \times 10^{-9}$) (Extended Data Fig. 4, Supplementary Figs. 32–36 and Supplementary Table 11). The gene-based analysis identified no significant genes (Extended Data Fig. 5 and Supplementary Fig. 37).

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 (Extended Data Fig. 6 and Supplementary Table 12). The nine genes significant in the gene-based analysis of the main GWAS all showed nominal significance in the female-only analysis (all $P_{female} < 8.15 \times 10^{-5}$), whereas *CCDC71*, *DEPDC1B* and *SGCD* were not significant in the male-only analysis (Supplementary Table 13).

Sensitivity analysis excluding BIP and SCZ

In the sensitivity analysis excluding individuals with BIP or SCZ ($n_{cases} = 8{,}618$; $n_{controls} = 1{,}027{,}690$), 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. 38–42 and Supplementary Table 14). All six lead SNPs of the main GWAS showed comparable effects (all $P_{sensitivity} < 3.32 \times 10^{-5}$; Extended Data Fig. 6 and Supplementary Table 12). Of the nine genes that reached significance in the gene-based analysis of the primary analysis, all remained nominally significant in the sensitivity analysis ($P_{sensitivity} < 0.0021$; Supplementary Table 13), with *FOXP2*, *EXD3*, *MMAB* and *MVK* reaching genome-wide significance, and *ZNF626* being the only additional genome-wide significant gene (Extended Data Fig. 8 and Supplementary Fig. 43).

Enrichment analysis

Enrichment of gene sets and tissue expression using Genotype-Tissue Expression (GTEx; v.8) data for 54 tissue types[24] and expression data from 29 different ages and 11 general developmental stages (BrainSpan)[25] was tested with MAGMA (v.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] (29 genes; standardized effect size $\beta = 0.03$; $P = 1.91 \times 10^{-6}$; $P_{adj} = 0.018$; Supplementary Table 15). 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. 44 and Supplementary Table 16). The analysis using the two BrainSpan datasets 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 Figs. 45 and 46 and Supplementary Tables 17 and 18).

Analyses based on Human Brain Atlas single-nucleus RNA sequencing data[27] showed SNP-h^2 enrichment using stratified LDSC[28,29] in 2 out of the 31 tested superclusters ('medium spiny neurons': enrichment = 5.99, $z = 3.58$, $P = 0.00017$, $P_{adj} = 0.0054$; and '*LAMP5-LHX6* and

https://doi.org/10.1038/s41588-026-02654-3

Table 1 | Lead 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	s.e.	P	OR repl.	P repl.	OR comb.	s.e. comb.	P comb.	Mapped genes
Comb. 1	1	rs2143288	169084340	169082360..169084340	G/A	0.858	0.868	0.89	0.022	1.06×10^{-7}	0.86	0.035	0.89	0.021	**1.23×10^{-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	**1.99×10^{-8}**	0.93	0.183	0.92	0.014	**7.42×10^{-9}**	*LINC01680* (14.4)
Comb. 2	2	rs12466671	22583435	22548735..22606235	A/G	0.587	0.599	0.93	0.015	6.80×10^{-7}	0.85	0.002	0.92	0.014	**2.16×10^{-8}**	-
2	3	rs6549383	71266885	71261185..71267720	T/C	0.503	0.503	1.09	0.015	**1.53×10^{-8}**	1.02	0.664	1.08	0.014	**2.45×10^{-8}**	*FOXP1* (0.0)
3	5	rs2135029	155856538	155743538..155868438	A/G	0.594	0.599	0.92	0.014	**1.41×10^{-8}**	0.92	0.141	0.92	0.014	**4.30×10^{-9}**	**SGCD** (0.0)
4	7	rs4727799	114110568	114021868..114224568	T/C	0.663	0.67	1.10	0.015	**6.92×10^{-10}**	1.15	0.015	1.10	0.015	**4.05×10^{-11}**	***FOXP2*** (0.0)
Comb. 3	7	rs10953781	114974695	114963595..115112695	C/T	0.528	0.534	1.08	0.014	7.13×10^{-8}	1.06	0.313	1.08	0.014	**4.38×10^{-8}**	-
5	9	rs73581580	140251458	140242728..140278158	G/A	0.836	0.839	0.89	0.019	**8.09×10^{-10}**	0.98	0.798	0.89	0.019	**1.45×10^{-9}**	*NRARP* (−4.8), ***EXD3*** (0.0), *NOXA1* (16.4)
6	12	rs73044862	109988891	109851891..110050091	G/T	0.529	0.513	1.09	0.014	**5.71×10^{-10}**	1.09	0.114	1.09	0.014	**1.68×10^{-10}**	*MYO1H* (-52.7), *KCTD10* (−23.5), *UBE3B* (0.0), ***MMAB*** (0.0), ***MVK*** (0.0)
Comb. 4	17	rs72100027	65977053	65825053..66098053	T/C	0.77	0.782	0.92	0.017	2.54×10^{-7}	0.87	0.031	0.91	0.016	**3.12×10^{-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	9.47×10^{-8}	0.92	0.037	0.93	0.013	**2.05×10^{-8}**	*PDK3* (0), *PCYT1B* (0), *PCYT1B-AS1* (+5.0), *POLA1* (+48.8)

Results of the variance-weighted meta-analysis are shown (two-sided significance test). 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; s.e., standard error. Mapped genes, genes within r^2 of 0.6 and 50 kb of the index SNP (distance to the 50 kb 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 ($P<5\times10^{-8}$). OR, s.e. and *P* values are reported for the main analysis (n_{cases} = 12,339; $n_{controls}$ = 1,041,717), the replication analysis (n_{cases} = 685; total $n_{controls}$ = 107,750) and the combined meta-analysis (n_{cases} = 13,024; $n_{controls}$ = 1,149,467) of the two datasets.

Chandelier cells': enrichment = 5.50, $z = 3.53$, $P = 0.00021$, $P_{adj} = 0.0064$; Supplementary Fig. 47 and Supplementary Table 19).

Drug-target analysis

Of the 16 genes prioritized from the discovery GWAS meta-analysis through proximity to the six GWAS loci or in the gene-based test, none were highlighted as drug targets by Open Targets, although several showed potential tractability (Supplementary Fig. 48). 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, *MVK* 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 or somatic conditions (Supplementary Table 20; all $P_{adj} > 0.18$).

PGS

PGS for BPD were calculated using PRS-CS[31] in the datasets for which 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 datasets (Methods). For comparison, PGS were additionally calculated based on the 2017 BPD GWAS[18]. The explained variance was converted to the liability scale (population prevalence, 1.5%)[3].

PGS explained a weighted average of 4.6% (area under the receiver operator characteristic curve (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 and Supplementary Table 21). 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 ($P = 2.09\times10^{-25}$; AUC = 61.4%). The OR for BPD case–control status comparing the highest PGS decile to the lowest decile was 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.

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) and cognition. After Bonferroni correction for multiple testing ($\alpha = 0.05/50 = 0.001$), BPD showed significant genetic correlations with 43 out of the 50 tested phenotypes. Among the psychiatric disorders, BPD showed the strongest correlations with post-traumatic stress disorder (PTSD) ($r_g = 0.77$; 95% CI, [0.61–0.93]; $P = 2.97\times10^{-21}$), depression ($r_g = 0.74$; 95% CI, [0.69–0.79]; $P = 6.06\times10^{-163}$) and attention deficit hyperactivity disorder (ADHD) ($r_g = 0.67$; 95% CI, [0.59–0.75]; $P = 7.20\times10^{-63}$). 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 for body mass index (Fig. 4 and Supplementary Table 22).

https://doi.org/10.1038/s41588-026-02654-3

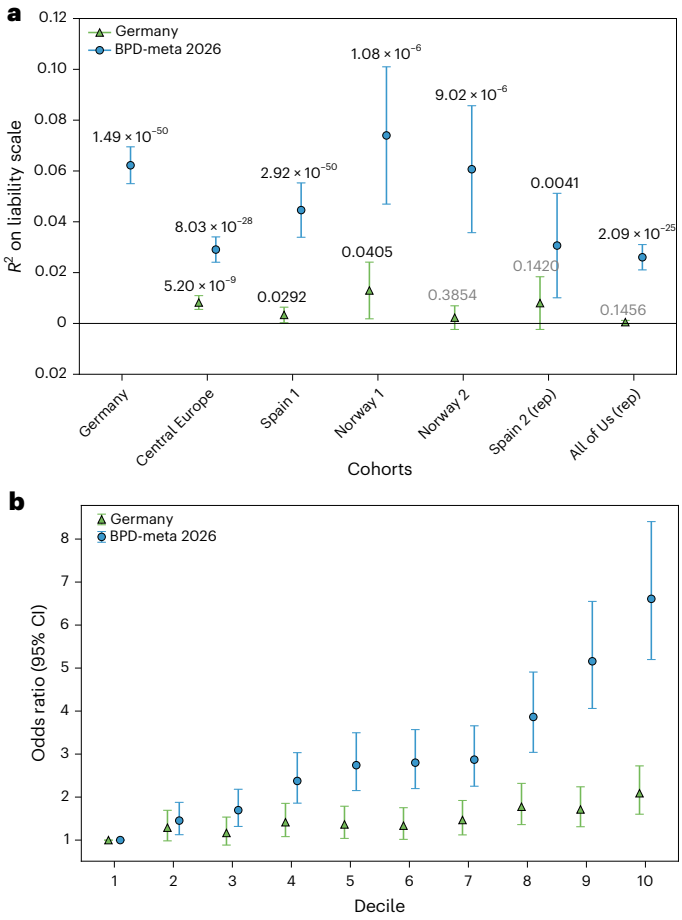


Fig. 3 | PGS analysis. a, The proportion of variance in case–control status explained by the PGS on the liability scale (*y* axis; Liability R^2). **b**, OR for BPD by PGS deciles, with decile 1 as reference. Leave-one-out PGS were calculated for the datasets with available individual-level genotype, including Germany (n_{cases} = 993; $n_{controls}$ = 1,539), Central Europe (n_{cases} = 1,285; $n_{controls}$ = 1,332), Spain 1 (n_{cases} = 306; $n_{controls}$ = 759), Norway 1 (n_{cases} = 57; $n_{controls}$ = 700) and Norway 2 (n_{cases} = 64; n = 270 controls), and for two independent replication (rep) datasets: Spain 2 (replication; n_{cases} = 46; $n_{controls}$ = 435) and All of Us (replication; n_{cases} = 639 cases; $n_{controls}$ = 107,315) (Supplementary Tables 1 and 2) using PRS-CS. For each prediction, the respective dataset was excluded from the used discovery GWAS meta-analysis (n_{cases} = 12,339; $n_{controls}$ = 1,041,717), while the full sample size was used for the two replication datasets (blue circles, 2026 meta-GWAS). 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; n_{cases} = 993; $n_{controls}$ = 1,539)[18]. In **a**, two-sided *P* values shown above each upper error bar were obtained from logistic regression assessing the associations between PGS and BPD case–control status. Nagelkerke's pseudo-R^2 was computed by comparing the full model (including PGS and ancestry principal components) with a reduced (covariate-only) model. The explained variance was converted to the liability scale of the population, assuming a lifetime disease risk of 1.5%. Error bars in **a**, s.e.m. of the proportion of variance in liability (Liability R^2) explained by the PGS. The analysis presented in **b** excluded the All of Us dataset, as its individual-level data could not be merged with the other datasets owing to data protection regulations. Error bars in **b**, 95% CI.

The sensitivity meta-analysis excluding individuals with BIP or SCZ showed comparable genetic correlations (Supplementary Fig. 49 and Supplementary Table 23) with the 50 disorders and traits. Notably, lower but still substantial genetic correlations were observed with BIP (r_g = 0.35; 95% CI, [0.28, 0.42] versus r_g = 0.47; 95% CI, [0.41–0.53]) and SCZ (r_g = 0.36; 95% CI, [0.31,0.41] versus r_g = 0.48; 95% CI, [0.43, 0.53]). The genetic correlation of the main meta-analysis and the sensitivity analysis showed r_g = 1.00 (95% CI, [0.98, 1.02]; $P < 1 \times 10^{-308}$).

Phenome-wide association studies

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'; that is, medical phenotypes based on ICD diagnoses documented in the electronic health records (EHR), using data from the Vanderbilt University Medical Center Biobank (BioVU)[32] (66,325 individuals; 214 with the phecode 301.20: 'antisocial/borderline personality disorder') and Hospital Episode Statistics in the UK Biobank (UKB) (316,635 individuals; 211 with the phecode 301.20) (Fig. 5).

PheWAS in BioVU. In the PheWAS analysis in BioVU, 69 out of the 1,431 tested diagnoses showed an association with BPD-PGS after Bonferroni correction ($P_{adj} < 0.05$; Supplementary Table 24). 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, PTSD and suicidal ideation or attempt. Associations were also observed in other categories, including neurological (for example, epilepsy) and somatic disorders (for example, type 2 diabetes, chronic airway obstruction, hypertension) and pain disorders and symptoms.

PheWAS in UKB. In the PheWAS analysis in the UKB, 317 out of the 1,250 tested diagnoses showed an association with BPD-PGS ($P_{adj} < 0.05$; Supplementary Table 25). 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 (for example, chronic airway obstruction), digestive (for example, esophagus, gastroesophageal reflux disease), circulatory, endocrine/metabolic and neurological diagnoses, as well as 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; 95% CI, [1.24, 1.63]; $P = 7.61 \times 10^{-7}$; OR_{UKB} = 1.54; 95% CI, [1.34, 1.76]; $P = 5.82 \times 10^{-10}$).

Discussion

The present GWAS, including 12,339 BPD cases and 1,041,717 controls, represents a major advance in BPD genetics, demonstrating substantial SNP heritability, identifying genome-wide significant risk loci, revealing pathway and tissue enrichment and considerably improving the predictive value of derived PGS. To our knowledge, this study represents the first sex-stratified and X-chromosomal GWAS of BPD, as well as the first systematic investigation of shared genetic risk for 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. Building on our previous study, which included only clinical studies that explicitly recruited patients with BPD[18], we extended the approach to also include biobank and cohort studies in which diagnoses were primarily derived from electronic health records linked to genetic data. The high genetic correlation indicates that 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

https://doi.org/10.1038/s41588-026-02654-3

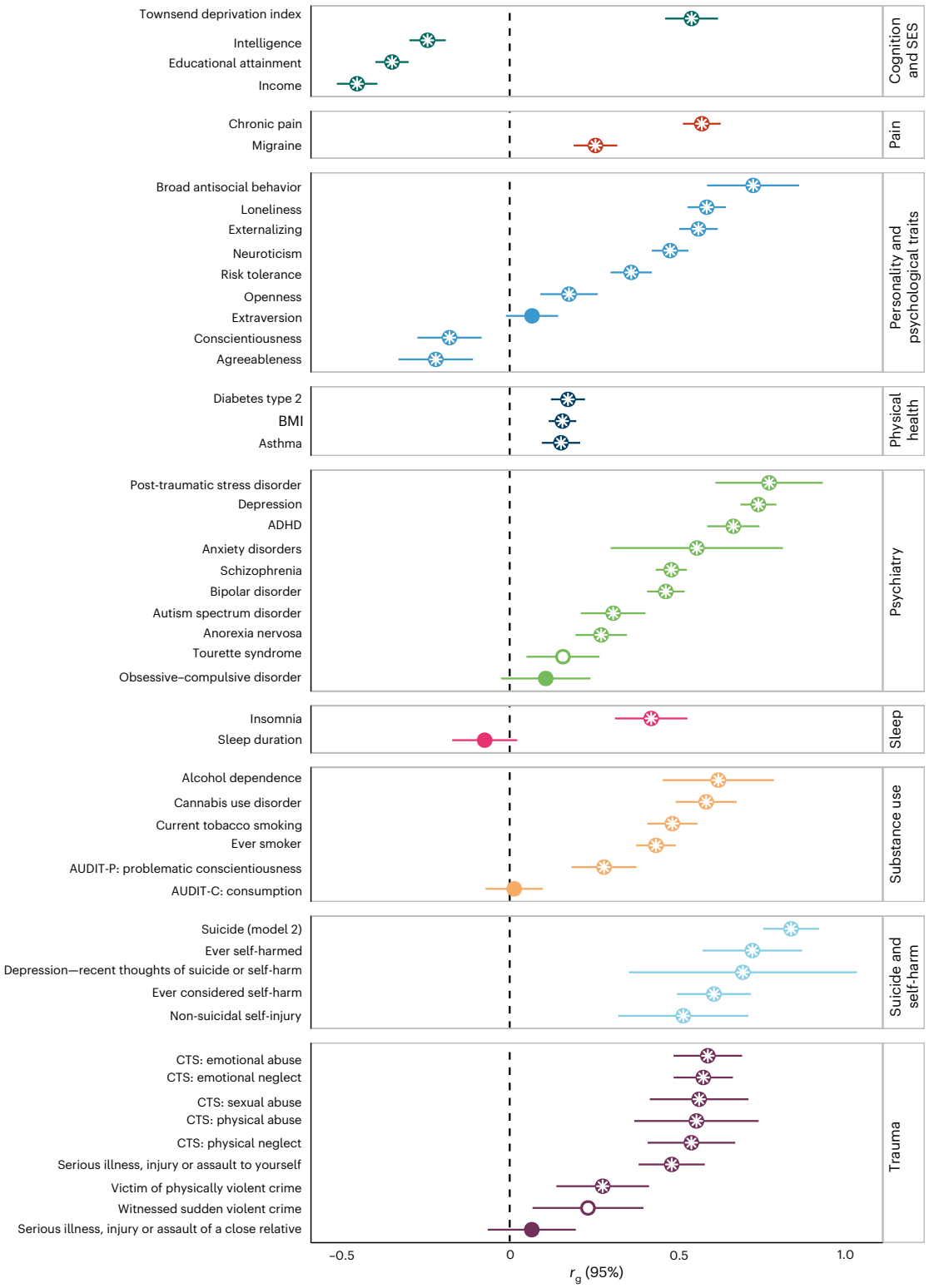


Fig. 4 | Genetic correlations of BPD with other phenotypes. Genetic correlations of BPD (total n_{cases} = 12,339; $n_{controls}$ = 1,041,717) with 50 disorders and traits: within each group, disorders and traits are sorted by their genetic correlation. Data are presented as point estimates (genetic correlations) with error bars indicating 95% CIs. Two-sided significance of the genetic correlation is indicated by a white dot ($P < 0.05$) or a white star ($P < 0.001$; 0.05/50 tested correlations). Source and sample size of the used GWAS, as well as the exact P values of the tested genetic correlations, are listed in Supplementary Table 22. SES, socioeconomic status. AUDIT, Alcohol Use Disorders Identification Test; BMI, body mass index; CTS, Childhood Trauma Screener.

other psychiatric disorders, in which SNP heritability often accounts for approximately one-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 (1.03) and attenuation ratio of 11% suggest that the association is largely driven by the polygenic signal and not by confounding.

https://doi.org/10.1038/s41588-026-02654-3

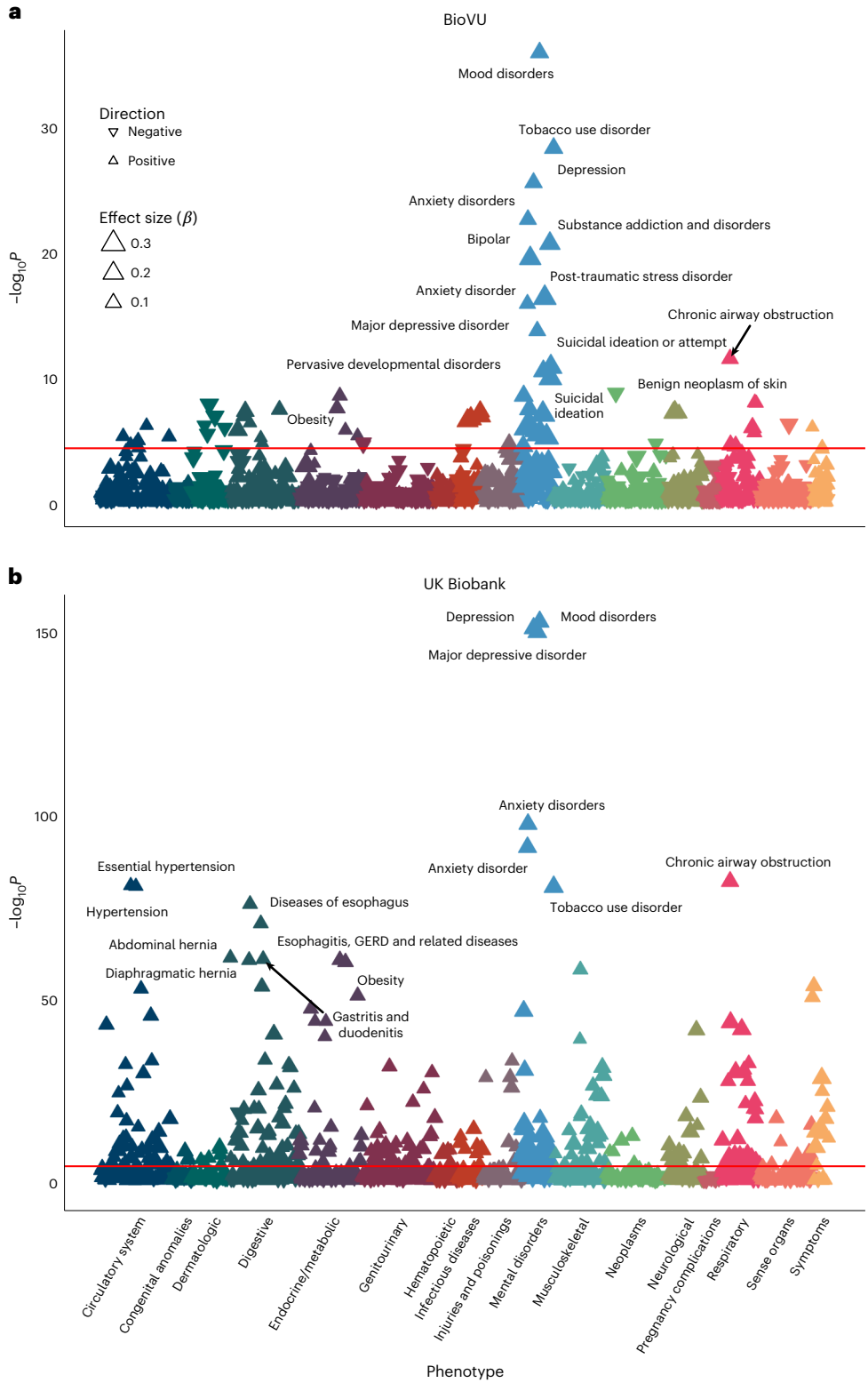


Fig. 5 | PheWAS of BPD-PGS. a,**b**, Association of BPD-PGS with 1,431 tested phecodes in BioVU (**a**) and 1,250 Hospital Episode Statistics phecodes in UKB (**b**). The two-sided statistical significance ($-\log_{10}P$) from the logistic regression models is plotted on the *y* axis, and diagnoses are grouped by category. The red line indicates the Bonferroni-corrected significance threshold for the tested phecodes (BioVU: $P < 3.49 \times 10^{-5}$ (0.05/1,431); UKB: $P < 4.00 \times 10^{-5}$ (0.05/1,250)). 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 (β).

In the context of identifying genetic risk variants and genes, the present study is the first to identify 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 *SGCD* 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 central nervous system in both neurons and astrocytes[41,42] and may be involved in GABAergic neurotransmission[41] and cerebrovascular function[43].

Among the genes mapped by proximity to the lead SNPs or in the MAGMA gene-based analysis, the highest PoPS scores were observed for *SGCD*, *FOXP2*, *FOXP1* and *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]. *FOXP2* 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 has 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 showing significant brain GWAS and quantitative trait locus associations in a transcriptomic analysis of major depressive disorder[61].

We present the first GWAS of BPD 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 substantially higher SNP heritability in females, which has also been described for PTSD and depression, the two mental disorders showing the strongest genetic correlation with BPD[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, as indicated by a SNP-heritability z-score of <4 (ref. 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, we comprehensively analyzed genetic variation on the X chromosome and identified the gene encoding choline-phosphate cytidylyltransferase B (*PCYT1B*) as significantly associated with BPD in the gene-based analysis, as well as an associated SNP in the gene with genome-wide significance in the combined analysis. *PCYT1B* is expressed in the brain[65,66] but has not previously been reported to be associated with mental health phenotypes. However, many previous GWAS analyses have not reported X chromosome data[64].

We also used 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' and 'LAMP5-LHX6 and Chandelier cells'. Both clusters have previously been demonstrated to be enriched in mental health phenotypes in a systematic analysis, namely in SCZ, IQ, educational attainment and neuroticism[29], with 'medium spiny neurons' additionally showing enrichment in major depressive disorder[29] and the 'Lamp5-LHX6 and Chandelier cell' cluster being significantly enriched in the most recent BIP GWAS[67]. Together, these findings suggest that specific neuronal cell types with established relevance for psychiatric disorders may also have 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. 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 shown 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 medications for neurological disorders and pain; however, 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 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[14,76]. 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[5,7,77]. The associations with suicide and self-harm are consistent with the clinical presentation of BPD[78], the correlations with trauma phenotypes highlight the role these experiences have in BPD[10] and the strong correlations of BPD with disorders from the internalizing–externalizing spectrum disorders suggest that BPD risk is influenced by the liability for both dimensions[78,79].

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 and 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 cases has been previously described[7]. The relatively small number of BPD cases in the PheWAS target samples makes it probable that these associations are driven by shared genetic risk and not solely by comorbidities of BPD cases in the target samples. These results provide a promising starting point for investigating shared genetic risks between BPD and somatic health, which is particularly relevant given that a substantial portion of the reduced life expectancy in individuals with BPD is attributable to physical health conditions[80]. 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 disentangle the extent to which comorbidities, overlapping symptoms and potential diagnostic misclassifications might have influenced our results[81]. Although 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 study participants 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 individuals with either a diagnosis of BIP or SCZ, which are common exclusion criteria in 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 BIP or SCZ comorbidity substantially biased the GWAS hits. Comparing genetic correlations between the two iterations of the BPD GWAS and the GWAS of BIP and SCZ, we observe attenuated but still substantial genetic correlations 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, there were several limitations. First, the sample size is relatively small compared with studies involving other mental disorders[16], and statistical power was limited, especially for variants with lower frequency. The inclusion of additional samples may lead to a substantial increase in the number of loci identified[82,83]. Furthermore, expanding the analyses to non-European ancestries is necessary to improve our understanding of the underlying genetic architecture and to facilitate the generalizability of the results[84,85]. Additionally, in this study, we investigated BPD as a categorical diagnosis, as it is 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; for example, by examining individual symptoms or symptom clusters[86,87]. In addition, functional dimensions such as RDoC[88] and classification systems such as HiTop[89,90], as well as borderline personality traits or symptom dimensions[2,17,91] should be incorporated.

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

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41588-026-02654-3.

References

1. Bohus, M. et al. Borderline personality disorder. *Lancet* **398**, 1528–1540 (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. in *Personalized Psychiatry* (ed. Baune, B. T.) 227–237 (Elsevier, 2020); https://doi.org/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–D679 (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. *Nat. Commun.* **16**, 395 (2025).
30. Bell, N., Uffelmann, E., van Walree, E., de Leeuw, C. & Posthuma, D. Using genome-wide association results to identify drug repurposing candidates. Preprint at *medRxiv* https://doi.org/10.1101/2022.09.06.22279660 (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* **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**, 716 (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.* **8**, 2235–2249 (2024).
57. Schwaba, T. et al. Robust inference and widespread genetic correlates from a large-scale genetic association study of human personality. Preprint at *bioRxiv* https://doi.org/10.1101/2025.05.16.648988 (2025).
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. *Trans. Psychiatry* **12**, 153 (2022).
72. Schulze, A. et al. Evidence for a shared genetic contribution to loneliness and borderline personality disorder. *Trans. Psychiatry* **13**, 398 (2023).
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. Garcia-Argibay, M. et al. Attention-deficit/hyperactivity disorder and major depressive disorder: evidence from multiple genetically informed designs. *Biol. Psychiatry* **95**, 444–452 (2024).
78. Bozzatello, P., Bellino, S., Bosia, M. & Rocca, P. Early detection and outcome in borderline personality disorder. *Front. Psychiatry* **10**, 710 (2019).
79. 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).
80. 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**, 18m12436 (2019).
81. 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).
82. Kendall, K. M. et al. The genetic basis of major depression. *Psychol. Med.* **51**, 2217–2230 (2021).
83. Visscher, P. M. et al. 10 Years of GWAS discovery: biology, function, and translation. *Am. J. Hum. Genet.* **101**, 5–22 (2017).
84. Peterson, R. E. et al. Genome-wide association studies in ancestrally diverse populations: opportunities, methods, pitfalls, and recommendations. *Cell* **179**, 589–603 (2019).
85. Fatumo, S. et al. A roadmap to increase diversity in genomic studies. *Nat. Med.* **28**, 243–250 (2022).
86. Wright, A. G. C. et al. Clarifying interpersonal heterogeneity in borderline personality disorder using latent mixture modeling. *J. Pers. Disord.* **27**, 125–143 (2013).
87. 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).
88. 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).
89. 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).
90. 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).
91. Lubke, G. H. et al. Genome-wide analyses of borderline personality features. *Mol. Psychiatry* **19**, 923–929 (2014).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2026

Fabian Streit[1,2,3,4,137] ✉, Swapnil Awasthi[5,6,137], Alisha S. M. 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 A. N. Christoffersen[30,31], Lucía Colodro-Conde[32,33], Brandon J. Coombes[24], Elizabeth C. Corfield[21,34,35], Norbert Dahmen[36], Maria Didriksen[37,38], Khoa M. Dinh[38,39], Srdjan Djurovic[40,41], Joseph Dowsett[38], Ole Kristian Drange[41,42], Helene Dukal[3], Susanne Edelmann[43,44], Christian Erikstrup[7,39], 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[54,55], R. Patrick Harper[56], Alexandra Havdahl[21,23,34], Urs Heilbronner[57], Stefan Herms[58,59], Henrik Hjalgrim[26,60,61], Christopher Hübel[20,62,63,64], Gitta A. Jacob[65], Bitten Aagaard[66], Anders Jorgensen[26,67], Martin Jungkunz[68,69], Nikolaus Kleindienst[70], Nora Knoblich[43], Stefanie Koglin[71], Julia Kraft[5], Kristi Krebs[72], Christopher W. Lee[73], Yuhao Lin[20,62], Stefanie Lis[70,74], Amanda Lisoway[75,76], Ioannis A. Malogiannis[77], Amy Martinsen[78,79,80], Tolou Maslahati[71], Katharina Merz[81], Andreas Meyer-Lindenberg[2], Susan Mikkelsen[39], Christina Mikkelsen[38], Arian Mobascher[36], Gerard Muntané[50,51,82], Asmundur Oddsson[83], Sisse R. Ostrowski[26,38], Teemu Palviainen[84], Ole B. V. Pedersen[26,30], Geir Pedersen[85,86], Liam Quinn[30], Matthias A. Reinhard[81,87], Florian A. Ruths[88], Björn H. Schott[89,90,91,92], Michael Schredl[93],

Emanuel Schwarz[1,2,4], Cornelia E. Schwarze[94], Michael Schwinn[38], Tabea Sarah Send[2], Engilbert Sigurdsson[54,55], Katja Simon-Keller[3], Astros T. Skuladottir[83,95], Joaquim Soler[96,97,98], Anne Sonley[99,100], Erik Sørensen[38], Hreinn Stefansson[83], Peter Straub[9], Jaana Suvisaari[101], Martin Tesli[102,103], Jacob Træholt[38], Henrik Ullum[104], Maja P. Völker[3], G. Bragi Walters[83], Rujia Wang[20,62], Christian C. Witt[105], Gerhard Zarbock[106], Peter Zill[81], John-Anker Zwart[78,79,80], DBDS Genomic Consortium*, the GLAD Study*, HUNT All-In Psychiatry*, Ole A. Andreassen[41,107], Arnoud Arntz[108], Joanna M. Biernacka[24,109], Martin Bohus[70], Gerome Breen[20,62], Alexander L. Chapman[110,111], Sven Cichon[112,113,114], Lea K. Davis[9,115], Michael Deuschle[2,4], Sebastian Euler[116], Sabine C. Herpertz[4,117], Benjamin Hummelen[85], Andrea Jobst[81], Jaakko Kaprio[84], James L. Kennedy[75,76,100], Kelli Lehto[72], Klaus Lieb[36], Lourdes Martorell[50,51], Shelley McMain[100,118], Richard Musil[81,119], Vanessa Nieratschker[43,44], Markus M. Nöthen[120], Frank Padberg[81,87], Aarno Palotie[84,121,122], Juan C. Pascual[96,97,123], Nader Perroud[124], Josep A. Ramos-Quiroga[51,125,126,127], Ted Reichborn-Kjennerud[21,128], Marta Ribases[51,125,126,129], Stefan Roepke[71,130], Dan Rujescu[131], Sandra Sanchez-Roige[9,15,132], Claudia Schilling[93], Christian Schmahl[4,70], Kari Stefansson[83,95], Thorgeir E. Thorgeirsson[83], Gustavo Turecki[133], Elisabet Vilella[50,51], Thomas Werge[26,31], Bendik S. Winsvold[78,80,134], Johannes Wrege[135], Marcella Rietschel[3,138], Stephan Ripke[5,6,136,138] & Stephanie H. Witt[3,4,138] ✉

[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, 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, CA, 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, 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, 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, Queensland, Australia. [33]School of Psychology, The University of Queensland, Brisbane, Queensland, Australia. [34]Psychiatric Genetic Epidemiology Group, Research Department & Nic Waals Institute, Lovisenberg Diaconal Hospital, Oslo, Norway. [35]MRC Integrative Epidemiology Unit, Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK. [36]Department of Psychiatry and Psychotherapy, University Medical Center, University of Mainz, Mainz, Germany. [37]Department of Neuroscience, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark. [38]Department of Clinical Immunology, Copenhagen University Hospital, Rigshospitalet, Copenhagen, Denmark. [39]Department of Clinical Immunology, Aarhus University Hospital, Aarhus, Denmark. [40]Department of Medical Genetics, Oslo University Hospital, Oslo, Norway. [41]Centre for Precision Psychiatry, Division of Mental Health and Addiction, Oslo University Hospital and Institute of Clinical Medicine, University of 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, Alberta, Canada. [49]Neuroscience and Mental Health Institute, University of Alberta, Edmonton, Alberta, Canada. [50]Hospital Universitari Institut Pere Mata (HUIPM), Institute de Recerca Biomédica Catalunya Sud (IRB-CatSud) previously 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]Faculty of Medicine, Department of Psychiatry, School of Health Sciences, University of Iceland, Reykjavik, Iceland. [56]Bradford District Care NHS Foundation Trust, Bradford, UK. [57]Institute of Psychiatric Phenomics and Genomics (IPPG), LMU University Hospital, LMU Munich, Munich, Germany. [58]Human Genomics Research Group, Department of Biomedicine, University of Basel, Basel, Switzerland. [59]Institute of Human Genetics, University Hospital Bonn & University of Bonn, Bonn, Germany. [60]Danish Cancer Institute, Copenhagen, Denmark. [61]Department of Haematology, Copenhagen University Hospital, Rigshospitalet, Copenhagen, Denmark. [62]UK National Institute for Health Research (NIHR) Biomedical Research Centre, South London and Maudsley Hospital and King's College London, London, UK. [63]National Centre for Register-based Research, Aarhus University, Aarhus, Denmark. [64]Clinic for Child and Adolescent Psychiatry, German Red Cross Hospitals Westend, Berlin, Germany.

[65]Institute for Psychology, Department of Clinical Psychology and Psychotherapy, University of Freiburg, Freiburg, Germany. [66]Department of Clinical Immunology, Aalborg University Hospital, Aalborg, Denmark. [67]Psychiatric Center Copenhagen, Copenhagen, Denmark. [68]Institute for Medical and Data Ethics, Heidelberg University, Faculty of Medicine, Heidelberg, Germany. [69]German Cancer Research Center (DKFZ), Heidelberg, Germany. [70]Department of Psychosomatic Medicine and Psychotherapy, Central Institute of Mental Health, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany. [71]Department of Psychiatry and Neurosciences, Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt- Universität zu Berlin, Campus Benjamin Franklin, Berlin, Germany. [72]Estonian Genome Centre, Institute of Genomics, University of Tartu, Tartu, Estonia. [73]Faculty of Health and Medical Sciences, University of Western Australia, Perth, Western Australia, Australia. [74]Department of Clinical Psychology, Central Institute of Mental Health, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany. [75]Molecular Brain Science, Campbell Family Mental Health Research Institute, Centre for Addiction and Mental Health, Toronto, Ontario, Canada. [76]Institute of Medical Science, University of Toronto, Toronto, Ontario, Canada. [77]First Department of Psychiatry, Eginition Hospital Medical School, National and Kapodistrian University of Athens, Athens, Greece. [78]Department of Research and Innovation, Division of Clinical Neuroscience, Oslo University Hospital, Oslo, Norway. [79]Institute of Clinical Medicine, Faculty of Medicine, University of Oslo, Oslo, Norway. [80]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. [81]Department of Psychiatry and Psychotherapy, LMU University Hospital Munich, Ludwig Maximilian University Munich, Munich, Germany. [82]Institut de Biologia Evolutiva (UPF-CSIC), Department of Medicine and Life Sciences, Universitat Pompeu Fabra, Parc de Recerca Biomèdica de Barcelona, Barcelona, Spain. [83]deCODE genetics / AMGEN, Reykjavik, Iceland. [84]Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland. [85]Department of Research and Innovation, Division of Mental Health and Addiction, Oslo University Hospital, Oslo, Norway. [86]Institute of Basic Medical Sciences, Faculty of Medicine, University of Oslo, Oslo, Norway. [87]German Center for Mental Health (DZPG), Partner Site Munich-Augsburg, Munich-Augsburg, Germany. [88]South London and Maudsley NHS Foundation Trust, London, UK. [89]Department of Psychiatry and Psychotherapy, University Medical Center Göttingen, Göttingen, Germany. [90]German Center for Neurodegenerative Diseases (DZNE), Göttingen, Germany. [91]Leibniz Institute for Neurobiology, Magdeburg, Germany. [92]Department of Psychiatry and Psychotherapy, Otto von Guericke University, Magdeburg, Germany. [93]Department of Psychiatry and Psychotherapy, Sleep Laboratory, Central Institute of Mental Health, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany. [94]Department of Clinical Psychology and Psychotherapy, Charlotte Fresenius University, Wiesbaden, Germany. [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, Ontario, 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]KG Jebsen Centre for Neurodevelopmental Disorders, University of Oslo, Oslo, Norway. [108]Department of Clinical Psychology, University of Amsterdam, Amsterdam, the Netherlands. [109]Department of Psychiatry & Psychology, Division of Computational Biology, Mayo Clinic, Rochester, MN, USA. [110]Department of Psychology, Faculty of Arts and Social Sciences, Simon Fraser University, Vancouver, British Columbia, Canada. [111]DBT Centre of Vancouver, Vancouver, British Columbia, Canada. [112]Department of Biomedicine, University of Basel, Basel, Switzerland. [113]Medical Genetics, Institute of Medical Genetics and Pathology, University Hospital Basel, Basel, Switzerland. [114]Institute of Neuroscience and Medicine (INM-1), Research Center Juelich, Juelich, Germany. [115]Vanderbilt Genetics Institute, Vanderbilt University Medical Center, Nashville, TN, USA. [116]Department of Consultation Psychiatry and Psychosomatics, University Hospital Zurich, Zurich, Switzerland. [117]Department of General Psychiatry, University Hospital Heidelberg, Heidelberg University, Heidelberg, Germany. [118]Borderline Personality Disorder Clinic, General Adult Psychiatry and Health Systems Division, Centre for Addiction and Mental Health, Toronto, Ontario, Canada. [119]Oberberg Fachkliniken for Psychiatry, Psychosomatics and Psychotherapy, Bad Tölz, Germany. [120]Insitute of Human Genetics, University Hospital Bonn & University of Bonn, Bonn, Germany. [121]Analytic and Translational Genetics Unit, Department of Medicine, Department of Neurology and Department of Psychiatry, Massachusetts General Hospital, Boston, MA, USA. [122]The Stanley Center for Psychiatric Research and Program in Medical and Population Genetics, The Broad Institute of MIT and Harvard, Cambridge, MA, USA. [123]Mental Health and Psychiatry Department, Vic Hospital Consortium, Vic, Spain. [124]Department of Psychiatric, University Hospitals of Geneva, Geneva, Switzerland. [125]Psychiatric Genetics Unit, Group of Psychiatry, Mental Health and Addiction, Vall d'Hebron Research Institute (VHIR), Universitat Autònoma de Barcelona, Barcelona, Spain. [126]Department of Psychiatry, Hospital Universitari Vall d'Hebron, Barcelona, Spain. [127]Department of Psychiatry and Forensic Medicine, Universitat Autònoma de Barcelona (UAB), Barcelona, Spain. [128]Institute of Clinical Medicine, University of Oslo, Oslo, Norway. [129]Department of Genetics, Microbiology, and Statistics, Faculty of Biology, Universitat de Barcelona, Barcelona, Spain. [130]Department of Psychiatry, Oberberg Fachkliniken for Psychiatry, Psychosomatics and Psychotherapy, Wendisch Rietz, Germany. [131]Department of Psychiatry, University of Halle, Halle, Germany. [132]Institute for Genomic Medicine, University of California San Diego, La Jolla, CA, USA. [133]Douglas Institute, Department of Psychiatry, McGill University, Montreal, Quebec, Canada. [134]Department of Neurology, Oslo University Hospital, Oslo, Norway. [135]Medical Faculty, University of Basel, Basel, Switzerland. [136]Analytic and Translational Genetics Unit, Massachusetts General Hospital, Boston, MA, USA. [137]These authors contributed equally: Fabian Streit, Swapnil Awasthi. [138]These authors jointly supervised this work: Marcella Rietschel, Stephan Ripke, Stephanie H. Witt. *Lists of authors and their affiliations appear at the end of the paper. ✉e-mail: fabian.streit@zi-mannheim.de; stephanie.witt@zi-mannheim.de

DBDS Genomic Consortium

Søren Brunak[28], Mie T. Bruun[29], Lea A. N. Christoffersen[30,31], Maria Didriksen[37,38], Khoa M. Dinh[38,39], Joseph Dowsett[38], Ole Kristian Drange[41,42], Susanne Edelmann[43,44], Christian Erikstrup[7,39], Thomas F. Hansen[52,53], Henrik Hjalgrim[26,60,61], Bitten Aagaard[66], Susan Mikkelsen[39], Christina Mikkelsen[38], Sisse R. Ostrowski[26,38], Ole B. V. Pedersen[26,30], Liam Quinn[30], Michael Schwinn[38], Erik Sørensen[38], Hreinn Stefansson[83], Henrik Ullum[104], Kari Stefansson[83,95] & Thomas Werge[26,31]

A full list of members and their affiliations appears in the Supplementary Information.

the GLAD Study

Zain-Ul-Abideen Ahmad[20], Christopher Hübel[20,62,63,64], Yuhao Lin[20,62], Rujia Wang[20,62] & Gerome Breen[20,62]

A full list of members and their affiliations appears in the Supplementary Information.

HUNT All-In Psychiatry

Ole A. Andreassen[41,107]

A full list of members and their affiliations appears in the Supplementary Information.

Methods

All study participants gave informed consent, and the studies were approved by the respective ethical committees (for details, see Supplementary Table 1 and Supplementary Note).

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). 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 ten large-scale biobank or cohort studies that provided summary statistics from GWAS of BPD (n_{cases} = 9,634; $n_{controls}$ = 1,037,117). Of those, nine cohorts identified BPD status following ICD codes (ICD-10: F60.3; ICD-9: 3018D/301.83; 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, along with country of origin for each sample, are documented in Supplementary Table 1. A detailed description of the study design, ascertainment of cases and controls, genotyping arrays used in each study and study-specific quality control and imputation procedures for studies sharing summary statistics is provided in Supplementary Table 2 and Supplementary Note.

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 2). Quality control and imputation were carried out using the RICOPILI pipeline[92]. A detailed description can be found in Supplementary Note.

GWAS

For studies providing genotype data at the individual level, GWAS analyses were conducted within each dataset using an additive logistic regression model implemented in PLINK (v.1.9)[93] on imputed genetic dosage data, with the relevant ancestry principal components (details in Supplementary Note) included as covariates. Details on the association analyses for the cohorts providing summary statistics can be found in Supplementary Table 2 and Supplementary Note. In addition to the main analysis, we performed sex-stratified analyses. Dedicated studies of BPD often exclude individuals with BIP and SCZ. To explore the influence of comorbidity with these severe mental disorders on our results, as a sensitivity analysis, association analyses were performed excluding individuals 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 3–5). In the case of cohorts in which 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), individuals 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[94] as implemented in RICOPILI[92] with a genome-wide significance threshold of 5×10^{-8}. To evaluate consistency of effect sizes, we computed Cochran's Q-statistics and the corresponding heterogeneity P values and I^2 statistics, which quantify the percentage of total variation owing to heterogeneity. This procedure excluded SNPs with minor allele frequency of < 1% or an imputation quality score of <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 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, as those 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 ten 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 ($P < 5 \times 10^{-8}$) 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 previously described[95]) and the required sample size for a power over 80% (for details, see Supplementary Table 6 and Supplementary Fig. 1).

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 n_{eff} = 43,234 (2 × 21,617)[95], a corresponding sample prevalence of 0.5 and assuming a population prevalence of 1.5% (ref. 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-to-male ratio of 3:1 (refs. 4,5). To assess the contribution of confounding to the observed signal, we calculated the attenuation ratio ($\frac{(\text{Intercept}-1)}{(\text{Mean}\,\chi^2-1)}$).

To test whether the SNP heritability estimates of the sex-stratified GWAS differed (two-sided test), and to test whether the genetic correlations between GWAS subsets (males versus females; individual-level versus summary cohorts) were smaller than 1 (one-sided test), we used the block jackknife extension of LDSC[20], which has previously been described[96].

To assess replication of the observed SNP associations, lead SNPs for the thresholds $P < 1 \times 10^{-6}$ and $P < 5 \times 10^{-8}$ were tested for replication in a meta-analysis of two independent datasets (Spain 2 and All of Us; total n_{cases} = 685; total $n_{controls}$ = 107,750). Additionally, for SNPs with $P < 1 \times 10^{-6}$ in the discovery GWAS analysis, discovery and replication datasets were analyzed in a combined GWAS meta-analysis, and loci with $P < 5 \times 10^{-8}$ 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 (v.1.07)[21] as implemented in FUMA (v.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 kb upstream and 10 kb downstream of the genes. A total of 9,237 gene sets were analyzed (MSigDB v.2023.1Hs; 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, $P < 2.5 \times 10^{-6}$ (0.05/19,843); gene set, $P < 5.4 \times 10^{-6}$ (0.05/9,237)).

Additionally, 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 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 GTEx (v.8) data for 54 tissue types[24], and data from BrainSpan, representing 29 different ages and 11 general developmental stages[25] as implemented in FUMA (v.1.5.2)[22]. In addition, we used 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 cell types at the level of

31 'superclusters' and retained the existing cell-type nomenclature. We calculated expression proportion values for each cell type ('gene specificity', previously described in detail[29,97]), resulting in ~1,300 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 for 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[98]. Additionally, we performed a tractability analysis (that is, 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 24 June 2024. Furthermore, we used the GREP (https://github.com/saorisakaue/GREP) pipeline to test for drug-target enrichment across clinical ATC or ICD categories[99]. Finally, we queried the DGIdb[100].

In addition to these approaches, focusing 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 DGIdb, 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 (v.1.08), 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 $P < 6.80 \times 10^{-5}$ (0.05/735).

PGS

We applied PGS to predict case–control status in the datasets for which individual-level genotype information was available. PGS 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 information[31]. The posterior effect sizes account for the linkage disequilibrium between SNPs using external linkage disequilibrium reference panels constructed from the 1000 Genomes Project Phase 3 European samples. PGS were calculated based on both the present meta-analysis and the prior BPD GWAS from 2017 (consisting of the Germany dataset)[18] for comparison. Using leave-one-out results of the present meta-analysis (2026 meta-GWAS), whereby 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 and All of Us; total $n_{cases} = 685$; total $n_{controls} = 107{,}750$); Supplementary Tables 1 and 2). 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, Spain 2, All of Us).

As an effect measure of the association between BPD-PGS and case–control status, the Nagelkerke-pseudo-R^2 (NkR^2) from logistic regression 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[101] of the population, assuming a lifetime disease risk of 1.5% (ref. 3). Additionally, the 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 OR of the deciles of the PGS distribution, PRS-CS were normalized to have a mean of zero 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 2026 ($n_{controls} = 5{,}035$; $n_{cases} = 2{,}751$). This part of the PGS analysis excluded the All of Us dataset, as its individual-level data could not be merged owing 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 one (lowest PRS-CS) to ten (highest PRS-CS). We then created dummy variables by coding observations within each decile as cases; those outside that decile were coded as controls. The dummy variable ranged from deciles two (Q2) to ten (Q10), with decile one (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 ORs for each decile were calculated by exponentiating the coefficients from the logistic regression model, controlling for principal components (1, 2, 3, 4, 5, 6, 8, 10, 12, 14, 15 and 20). For the reference comparison (Q1/Q1), an OR of 1 was assigned, indicating no effect. The ORs for each of the remaining deciles (Q2 to Q10) were computed accordingly: $OR_{Q1/Qj} = \exp(\beta j)$ for $j = 2,3,\ldots,10$. Additionally, 95% CI for the odds ratios were calculated using the standard errors of the coefficients ($CI_{lower} = \exp(\beta j - 1.96 \times \text{s.e.}j)$; $CI_{upper} = \exp(\beta j + 1.96 \times \text{s.e.}j)$). As an additional measure of effect, the highest decile (Q10) was compared to the middle 10% of the distribution as a 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 LDSC[20]. Calculations were carried out with a free intercept and the 1000 Genomes dataset (EUR) as a reference linkage disequilibrium structure panel[102].

In a targeted approach, genetic correlations were calculated with 50 GWAS incorporating a range of other disorders and traits relevant to BPD, selected based on reported phenotypic and genetic associations in the literature, theoretical considerations 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) and cognition (Supplementary Table 22). A Bonferroni *P* value threshold, corrected for the number of respective tests, was applied ($P < 1 \times 10^{-3}$ (0.05/50)).

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 BioVU[32] were used to test the association of BPD-PGS with medical phenotypes based on two or more ICD diagnoses ('phecodes'; Phecode Map 1.2) documented in the EHRs. For each phecode, a case was defined as having ≥1 ICD code in the phecode group on ≥2 distinct dates, and a control was defined as having no codes in that phecode group. Individuals with only one 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 individuals 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 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 ten genetic principal components as covariates. A total of 1,431 phecodes were included, and a Bonferroni-corrected threshold of $P < 3.49 \times 10^{-5}$ (0.05/1,431) was applied.

PheWAS in UKB

An additional PheWAS analysis was carried out using data from the UKB. PGS for BPD were calculated using PRS-CS based on summary statistics from the discovery meta-analysis, excluding the UKB samples (12,157 cases; 1,039,897 controls). The PheWAS was calculated using phecodes (Phecode Map 1.2) based on diagnoses recorded in Hospital Episode Statistics in a maximum sample of 316,635 individuals (211 individuals with phecode 301.20). For each phecode, a case was defined as having ≥1 ICD code in the phecode group, and a control was defined as having no codes in that phecode group. A total of 1,250 phecodes with at least 100 cases were tested for association with sex, age, batch, the first 40 genetic principal components and assessment center as covariates using a Bonferroni-corrected threshold of $P < 4.00 \times 10^{-5}$ (0.05/1,250).

Reporting summary

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

Data availability

Individual-level data are not publicly available owing to ethical restrictions. The results of the meta-analysis and PGS weights are publicly available through the website of the Psychiatric Genomics Consortium (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 22). The exceptions are the GWAS summary statistics for self-harm[103], 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. Source data are provided with this paper.

Code availability

No custom code was developed for the present analyses. GWAS in datasets with individual data were carried out within each dataset using PLINK (v.1.9; http://pngu.mgh.harvard.edu/purcell/plink). Quality control and imputation were carried out using RICOPILI (https://sites.google.com/a/broadinstitute.org/ricopili; https://github.com/Ripkelab/ricopili). Prephasing and imputation was carried out using EAGLE (v.2.4.1; https://alkesgroup.broadinstitute.org/Eagle) and MINIMAC3 (http://genome.sph.umich.edu/wiki/minimac3) using the Haplotype Reference Consortium reference panel (Release 1.1, dataset EGAD00001002729). 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 (v.1.5.2) was used for downstream analyses, including gene-based and gene-set analyses (https://fuma.ctglab.nl), based on MAGMA (v.1.07; http://ctglab.nl/software/magma). Functional annotation within FUMA used data from MSigDB v2023.1Hs, GTEx (v.8; https://fuma.ctglab.nl/downloadPage) and BrainSpan (https://www.brainspan.org). Additionally, genes were prioritized using PoPS (v.0.2; 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).

References

92. Lam, M. et al. RICOPILI: Rapid Imputation for COnsortias PIpeLIne. *Bioinformatics* **36**, 930–933 (2020).
93. Chang, C. C. et al. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**, 7 (2015).
94. Willer, C. J., Li, Y. & Abecasis, G. R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190–2191 (2010).
95. Grotzinger, A. D., Fuente, J., Privé, F., Nivard, M. G. & Tucker-Drob, E. M. Pervasive downward bias in estimates of liability-scale heritability in genome-wide association study meta-analysis: a simple solution. *Biol. Psychiatry* **93**, 29–36 (2023).
96. 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).
97. 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).
98. Gaulton, A. et al. ChEMBL: a large-scale bioactivity database for drug discovery. *Nucleic Acids Res.* **40**, D1100–D1107 (2012).
99. Sakaue, S. & Okada, Y. GREP: genome for repositioning drugs. *Bioinformatics* **35**, 3821–3823 (2019).
100. 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).
101. 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).
102. 1000 Genomes Project Consortium. A global reference for human genetic variation. *Nature* **526**, 68–74 (2015).
103. Campos, A. I. et al. Genetic aetiology of self-harm ideation and behaviour. *Sci. Rep.* **10**, 9713 (2020).

Acknowledgements

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. We thank the Estonian Biobank participants and the EstBB Research Team. This work was endorsed by the German Center for Mental Health (DZPG). This research has been conducted using the UKB Resource under application numbers 16406 and 96802. The GTEx Project was supported by the Common Fund of the Office of the Director of the US National Institutes of Health and by NCI, NHGRI, NHLBI, NIDA, NIMH and NINDS. This work was supported by the Brain & Behavior Research Foundation with support from the Families for Borderline Personality Disorder Research (grant no. 31537 to F.S.); the Hector Foundation II (to E.S. and A.M.L.); the German Research Foundation (DFG CRC 1436 SP A05 to B.H.S.; grant no. 402170461 to S. Ripke; project no. 514201724 to U.H.; grant no. NI 1332/16-1 to V.N.); the European Union's Horizon Programme (grant no. 101057454 to S. Ripke.; grant no. 847776 to K. Lehto); the European Research Council (grant agreement ID: 101042183, to Y.L.), the Berlin Institute of Health at Charité (to S. Ripke); the Research Council of Norway (grant no. 274611 to E.C.C. and T.R.K.; grant no. 223273 to O.K.D.; grant nos. 324499, 324252 and 296030 to O.A.A., grant no. 336085 to A. Havdahl); the South-Eastern Norway Regional Health Authority (grant no. 2021045 to E.C.C.; grant nos. 2020022, 2922083, 2022039 and 2019097 to A. Havdahl); the European Union's Horizon Europe

Programme (grant nos. FAMILY 101057529; HOMME 101142786; Marie Skłodowska-Curie grant ESSGN 101073237 to A. Havdahl); Nordforsk (grant no. 164218 to O.A.A.); South East Norway Health Authority (2023-031 to O.A.A.); the European Union's Horizon 2020 Programme (grant no. 945151 to U.H.); the Novo Nordisk Foundation (grant nos. NNF17OC0027594 and NNF14CC0001 to S.B.); the Estonian Research Council (grant no. PSG615 to K. Lehto); the Estonian Centre of Excellence for Well-Being Sciences, funded by the Estonian Ministry of Education and Research (grant no. TK218 to K. Lehto); the US National Institutes of Health (grant no. T32GM139790 to J. Gonzalez); the German Federal Ministry of Education and Research through support of the German Center for Mental Health (grant no. 01EE2306B to V.N.); and the Australian National Health and Medical Research Council (grant nos. 1172917 and 2025674 to L.C.C.). E.C.C. is a member of the MRC Integrative Epidemiology Unit at the University of Bristol, which is supported by the Medical Research Council and the University of Bristol (MC_UU_00032/1). This research was also supported by the National Institute of Mental Health of the National Institutes of Health under award number R01 MH124873 (to S. Ripke). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Author contributions

F.S. and S.A. conceived of and designed the study, led the project and coordinated analyses. F.S., S.A., A.S.M.H., A. Braun, M.N., E.M., O.B., J.F., L.Z., C.M.C., D.A., E.Z., J.N., J.G., A. Harder and Y. Lu performed central analyses and contributed to data interpretation. F.S., S.A., A.S.M.H., A. Braun, M.N., E.M., L.C.-C., M.K.E., D.S.F.d.S., J.C.F., S.L., B.H.S., M. Schredl, E. Schwarz, M.P.V., C.C.W., S.S.-R., S. Ripke and S.H.W. drafted and revised the paper. The following authors contributed to the individual cohorts included in the study: Z.A., Z.-U.-A.A., H.A., A. Batzler, M.E.B., O.M.B.-d.W., S.B., M.T.B., L.A.N.C., B.J.C., E.C.C., N.D., M. Didriksen, K.M.D., S.D., J.D., O.K.D., H.D., S. Edelmann, C.E., E.F., A.F., J.C.F., M.G., A.G.-Z., T.F.H., M.H., R.P.H., A. Havdahl, U.H., S.H., H.H., C.H., G.A.J., B.A., A. Jorgensen, M.J., N. Kleindienst, N. Knoblich, S.K., J. Kraft, K.K., C.W.L., Y. Lin, S.L., A.L., I.A.M., A. Martinsen, T.M., K.M., A.M.-L., S. Mikkelsen, C.M., A. Mobascher, G.M., A.O., S.R.O., T.P., O.B.V.P., G.P., L.Q., M.A.R., F.A.R., B.H.S., M. Schredl, C.E.S., M. Schwinn, T.S.S., E. Sigurdsson, K.S.-K., A.T.S., J. Soler, A.S., E. Sørensen, H.S., P.S., J. Suvisaari, M.T., J.T., H.U., G.B.W., R.W., C.C.W., G.Z., P.Z. and J.-A.Z. performed cohort collection, genotyping, data generation and cohort-level analyses. O.A.A., A.A., J.M.B., M.B., G.B., A.L.C., S.C., L.K.D., M. Deuschle, S. Euler, S.C.H., B.H., A. Jobst, J. Kaprio, J.L.K., K. Lehto, K. Lieb, L.M., S. McMain, R.M., V.N., M.M.N., F.P., A.P., J.C.P., N.P., J.A.R.-Q., T.R.-K., M. Ribases, S. Roepke, D.R., S.S.-R., C. Schilling, C. Schmahl, K. Stefansson, T.E.T., G.T., E.V., T.W., B.S.W. and J.W. were the principal investigators responsible for individual cohorts and resources. M. Rietschel, S. Ripke and S.H.W. jointly supervised the study. All authors reviewed and approved the final paper.

Competing interests

O.A.A. is a consultant to Precision Health and has received speaker's honoraria from BMS, Lilly, Lundbeck, Janssen, Otsuka and Sunovion. S.B. has ownerships in Hoba Therapeutics Aps, Novo Nordisk, Lundbeck and Eli Lilly and Company. J.L.K. is a Scientific Advisory Board member of Myriad Neuroscience. R.M. 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. F.P. is a member of the Scientific Advisory Boards of Brainsway and Sooma. He has received speaker's honoraria from Mag&More and the neuroCare Group. J.A.R.-Q. was on the speakers' bureau and/or acted as a consultant for Biogen, Idorsia, Casen-Recordati, Janssen-Cilag, Novartis, Takeda, Bial, Sincrolab, Neuraxpharm, BMS, Medice, Rubió, Uriach, Technofarma and Raffo in the last 3 years. He also received travel awards for taking part in psychiatric meetings from Idorsia, Janssen-Cilag, Rubió, Takeda, Bial and Medice. The Department of Psychiatry, which he chaired, 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. C. Schilling received lecturer fees from Idorsia Pharmaceuticals (2024). M.M.N. has received fees for membership in an advisory board from HMG Systems Engineering, 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). M.M.N. receives salary payments from Life & Brain and holds shares in Life & Brain. A.O., A.S., H.S., G.B.W., T.E.T. and K.S. are employees of deCODE genetics/Amgen. The other authors declare no competing interests.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41588-026-02654-3.

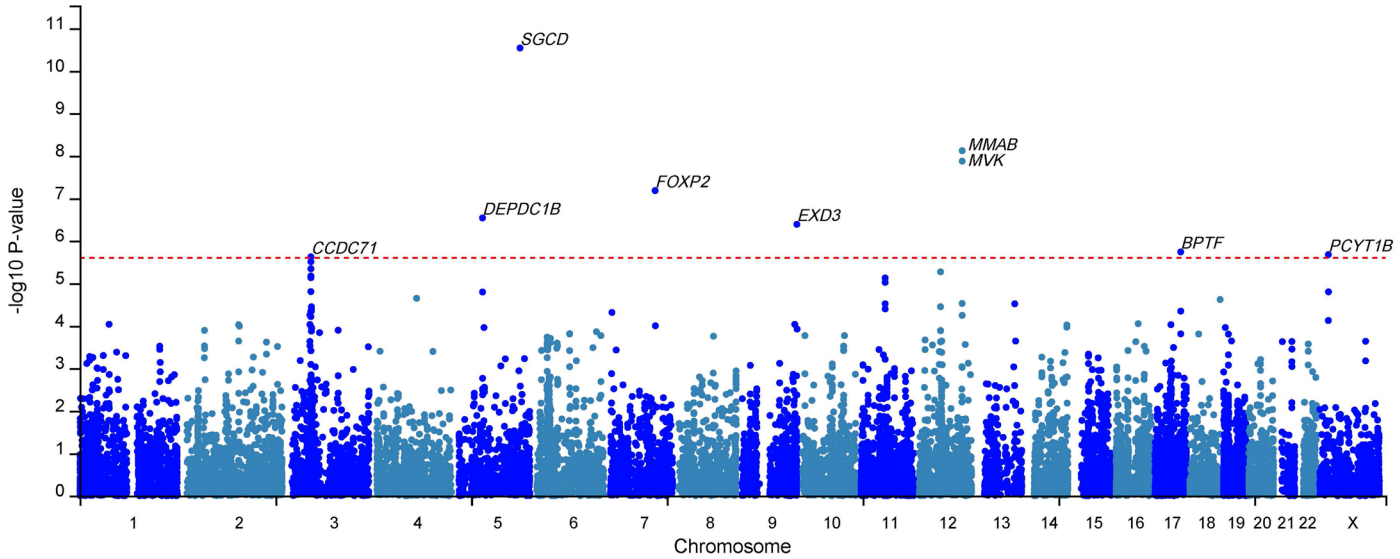
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41588-026-02654-3.

Correspondence and requests for materials should be addressed to Fabian Streit or Stephanie H. Witt.

Peer review information *Nature Genetics* thanks Nathan Gillespie, Daniel Levey and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

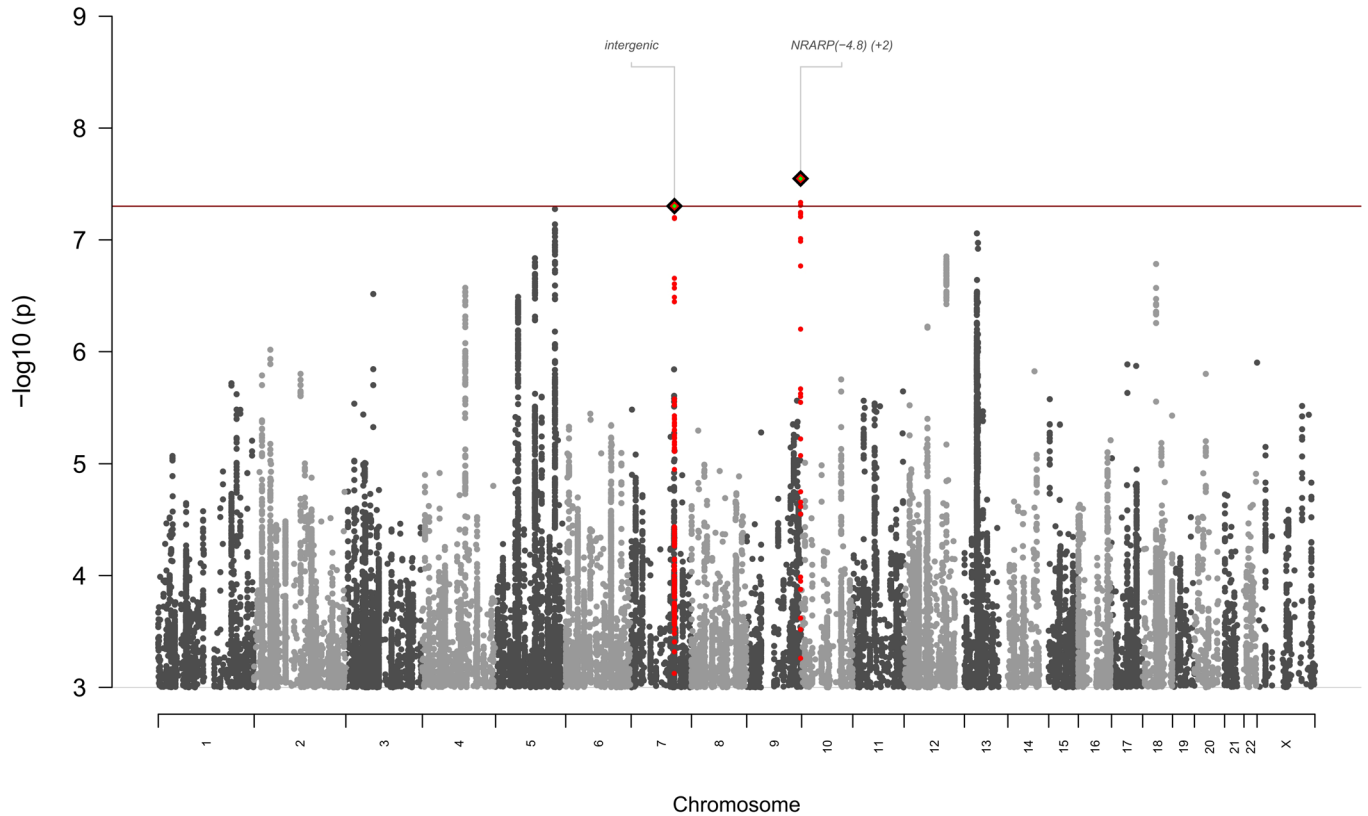
Reprints and permissions information is available at www.nature.com/reprints.

https://doi.org/10.1038/s41588-026-02654-3



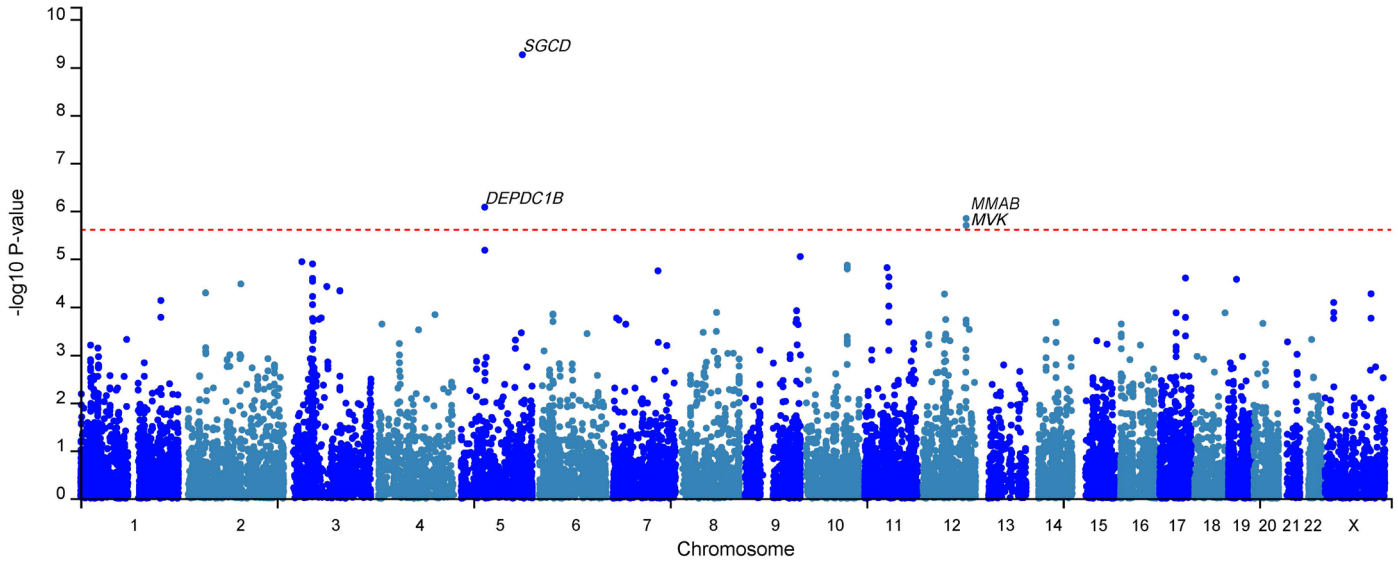
Extended Data Fig. 1 | Manhattan plot of the gene-based analysis (n_{cases} = 12,339, $n_{controls}$ = 1,041,717). The two-sided $-\log_{10}$ *P*-value of the MAGMA gene-based test (SNP-wise mean model) for each gene is indicated on the *y*-axis (chromosomal position shown on the *x*-axis). The red line indicates significance after correction for multiple testing ($P < 0.05/19{,}843 = 2.5\times10^{-6}$). Gene names are given for significant genes.

https://doi.org/10.1038/s41588-026-02654-3



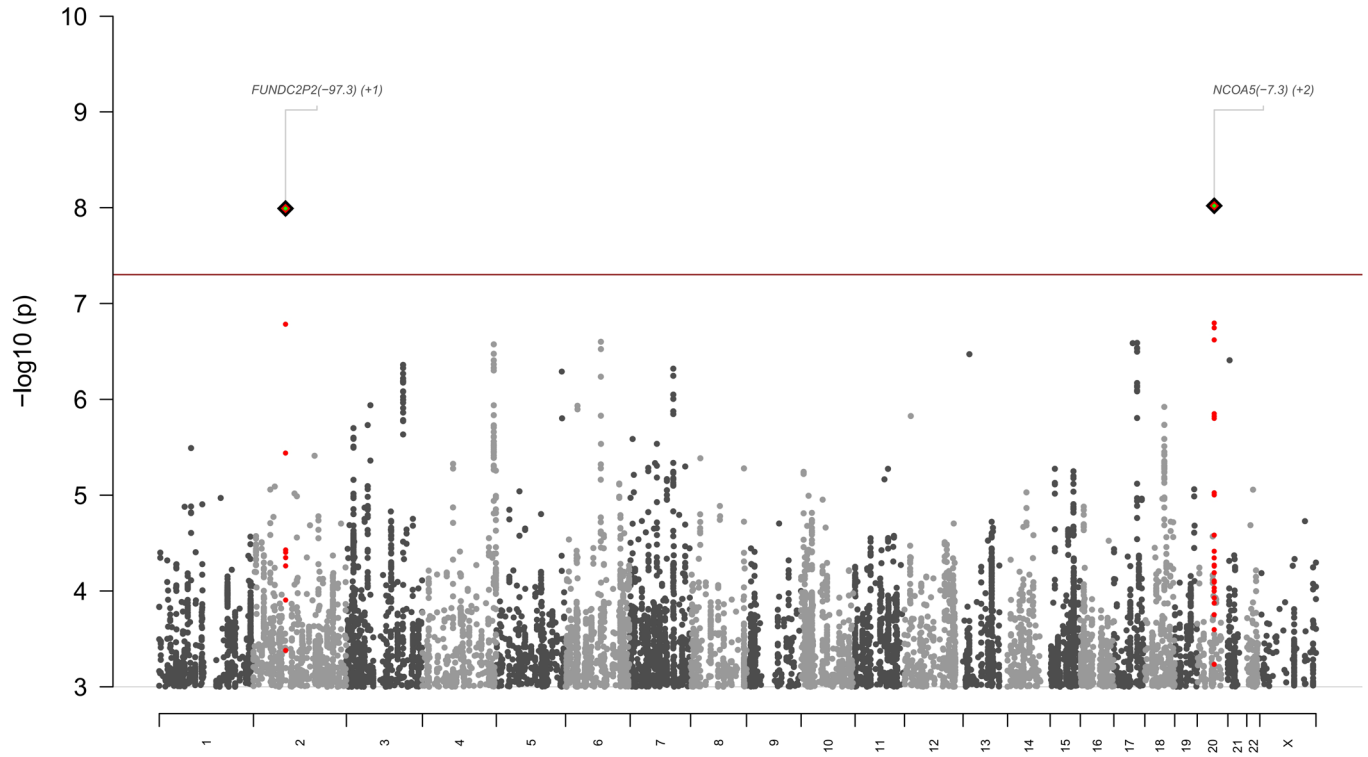
Extended Data Fig. 2 | Female subset Manhattan plot of the GWAS meta-analysis (n_{cases} = 10,025, $n_{controls}$ = 547,333). The two-sided $-\log_{10}$ *P*-value for each single-nucleotide polymorphism (SNP) in the inverse variance weighted GWAS meta-analysis is indicated on the *y*-axis (chromosomal 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 as diamonds, and SNPs in linkage disequilibrium with the index SNPs are highlighted in green.

https://doi.org/10.1038/s41588-026-02654-3



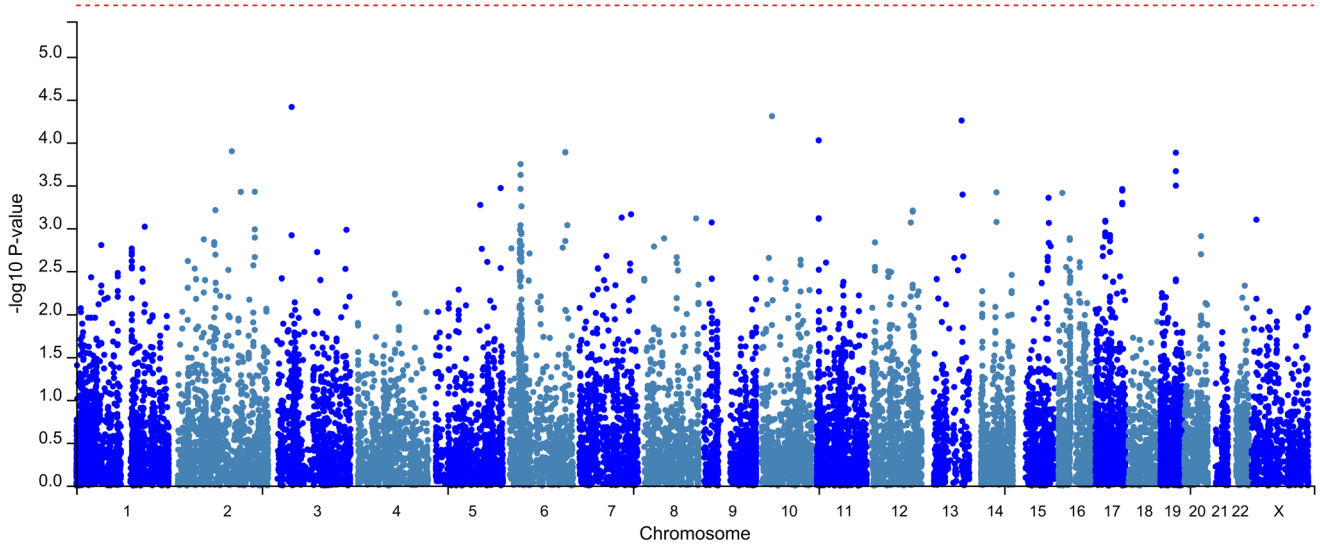
Extended Data Fig. 3 | Female subset Manhattan plot of the gene-based analysis (n_{cases} = 10,025, $n_{controls}$ = 547,333). The two-sided $-\log_{10}$ *P*-value of the MAGMA gene-based test (SNP-wise mean model) for each gene is indicated on the *y*-axis (chromosomal position shown on the *x*-axis). The red line indicates significance after correction for multiple testing ($P < 0.05/19{,}844 = 2.5\times10^{-6}$). Gene names are given for significant genes.

https://doi.org/10.1038/s41588-026-02654-3



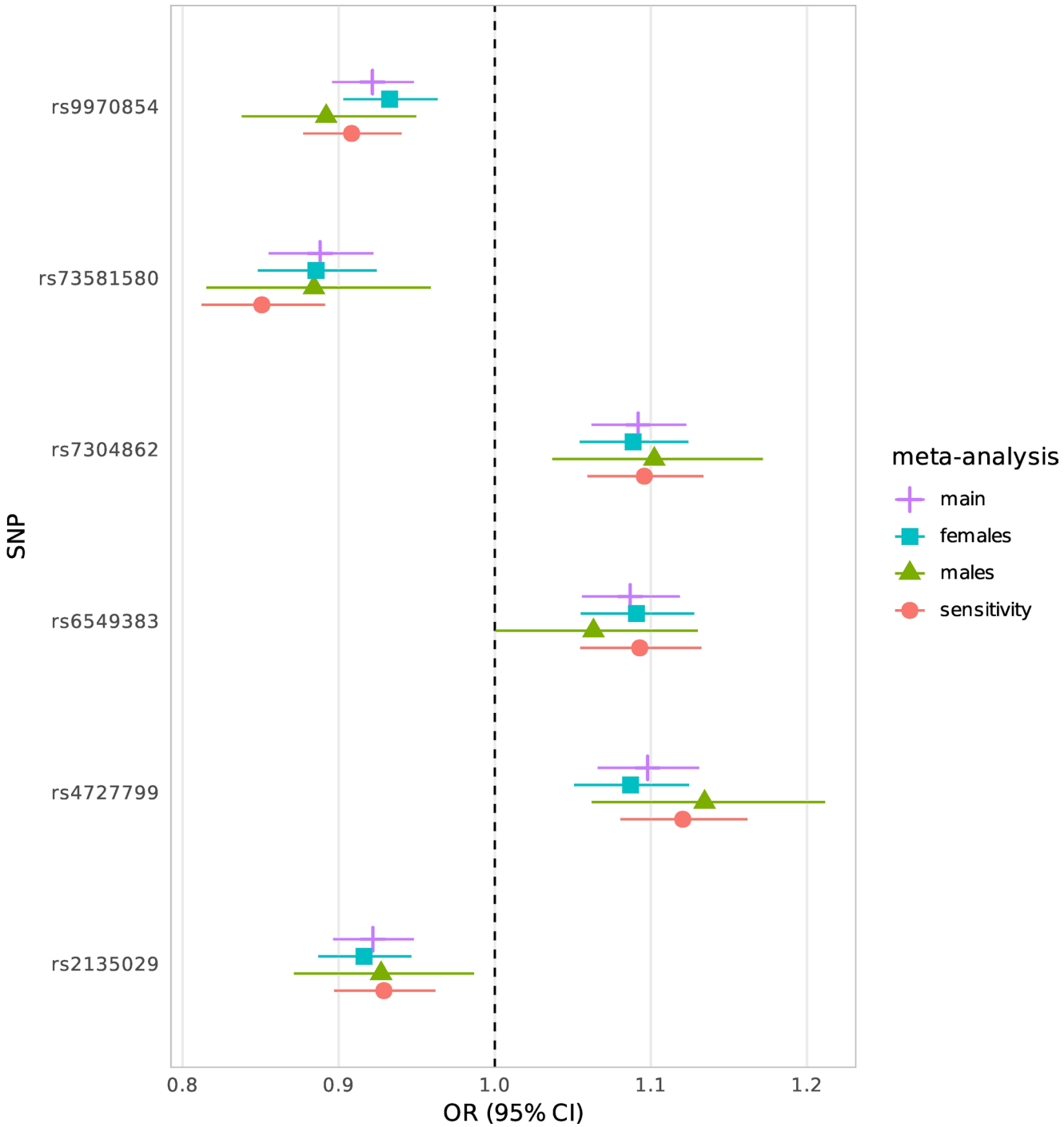
Extended Data Fig. 4 | Male subset Manhattan plot of the GWAS meta-analysis (n_{cases} = 2,260, $n_{controls}$ = 485,444). The two-sided $-\log_{10}$ *P*-value for each single-nucleotide polymorphism (SNP) in the inverse variance weighted GWAS meta-analysis is indicated on the *y*-axis (chromosomal position 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 as diamonds, and SNPs in linkage disequilibrium with the index SNPs are highlighted in red.

https://doi.org/10.1038/s41588-026-02654-3



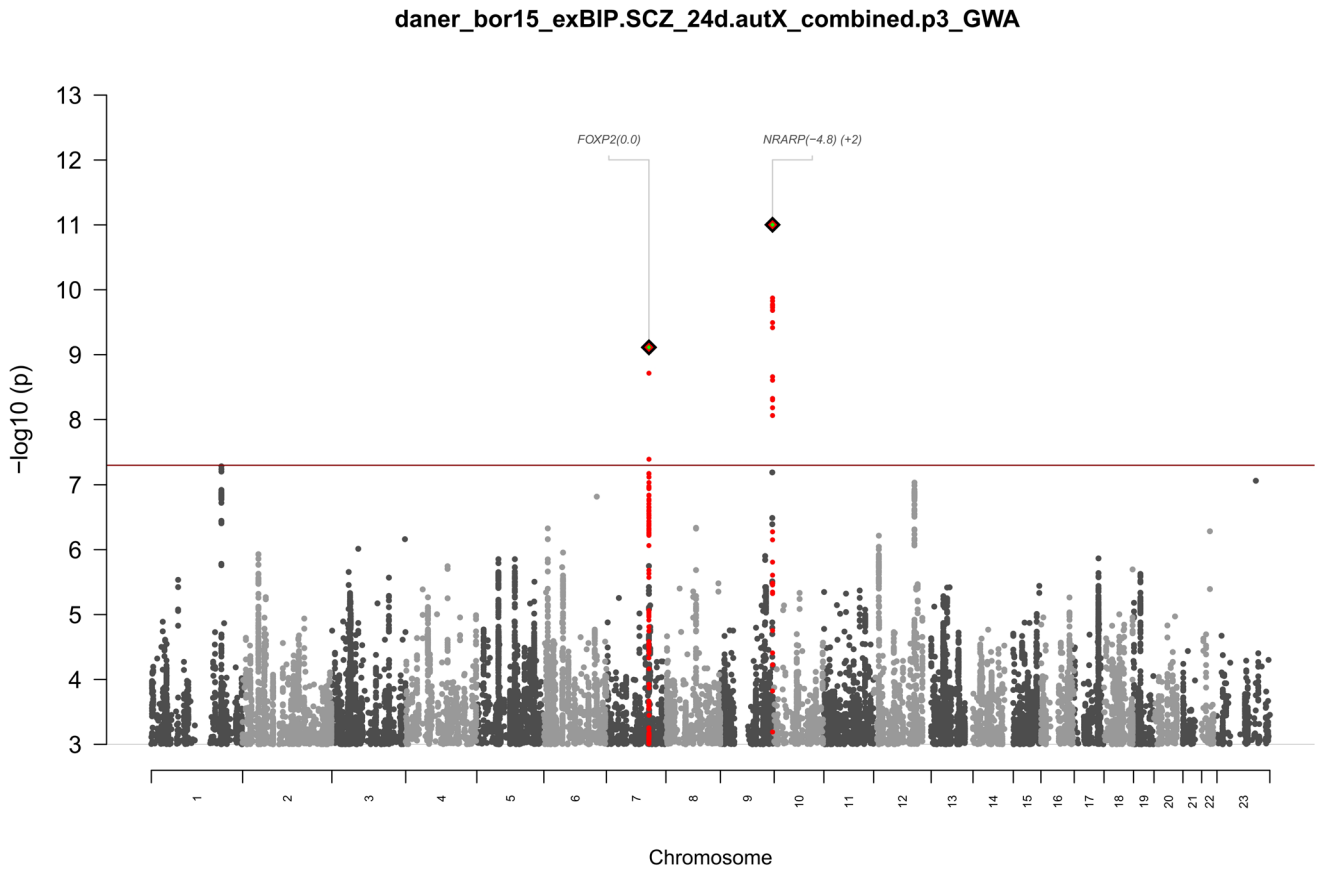
Extended Data Fig. 5 | Male subset Manhattan plot of the gene-based analysis (n_{cases} = 2,260, $n_{controls}$ = 485,444). The two-sided $-\log_{10}$ *P*-value of the MAGMA gene-based test (SNP-wise mean model) for each gene is indicated on the *y*-axis (chromosomal position shown on the *x*-axis). The red line indicates significance after correction for multiple testing ($P < 0.05/19{,}856 = 2.5\times10^{-6}$).

https://doi.org/10.1038/s41588-026-02654-3



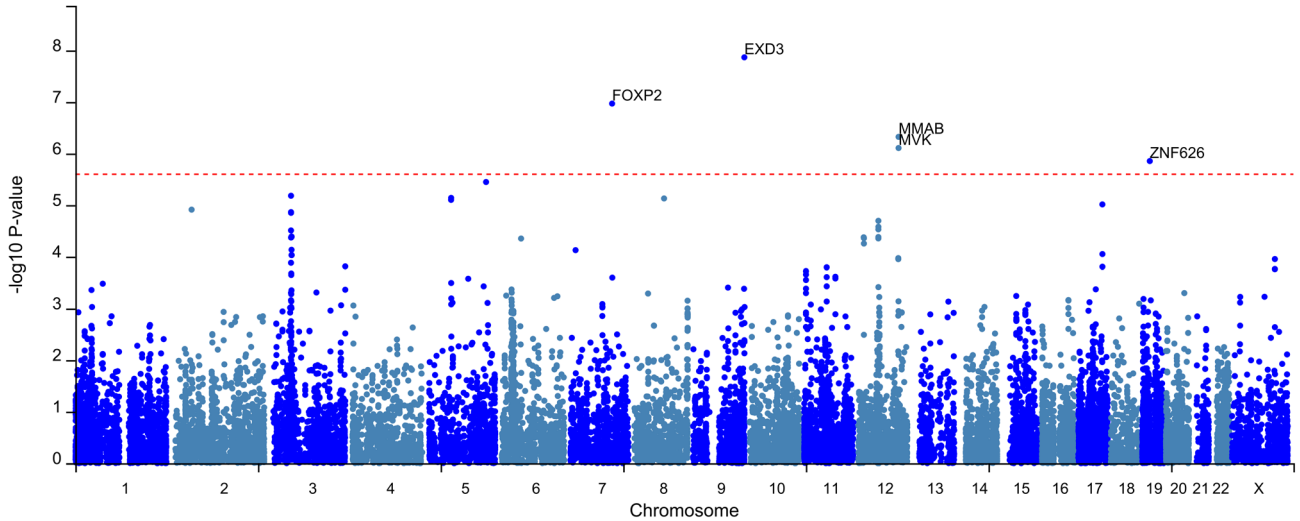
Extended Data Fig. 6 | Comparison of the effect sizes of the six GWAS lead SNPs. Odds ratio (OR) estimates from the inverse variance weighted GWAS meta-analyses with error bars indicating 95% confidence intervals are presented for the main meta-analysis, the sex-stratified analyses, and the sensitivity 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. Main analysis: n_{cases} = 12,339, $n_{controls}$ = 1,041,717; female subset: n_{cases} = 10,025, $n_{controls}$ = 547,333; male subset: n_{cases} = 2,260, $n_{controls}$ = 485,444; sensitivity analysis: n_{cases} = 8,618, $n_{controls}$ = 1,027,690. Detailed results can be found in Supplementary Table 12.

https://doi.org/10.1038/s41588-026-02654-3



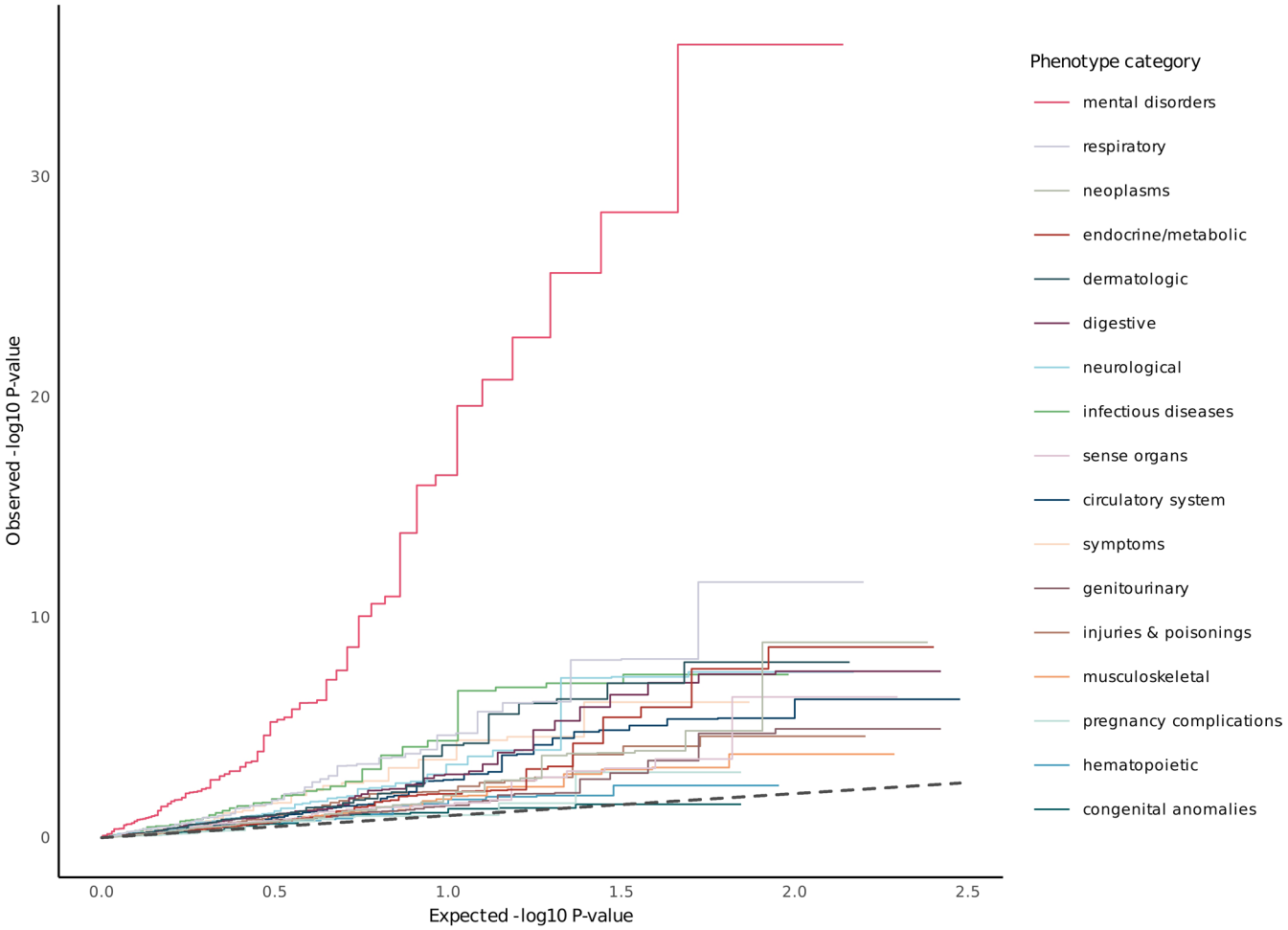
Extended Data Fig. 7 | Sensitivity analysis Manhattan plot of the GWAS meta-analysis of BPD excluding individuals with schizophrenia or bipolar disorder (n_{cases} = 8,618, $n_{controls}$ = 1,027,690). The two-sided $-\log_{10}$ *P*-value for each single-nucleotide polymorphism (SNP) in the inverse variance weighted GWAS meta-analysis 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 as diamonds, and SNPs in linkage disequilibrium with the index SNPs are highlighted in red.

https://doi.org/10.1038/s41588-026-02654-3



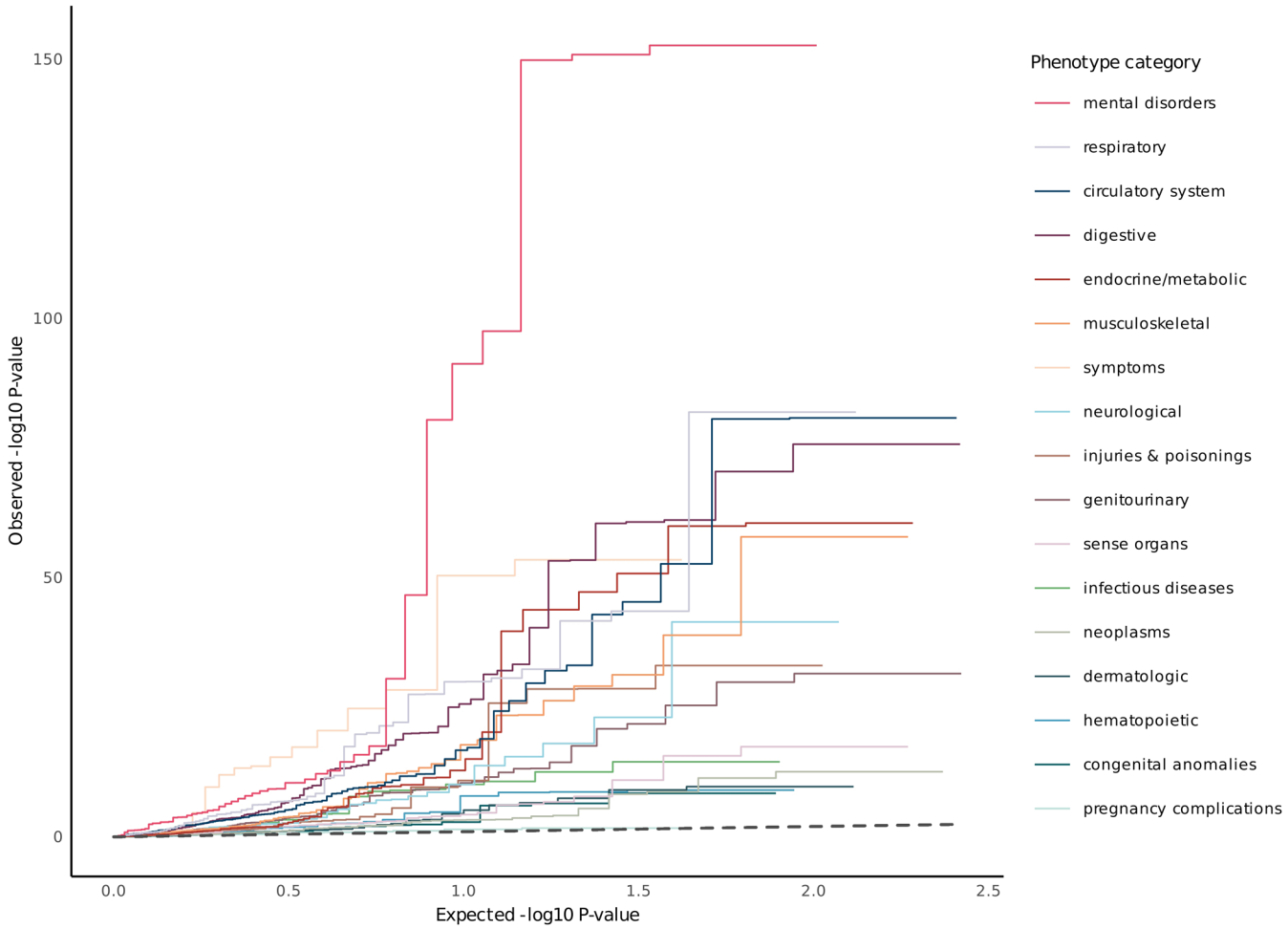
Extended Data Fig. 8 | Sensitivity analysis Manhattan plot of the gene-based analysis in the subset excluding individuals with schizophrenia or bipolar disorder (n_{cases} = 8,618, $n_{controls}$ = 1,027,690). The two-sided $-\log_{10}$ *P*-value of the MAGMA gene-based test (SNP-wise mean model) for 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.

https://doi.org/10.1038/s41588-026-02654-3



Extended Data Fig. 9 | Quantile–quantile plot of the PheWAS analysis in BioVU (n = 66,325), stratified by phecode category (1,431 tested phecodes). Observed two-sided $-\log_{10}$ P-values obtained from the logistic regression analysis are shown on the y-axis, and expected $-\log_{10}$ P-values are shown on the x-axis. Detailed results can be found in Supplementary Table 24.

https://doi.org/10.1038/s41588-026-02654-3



Extended Data Fig. 10 | Quantile–quantile plot of the PheWAS analysis in UKB (n = 316,635), stratified by phecode category (1,250 tested phecodes). Observed two-sided $-\log_{10}$ P-values obtained from the logistic regression analysis are shown on the y-axis, and expected $-\log_{10}$ P-values are shown on the x-axis. Detailed results can be found in Supplementary Table 25.

nature portfolio

Corresponding author(s): Fabian Streit

Last updated by author(s): Dec 1, 2025

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Data collection was not part of this study.
Data analysis	RICOPILI: https://sites.google.com/a/broadinstitute.org/ricopili/; https://github.com/Ripkelab/ricopili PLINK 1.9: http://pngu.mgh.harvard.edu/purcell/plink/ EAGLE 2.4.1: https://alkesgroup.broadinstitute.org/Eagle/ MINIMAC 3: http://genome.sph.umich.edu/wiki/minimac3 METAL: http://www.sph.umich.edu/csg/abecasis/metal/ MAGMA 1.07: http://ctglab.nl/software/magma FUMA 1.5.2: https://fuma.ctglab.nl/ Polygenic Priority Score (PoPS) v0.2: https://github.com/FinucaneLab/pops GREP: https://github.com/saorisakaue/GREP PRS-CS v.1.1.0: https://github.com/getian107/PRScs LDSC: https://github.com/bulik/ldsc DRUGSETS: https://github.com/nybell/drugsets

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

Individual-level data are not publicly available due to ethical restrictions. The results of the meta-analysis and PGS weights are publicly available via the website of the Psychiatric Genomics Consortium (PGC) 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 22). The exceptions are the GWAS summary statistics for self-harm104 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.

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	Biological sex was used in the study, and analyses stratified by sex are reported additionally.
Reporting on race, ethnicity, or other socially relevant groupings	All studies consisted of subjects of European ancestry. The analysis was controlled for effects of population stratification by a) excluding outliers based on ancestry principal components, and b) adjusting for ancestry principal components in the association analyses.
Population characteristics	Additional meta-analyses were calculated stratified by sex (female only, male only), and as a sensitivity analysis excluding subjects with comorbidity of schizophrenia or bipolar disorder. The number of female and male subjects in each analysis are listed in the Supplementary Tables S2-S5. Further details on the included populations can be found in Supplementary Table S1 and the Supplementary Note
Recruitment	Detailed descriptions of recruitment of all included 28 studies can be found in the Supplementary Table 1 ("study overview") and supplementary methods ("description of studies").
Ethics oversight	All subjects gave informed consent, and the studies were approved by the respective ethical committees (details see Supplementary Table 1 and Supplementary Note).

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Total sample size was defined as the sum of all participants included in each meta-analysis. To account for the disparity of cases and control groups, the effective sample size is additionally reported for each meta-analysis. To maximize power to detect genetic loci, we included all cohorts for which eligible data were available and collaboration was feasible; no a priori sample size calculation was performed. A post-hoc power analysis demonstrated >80% power to detect common variant effects (MAF 0.3-0.7) at genome-wide significance, indicating that the achieved sample size was sufficient for the study aims.
Data exclusions	We applied filtering criteria to samples and genetic markers (samples and SNPs with genotype call rate < 98%; difference in SNP genotyping call rate between cases and controls >2%, deviation of autosomal heterozygosity from the mean (|Fhet|>0.2), deviation from Hardy-Weinberg equilibrium (HWE; p<1×10−10 in cases; p<1×10−6 in controls); minor allele frequency <1% (5% for X chromosome); INFO score < 0.6; significant heterogeneity (p <0.001), effective sample size <85%). All analyses using individual-level genotype data were adjusted for ancestry principal components (GWAS, PGS analyses, PheWAS). In addition, PheWAS analyses and studies that provided summary statistics adjusted the analyses in their data sets for additional covariates where appropriate including age, sex, and technical or study-specific covariates (e.g., genotyping array, study center).
Replication	To maximize discovery, most available studies were included in the discovery analysis (12,339 cases and 1,041,717 controls). Two independent replication cohorts (n total: 685 cases and 107,750 controls) were used to replicate the lead loci, and polygenic risk score

prediction. All six genome-wide significant lead SNPs (6/6), and 77% of the lead SNPs associated with p<1x10-6 (24/31) showed effects in the same direction in the replication analysis. In the independent replication samples, PGS were associated with case-control status and explained 3.1% of the variance (p=0.0041, AUC=62.4%) in Spain 2, and 2.6% in All of Us (p=2.09x10-25, AUC=61.4%).

Randomization: There was no randomization in the used GWAS design.

Blinding: There was no blinding in the used (observational) design.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Plants

Seed stocks: *Report on the source of all seed stocks or other plant material used. If applicable, state the seed stock centre and catalogue number. If plant specimens were collected from the field, describe the collection location, date and sampling procedures.*

Novel plant genotypes: *Describe the methods by which all novel plant genotypes were produced. This includes those generated by transgenic approaches, gene editing, chemical/radiation-based mutagenesis and hybridization. For transgenic lines, describe the transformation method, the number of independent lines analyzed and the generation upon which experiments were performed. For gene-edited lines, describe the editor used, the endogenous sequence targeted for editing, the targeting guide RNA sequence (if applicable) and how the editor was applied.*

Authentication: *Describe any authentication procedures for each seed stock used or novel genotype generated. Describe any experiments used to assess the effect of a mutation and, where applicable, how potential secondary effects (e.g. second site T-DNA insertions, mosiacism, off-target gene editing) were examined.*