Article

Genomics yields biological and phenotypic insights into bipolar disorder

https://doi.org/10.1038/s41586-024-08468-9

Received: 22 December 2023

Accepted: 28 November 2024

Published online: 22 January 2025

Check for updates

Bipolar disorder is a leading contributor to the global burden of disease[1]. Despite high heritability (60–80%), the majority of the underlying genetic determinants remain unknown[2]. We analysed data from participants of European, East Asian, African American and Latino ancestries (n = 158,036 cases with bipolar disorder, 2.8 million controls), combining clinical, community and self-reported samples. We identified 298 genome-wide significant loci in the multi-ancestry meta-analysis, a fourfold increase over previous findings[3], and identified an ancestry-specific association in the East Asian cohort. Integrating results from fine-mapping and other variant-to-gene mapping approaches identified 36 credible genes in the aetiology of bipolar disorder. Genes prioritized through fine-mapping were enriched for ultra-rare damaging missense and protein-truncating variations in cases with bipolar disorder[4], highlighting convergence of common and rare variant signals. We report differences in the genetic architecture of bipolar disorder depending on the source of patient ascertainment and on bipolar disorder subtype (type I or type II). Several analyses implicate specific cell types in the pathophysiology of bipolar disorder, including GABAergic interneurons and medium spiny neurons. Together, these analyses provide additional insights into the genetic architecture and biological underpinnings of bipolar disorder.

Bipolar disorder (BD) is an often lifelong mood disorder that impairs quality of life, functional ability and is associated with suicidality[5]. Symptoms typically occur in early adulthood[5], with a similar prevalence and incidence rate across the world[6]. Current treatment options include pharmacotherapies such as mood stabilizers, antipsychotics and antidepressants, preferably administered in conjunction with psychosocial interventions[1,5]. However, approximately one-third of patients relapse within the first year of treatment[7].

The heterogeneous nature of the disorder is noted in the fifth edition of *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5), which includes the category 'bipolar and related disorders', encompassing BD type I (BDI), BD type II (BDII) and cyclothymic disorders[8]. The 11th revision of *International Classification of Diseases* (ICD-11) also recognizes BDI and BDII as distinct subtypes[9]. BDI is characterized by episodes of both mania and depression, whereas BDII includes episodes of hypomania and depression. Advances in genetics and neuroimaging have begun to make inroads into the underlying pathophysiology of BD. The Psychiatric Genomics Consortium (PGC) Bipolar Disorder Working Group has spearheaded genetic discoveries in BD[10,11]. A genome-wide association study (GWAS) of 41,917 individuals with BD and 371,549 control individuals identified 64 loci and highlighted calcium channel antagonists as potential targets for drug repurposing[3]. Brain imaging studies have shown decreased cortical thickness, lower subcortical volume and disrupted white matter integrity associated with BD, as well as brain alterations associated with medication use[12]. To date, this research has been conducted almost exclusively on individuals of European (EUR) ancestry.

Here we present the largest to date multi-ancestry GWAS meta-analysis of 158,036 individuals with BD and 2,796,499 control individuals, combining clinical, community and self-reported samples. We identified 337 linkage disequilibrium-independent genome-wide significant variants that map to 298 loci. We hypothesized that differences in source of patient ascertainment, BD subtype and genetic ancestry might lead to differences in genetic architecture, thus we also analysed these groups separately. We provide new insights into the genetic architecture and neurobiological mechanisms involved in BD, with the potential to inform the development of new treatments and precision medicine approaches.

Study population

The current GWAS meta-analysis includes 79 cohorts. Case definitions were based on a range of assessment methods: (semi-)structured clinical interviews (clinical), medical records, registries and questionnaire data (community) and self-reported surveys (self-reported). Details of the cohorts, including sample size, ancestry, and inclusion or exclusion criteria for individuals, are provided in Supplementary Tables 1 and 2 and the Supplementary Note. BD subtype data were available for a subset of individuals within the clinical and community groups. Of individuals with BD (cases), 82.5% in the clinical ascertainment group had BDI, as did 68.7% of individuals in the community ascertainment group ($\chi^2 = 730$, $P < 2.2 \times 10^{-16}$; Supplementary Table 2). The total number of samples available for analyses included 158,036 cases with BD and 2,796,499 controls (effective n (n_{eff}) = 535,720; see Methods).

Genetic architecture of BD

Given our hypothesis that samples ascertained and assessed by different methods could lead to differences in the genetic architecture, we performed meta-analyses separately for clinical, community and

A list of authors and their affiliations appears online. ✉e-mail: k.s.oconnell@medisin.uio.no; ole.andreassen@medisin.uio.no

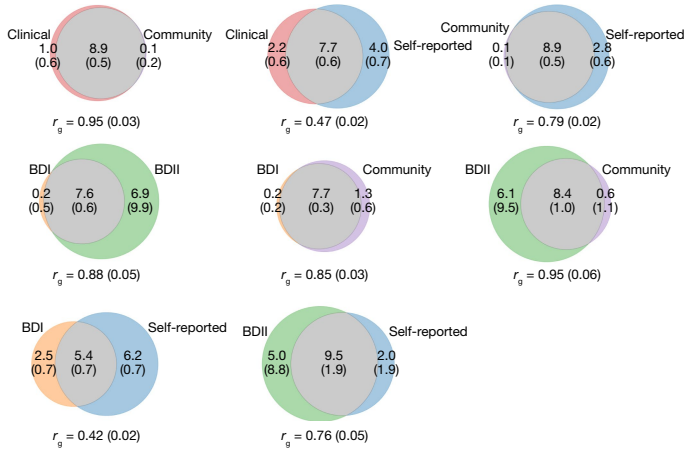


Fig. 1 | Genetic correlation and bivariate MiXeR estimates for the genetic overlap of BD ascertainment and subtypes. Trait-influencing genetic variants shared between each pair (grey) and unique to each trait (colours) are shown. The numbers within the Venn diagrams indicate the estimated number of trait-influencing variants (and standard errors; in thousands) that explain 90% of SNP-h^2 in each phenotype. The size of the circles reflects the polygenicity of each trait, with larger circles corresponding to greater polygenicity. The estimated genetic correlation (r_g) and standard error between BD and each trait of interest from LDSC are shown below the corresponding Venn diagram. Clinical and community samples were stratified into BDI and BDII subtypes if subtype data were available. Model fit statistics indicated that MiXeR-modelled overlap for bivariate comparisons including the BD subtypes (BDI and BDII) were not distinguishable from minimal or maximal possible overlap, and therefore are to be interpreted with caution (see Supplementary Table 4).

self-reported samples. Using linkage disequilibrium score regression (LDSC)[13] and assuming a population prevalence of 2%[14], BD ascertained from clinical samples was more heritable (single-nucleotide polymorphism heritability (SNP-h^2) = 0.22; s.e. = 0.01) than BD ascertained from community samples (SNP-h^2 = 0.05; s.e. = 0.003) or self-reported (SNP-h^2 = 0.08; s.e. = 0.003; Supplementary Table 3). We used genetic correlation[13] and MiXeR[15,16] analyses to further investigate the genetic architecture of BD based on assessment. Although there was a strong genetic correlation (r_g) between clinical and community samples (r_g = 0.95; s.e. = 0.03), the genetic correlation for self-reported BD was significantly greater ($P = 7.4 \times 10^{-28}$) with community samples (r_g = 0.79; s.e. = 0.02) than with clinical samples (r_g = 0.47; s.e. = 0.02; Extended Data Fig. 1).

MiXeR estimated the greatest polygenicity for BD ascertained from self-reported samples, followed by clinical and then community samples (Fig. 1 and Supplementary Table 4). BD in clinical samples was estimated to be the most discoverable, whereas self-reported BD had the lowest discoverability (Extended Data Fig. 2 and Supplementary Table 4). Almost all variants estimated to influence BD in community samples were shared with BD ascertained from clinical samples. The majority of clinical and community BD-influencing variants were also shared with self-reported BD (Fig. 1 and Extended Data Fig. 3). The mean correlation of variant effects in the shared components was high across all groups (community and self-reported $r_{g\,shared}$ = 0.95 (s.e. = 0.03), community and clinical $r_{g\,shared}$ = 0.99 (s.e. = 0.01) and clinical and self-reported $r_{g\,shared}$ = 0.74 (s.e. = 0.06; Supplementary Table 4).

To analyse BD subtypes, we used available GWAS summary statistics for BDI (25,060 individuals) and BDII (6,781 individuals)[3], which come from a subset of the clinical and community samples. Assuming a population prevalence of 1%[17], BDI was more heritable (SNP-h^2 = 0.21; s.e. = 0.01) than BDII (SNP-h^2 = 0.11; s.e. = 0.01). BDI and BDII were highly, but imperfectly, correlated (r_g = 0.88; s.e. = 0.05). The genetic correlations between both subtypes and the community samples were high (BDI r_g = 0.85; s.e. = 0.03, BDII r_g = 0.95; s.e. = 0.06). By contrast, the genetic correlation between BDI and self-reported BD (r_g = 0.42; s.e. = 0.02) was significantly lower ($P = 7.1 \times 10^{-13}$) than between BDII and self-reported BD (r_g = 0.76; s.e. = 0.05; Extended Data Fig. 1).

Given the difference in proportion of individuals with BDI and BDII within the clinical and community cohorts, we evaluated the genetic correlation between BD within clinical and community cohorts, and self-reported BD, after conditioning on the genetic risk for BDI and BDII. After conditioning, the genetic correlation between self-reported BD and BD within community cohorts (r_g = 0.92; s.e. = 0.09) was not significantly different (P = 0.10) than BD in clinical cohorts (r_g = 0.71; s.e. = 0.13).

We show that genetic architecture is different across ascertainment and subtypes, and that these differences appear to be driven by the proportion of BD subtype within the sample. Despite these observed differences, the high correlations of variant effects in the shared components across ascertainment groups support our decision to use a meta-analysis for all BD cases.

Ancestry-specific GWAS meta-analyses

We conducted separate meta-analyses in four ancestral groups. Because the self-reported data differed in genetic architecture from the clinical and community data, we performed separate meta-analyses with and without the inclusion of the self-reported data. Supplementary Table 2 provides a summary of the GWAS meta-analyses, and details of associated loci are described in Supplementary Tables 5–7. Ancestry-specific estimates of SNP heritability and cross-ancestry genetic correlations are provided in Supplementary Table 3.

We identified 261 independent genome-wide significant variants mapping to 221 loci associated with BD in EUR ancestry meta-analyses that included self-reported data, and 94 independent genome-wide significant variants mapping to 88 loci without self-reported data (Supplementary Tables 5 and 6). There were 92 of the 94 independent genome-wide significant variants available for meta-analysis in the self-reported cohorts, of which 78 (85%) were concordant for direction of effect (Supplementary Table 6).

In the East Asian (EAS) ancestry meta-analysis, we identified two BD-associated loci, one of which is novel with an ancestry-specific index variant (rs117130410, chromosome 4: 105734758, build GRCh37; Extended Data Fig. 4 and Supplementary Table 7). Although this variant had a frequency of 16% and 9% in EAS individuals with BD and controls, respectively, it is monomorphic in non-Asian populations. The second locus (rs174576, chromosome 11: 61603510, build GRCh37; Supplementary Table 7) was only identified when the self-reported data were excluded from the meta-analysis as the index variant was not available in the self-reported data. This locus has been identified previously and implicates the *FADS1* and *FADS2* genes[3,18]. No genome-wide significant loci were observed in the African American (AFR) or Latino (LAT) ancestry meta-analyses.

Multi-ancestry meta-analysis

A multi-ancestry meta-analysis of all the datasets identified 337 linkage disequilibrium-independent genome-wide significant variants mapping to 298 loci (Extended Data Fig. 4 and Supplementary Table 8). There was minimal test statistic inflation after correction for principal components in each dataset (LDSC intercept = 1.052 (s.e. = 0.016), attenuation ratio = 0.071 (s.e. = 0.013)).

Of the 298 loci identified in this multi-ancestry meta-analysis, 267 are novel for BD. Of the 64 previously reported BD-associated loci[3], 31 met genome-wide significance in the present analysis containing all samples, and of the 33 that did not, 25 met genome-wide significance in either the clinical samples or in the meta-analysis that excluded self-reported data (Supplementary Table 9). Moreover, the direction

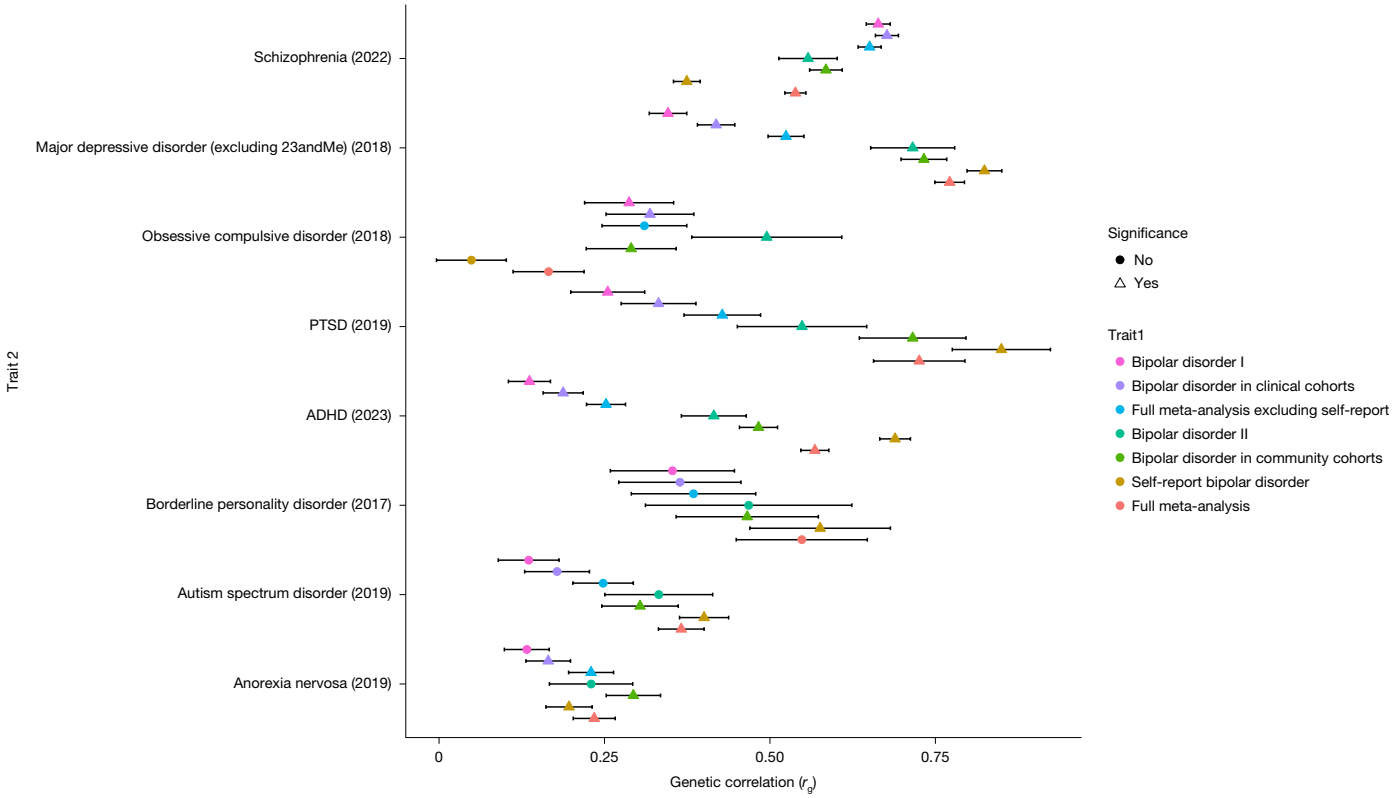


Fig. 2 | Genetic correlations (with standard errors) between BD and other psychiatric disorders. The *y* axis (trait 2) is ordered based on the significance and magnitude of genetic correlation of each trait with BDI. *P* values were calculated from the two-sided *z*-statistics computed by dividing the estimated genetic correlation by the estimated standard error, without adjustment. The standard error for a genetic correlation was estimated using a ratio block jackknife over 200 blocks. The triangles indicate significant results passing the Bonferroni-corrected significance threshold of two-sided $P < 3.6 \times 10^{-5}$. Error bars represent the standard error of the estimate. The year indicated in parentheses after each trait refers to the year in which the GWAS was published. Details are provided in Supplementary Table 13. PTSD, post-traumatic stress disorder.

of association for all top SNPs (12,151 SNPs with $P < 1 \times 10^{-5}$) from the previous GWAS was consistent with the direction of association in this multi-ancestry meta-analysis of all samples (Supplementary Table 9).

When considering the effect of ancestry on the discovery of these 298 loci, one locus (index SNP rs7248481, chromosome 19: 13079957–13122567) was most strongly associated in the EAS ancestry meta-analysis. For all other loci, the association was strongest in the EUR ancestry meta-analysis. The majority of the 298 loci were nominally significant ($P < 0.05$) within the AFR (290 of 298 loci), EAS (257 of 298 loci) and LAT (293 of 298 loci) ancestry-specific meta-analyses, highlighting consistency of signal across the ancestry groups (Supplementary Table 8).

We estimated the proportion of SNP-h^2 accounted for by SNPs within genome-wide significant loci[19]. Compared with only 8.3% accounted for by SNPs within the 64 previously identified loci[3], SNPs within the 298 loci account for 18.5% of the SNP-h^2 of BD (Supplementary Table 10). Moreover, SNPs within the 298 loci also accounted for higher proportions of SNP-h2 in the clinical (64 loci: 8.5%; 298 loci: 17.8%), BDI (64 loci: 8.3%; 298 loci: 17.5%), community (64 loci: 4.8%; 298 loci: 22.6%) and self-reported (64 loci: 2.0%; 298 loci: 21.1%) samples.

We carried out sensitivity meta-analyses excluding the self-reported samples (leaving 67,948 cases and 867,710 controls; $n_{eff} = 191{,}722$) and identified 116 independent genome-wide significant variants mapping to 105 loci (Supplementary Table 11). There was minimal test statistic inflation due to uncontrolled population stratification after correction for principal components in each dataset (LDSC intercept = 1.050; s.e. = 0.012, attenuation ratio = 0.086; s.e. = 0.018). Analysis of self-reported cohorts only (90,088 cases and 1,928,789 controls; $n_{eff} = 344{,}088$) identified 126 loci (Supplementary Table 12). Of the 116 independent genome-wide significant variants identified in the meta-analysis excluding the self-reported samples, 110 variants were available for meta-analysis in the self-reported samples, of which 96 (87%) were concordant (Supplementary Table 11).

Our multi-ancestry meta-analysis identified 298 loci, implicating 337 linkage disequilibrium-independent genome-wide significant variants.

Genetic correlations with other traits

Genome-wide genetic correlations were estimated between EUR ancestry BD GWAS (with and without self-reported data, and when stratified by ascertainment and subtypes) and human diseases and traits via the Complex Traits Genetics Virtual Lab (https://vl.genoma.io) web platform[20] (Fig. 2 and Supplementary Tables 13–15). Most psychiatric disorders, including major depressive disorder, post-traumatic stress disorder, attention deficit–hyperactivity disorder (ADHD), borderline personality disorder and autism spectrum disorder, were more strongly correlated with the full meta-analysis, BDII and BD in community and self-reported samples, than with BDI and BD in clinical cohorts (Fig. 2). By contrast, schizophrenia was more strongly genetically correlated with the full BD meta-analysis excluding self-reported data and with BDI and BD in clinical samples (Fig. 2). This pattern of correlations, together with the observed patterns of genetic architecture, suggest that the self-reported samples include a high proportion of people with BDII.

Polygenic association with BD

Polygenic risk score (PRS) analyses were performed using PRS-CS-auto[21] in 55 EUR ancestry cohorts for which individual-level genotype and phenotype data were available (40,992 cases and 80,215 controls), as well as one cohort of AFR ancestry (347 cases and 669 controls) and

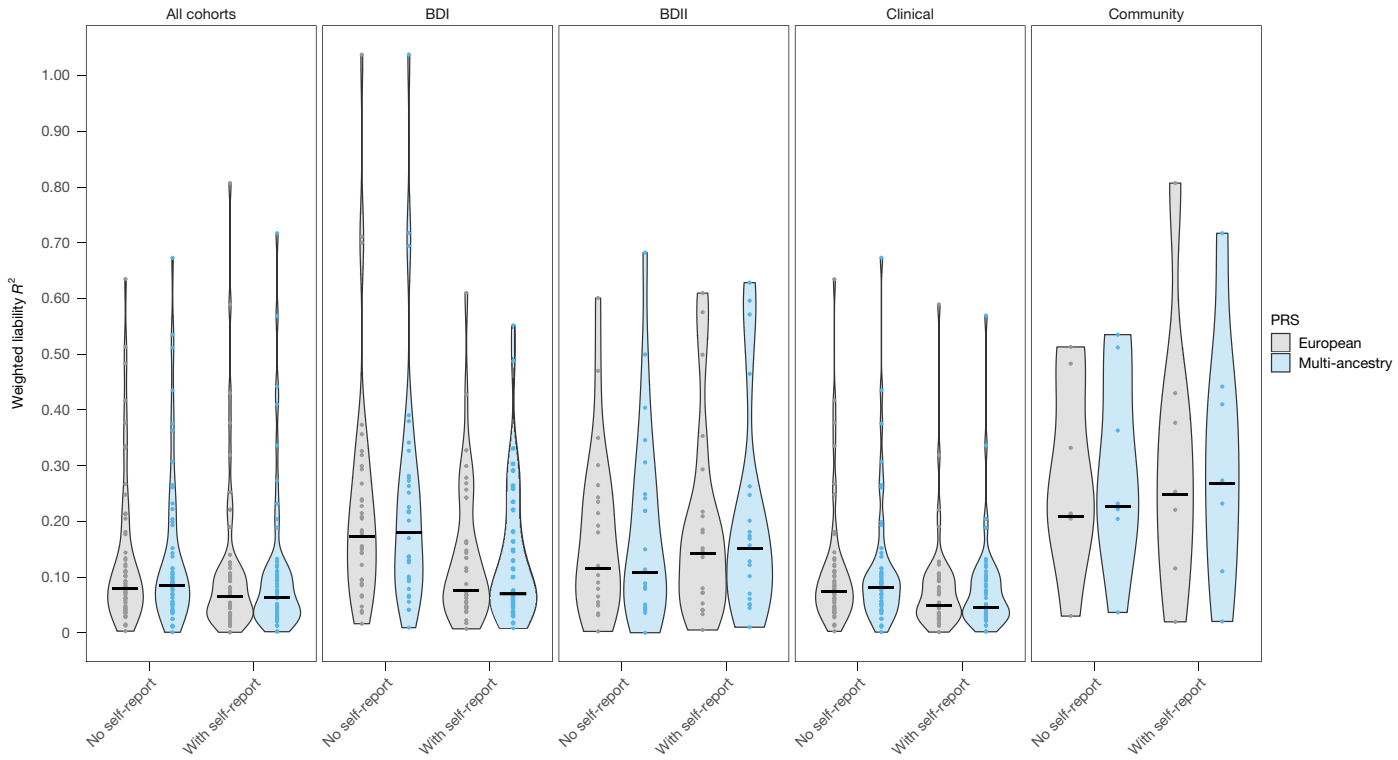


Fig. 3 | Phenotypic variance in BD in EUR cohorts explained by PRSs derived from the multi-ancestry and EUR meta-analyses (with and without self-reported data). Variance explained is presented on the liability scale, assuming a 2% population prevalence of BD. The results (all cohorts) are the median weighted liability R^2 values across all 55 EUR cohorts (40,992 cases and 80,215 controls; n_{eff} = 46,725). Similarly, BDI, BDII, clinical and community panels show the results across 36 BDI cohorts (12,419 cases and 33,148 controls; n_{eff} = 14,607), 21 BDII cohorts (2,549 cases and 23,385 controls; n_{eff} = 4,021), 48 clinical cohorts (27,833 cases and 46,623 controls; n_{eff} = 29,543) and 7 community cohorts (13,159 cases and 36,592 controls; n_{eff} = 17,178). All analyses were weighted by the effective *n* per cohort. The median liability R^2 is represented as a horizontal black line.

three cohorts of EAS ancestry (4,473 cases and 65,923 controls; Supplementary Tables 16–20). In the EUR ancestry cohorts, the variance explained by the multi-ancestry GWAS without the self-reported data (R^2 = 0.090, s.e. = 0.019) was significantly greater than that explained by both the multi-ancestry GWAS including self-reported data (R^2 = 0.058, s.e. = 0.017, $P = 2.72 \times 10^{-4}$) and the EUR ancestry GWAS excluding the self-reported data (R^2 = 0.084, s.e. = 0.018, $P = 5.62 \times 10^{-3}$; Fig. 3a and Supplementary Tables 16 and 21). Individuals in the top quintile (top 20%) for this multi-ancestry GWAS without the self-reported data PRS had an odds ratio of 7.06 (95% CI = 3.9–10.4) of being affected with BD compared with individuals in the middle quintile. The corresponding median area under the receiver operating characteristic curve was 0.70 (95% CI = 0.67–0.73). Therefore, the BD liability explained remains insufficient for diagnostic prediction in the general population.

Similarly, PRS derived from GWAS excluding self-reported data explained significantly more variance in cases of BDI (Fig. 3b and Supplementary Tables 17) and in clinical cohorts (Fig. 3d and Supplementary Tables 19) than when self-reported data were included. Conversely, inclusion of the self-reported data yielded greater median R^2 estimates for the PRS in cases of BDII (Fig. 3c and Supplementary Tables 18) and in community cohorts (Fig. 3e and Supplementary Tables 20); however, these differences were not significant. These results are probably due to increased phenotypic heterogeneity when the self-reported data were included in the PRS discovery sample (see Fig. 2).

PRS analysis of three clinically ascertained EAS cohorts revealed that the PRSs derived from GWAS excluding the self-reported data (Taiwan: EUR ancestry PRS (EUR-PRS) R^2 = 0.069, multi-ancestry PRS (multi-PRS) R^2 = 0.075; Japan: EUR-PRS R^2 = 0.027, multi-PRS R^2 = 0.025; Korea: EUR-PRS R^2 = 0.016, multi-PRS R^2 = 0.022) performed better than those that included self-reported data (Taiwan: EUR-PRS R^2 = 0.026, multi-PRS R^2 = 0.036; Japan: EUR-PRS R^2 = 0.015, multi-PRS R^2 = 0.015; Korea: EUR-PRS R^2 = 0.014, multi-PRS R^2 = 0.017; Supplementary Table 22).

In a clinically ascertained AFR target cohort, the inclusion of self-reported data increased the explained variance (R^2) by both the multi-PRS and the EUR-PRS from 0.010 to 0.23 or 0.22, respectively (Supplementary Table 22).

Pathway, tissue and cell-type enrichment

Gene set enrichment analyses were performed on the summary statistics derived from the multi-ancestry meta-analysis including self-reported data, using MAGMA[22]. We identified significant enrichment of six gene sets (Supplementary Table 23) related to synapse and transcription factor activity. The association signal was enriched among genes expressed in the brain (Supplementary Table 24), and specifically in the early-to-mid-prenatal stages of development (Supplementary Table 25). Single-cell enrichment analyses of brain cell types indicate involvement of neuronal populations from different brain regions, including hippocampal pyramidal neurons and interneurons of the prefrontal cortex and hippocampus (Supplementary Fig. 1), and were largely consistent with findings from the previous PGC-BD GWAS[3]. Similar patterns of enrichment were observed based on ascertainment and subtype (Supplementary Fig. 2). In addition, GSA-MiXeR[19] highlighted enrichment of specific dopamine-related and calcium-related biological processes and molecular functions, as well as GABAergic interneuron development, respectively (Supplementary Table 26).

A recent study[23] has analysed single-nucleus RNA sequencing data of 3.369 million nuclei from 106 anatomical dissections within 10 brain regions and divided cells into 31 superclusters and 461 clusters, respectively, based on principal component analysis of sequenced genes.

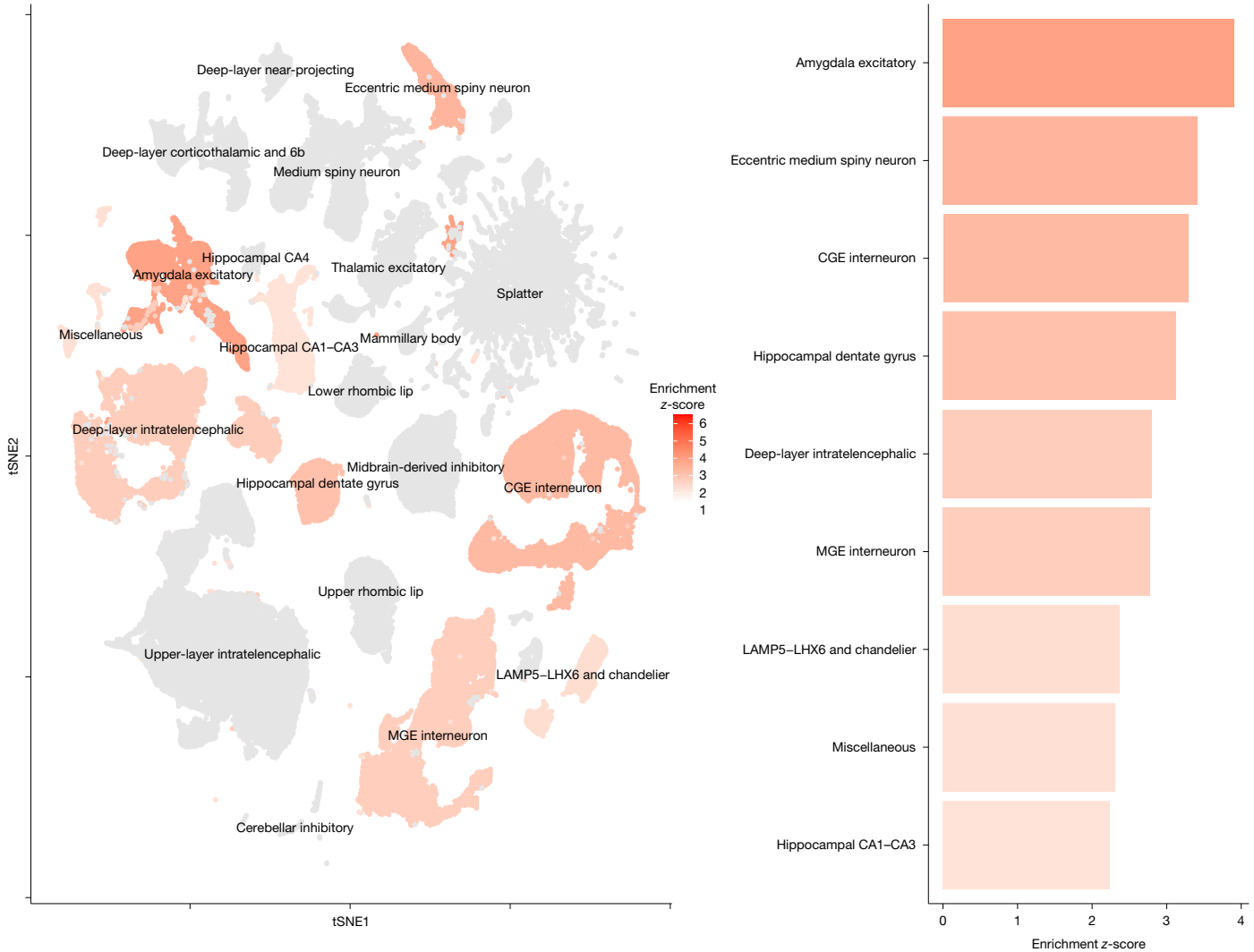


Fig. 4 | Supercluster-level SNP-h^2 enrichment for BD. The *t*-distributed stochastic neighbour embedding (tSNE) plot (naming convention and source of the single-cell data from Siletti et al.[23]; left) is coloured by the enrichment *z*-score. Grey indicates non-significantly enriched superclusters (false discovery rate > 0.05). The bar plot (right) shows the nine significantly enriched superclusters. CGE, caudal ganglionic eminence; MGE, medial ganglionic eminence.

These superclusters were then annotated based on their regional composition within the brain (Fig. 4). We used stratified LDSC[24] to estimate SNP-h^2 enrichment for the top decile of expression proportion genes in each of the 31 superclusters and 461 clusters, as previously described[25]. Heritability was significantly enriched in 9 of the 31 superclusters (Fig. 4), and 49 of the 461 clusters (Extended Data Fig. 5). No enrichment was seen in non-neuronal clusters. Two clusters within the medium spiny neurons, not observed at the supercluster level, were significantly enriched, further supporting the involvement of striatal processes in BD.

Together, these results implicate the synapse, interneurons of the prefrontal cortex and hippocampus, and hippocampal pyramidal neurons as particularly relevant in the molecular biology of BD.

Single-cell enrichment analysis in 914 cell types across 29 non-brain mouse tissues identified significant enrichment in the enteroendocrine cells of the large intestine and delta cells of the pancreas, which remained significant after cross-dataset conditional analyses with a mouse brain tissue dataset (Supplementary Table 27).

Fine-mapping

We performed functional fine-mapping using Polyfun+SuSiE[26] (Supplementary Tables 28 and 29). At a threshold of posterior inclusion probability (PIP) > 0.50, we identified 80 putatively causal fine-mapped SNPs for the multi-ancestry meta-analyses including self-reported data. At the more stringent threshold of PIP > 0.95, we identified 20 putatively causal SNPs. When comparing the number of SNPs within 95% credible sets, the inclusion of multi-ancestry and self-reported data led to smaller credible sets (that is, credible sets with fewer numbers of SNPs). For example, we identified 175 95% credible sets of less than 20 SNPs in the multi-ancestry dataset with self-reported data, compared with 122 in the EUR dataset with self-reported data (Extended Data Fig. 6). Putatively causal SNPs with a PIP > 0.5 were mapped to genes by performing variant annotation with Variant Effect Predictor (VEP; GRCh37) Ensembl release 109 (ref. 27), based on their position relative to annotated Ensembl transcripts and known regulatory features. This analysis identified 71 unique genes annotated to fine-mapped SNPs from the multi-ancestry meta-analysis including self-reported data (Supplementary Table 29).

Common and rare variation convergence

Within loci associated with BD in the multi-ancestry meta-analysis, the 71 genes annotated to putatively causal fine-mapped SNPs (Supplementary Table 29) were enriched for ultra-rare (5 or less minor allele

count) damaging missense and protein-truncating variants in cases of BD in the Bipolar Exome (BipEx) consortium dataset[4] (OR = 1.16, 95% CI = 1.05–1.28, $P = 0.002$), and in cases of schizophrenia in the Schizophrenia Exome Meta-analysis (SCHEMA) dataset[28] (OR = 1.21, 95% CI = 1.02–1.43, $P = 0.024$). This enrichment is similar to that observed for schizophrenia[28] and ADHD[29].

Credible BD-associated genes

In addition to the 71 genes annotated to the fine-mapped putatively causal SNPs as described above, we annotated a further 45 genes to the 80 fine-mapped SNPs by summary data-based Mendelian randomization using expression quantitative trait locus (eQTL) and splicing QTL (sQTL) data, as well as by proximity, that is, the nearest gene to each SNP (Extended Data Fig. 7 and Supplementary Tables 30 and 31). No genes were annotated to the CpGs identified by the methylation QTL (mQTL) analysis (Supplementary Table 30). We then determined whether any of these 116 genes were also identified through the genome-wide gene-based analysis using MAGMA[22], eQTL analyses using transcriptome-wide association study (TWAS) as implemented in FUSION[30] and isoTWAS[31], or through enhancer–promoter interactions[32,33]. This resulted in seven possible approaches by which loci could be mapped to genes, including eQTL evidence (eQTL or TWAS or FOCUS or isoTWAS), mQTL, sQTL, VEP, proximity, MAGMA and enhancer–promoter interactions.

We integrated the results from the post-GWAS analyses described above and identified a credible set of 36 genes identified by at least three of the described approaches (Supplementary Table 31). The *SP4* gene was identified by six of these approaches, and astrocyte and GABAergic neuron-specific regulation of *SP4*, by the genome-wide significant variant rs2107448, was identified from cell-type-specific enhancer–promoter interaction results (Supplementary Table 31). Moreover, the *TTC12* and *MED24* genes were identified by five of the approaches. Eight of the 36 credible genes have synaptic annotations in the SynGO database[34]. Three genes (*HTT*, *ERBB4* and *LR5NF*) were mapped to both postsynaptic and presynaptic compartments. One gene (*CACNA1B*) was mapped to only the presynapse, and four genes (*SHANK2*, *OLFM1*, *SHISA9* and *SORCS3*) were mapped to only the postsynapse (Supplementary Table 32).

On the basis of the lifespan gene expression data from the Human Brain Transcriptome project (www.hbatlas.org)[35], suggestive evidence for two clusters of credible genes was observed based on temporal expression (Extended Data Fig. 8 and Supplementary Table 31). The first cluster showed reduced prenatal gene expression, with gene expression peaking at birth and remaining stable over the life course. Conversely, the second cluster showed a peak in gene expression during fetal development with a drop-off in expression before birth. However, both clusters showed high variability in gene expression across the lifespan.

Together, these results implicate 36 credible genes in BD.

Drug target analyses

Gene set analyses were performed restricted to genes targeted by drugs, assessing individual drugs and grouping drugs with similar actions, as previously described[3,36]. Gene-level and gene set analyses of the multi-ancestry GWAS summary statistics including self-reported data were performed in MAGMA[22], and identified significant enrichment in the targets of anticonvulsant pregabalin (Supplementary Table 33). There was also significant enrichment in the targets of antipsychotics and anxiolytics (Supplementary Table 34).

Examination of the Drug Gene Interaction Database (DGIdb)[37] to identify drug–gene interactions using the credible genes as input genes showed that 15 out of 36 genes were interacting with a total of 528 drugs. Gene set enrichment analysis of these drug–gene interactions showed a significant enrichment ($P < 0.0001$) for targets of the atypical antipsychotic drugs nemonapride and risperidone (Supplementary Table 35). However, after correction for the total number of drugs ($n = 69{,}018$), the enrichment was not significant (false discovery rate > 0.05). In addition, 16 of the 36 credible genes had evidence of tractability with a small molecule in the OpenTargets dataset, including *FURIN*, *MED24*, *THRA*, *ALDH2*, *ANKK1*, *ARHGAP15*, *CACNA1B*, *ERBB4*, *ESR1*, *FES*, *GPR139*, *HTT*, *MLEC*, *MSH6*, *PSMD14* and *TOMM2*.

Among the 36 credible genes, two genes (*ALDH2* and *ESR1*) were within the list of 139 lithium target and interaction partner genes. The results of the network-based separation (S_{AB}) analysis do not indicate a general overlap between the credible genes and lithium target genes in the human protein interactome ($S_{AB} = 0.124$, $z = 1.710$, $P = 0.044$). The positive S_{AB} value indicates that the lithium target genes and the 36 credible genes are separated from each other in the network of protein–protein interactions.

As the credible gene list is primarily derived from our fine-mapping analysis, it is possible that lithium target genes (and interaction partners) are within loci for which significant fine-mapped putatively causal SNPs were not identified. The identification of evidence of tractability with small molecules for some of the credible genes indicates opportunities for novel drug development.

Discussion

We performed, to our knowledge, the largest GWAS of BD, including diverse samples of EUR, EAS, AFR and LAT ancestry, resulting in an over fourfold increase in the number of BD-associated loci: 337 linkage disequilibrium-independent genome-wide significant variants mapping to 298 loci. In the meta-analysis of EUR, the largest ancestry group, we identified over 200 genome-wide significant loci. We also found a novel ancestral-specific association in the EAS cohort. We confirmed our hypothesis that differences in ascertainment and BD subtype might lead to differences in genetic architecture. Post-GWAS analyses provide novel insights into the biological underpinnings and genetic architecture of BD and highlight differences depending on ascertainment of participants and BD subtype. We also showed that multi-ancestry data improved fine-mapping and polygenic prediction.

Enrichment of the common variant associations from this multi-ancestry meta-analysis highlights the synapse, interneurons of the prefrontal cortex and hippocampus, and hippocampal pyramidal neurons as particularly relevant. Exploratory analyses[19] suggest enrichment of dopamine-related and calcium-related biological processes and development of GABAergic interneurons. These findings were further corroborated by enrichment analyses in single-nucleus RNA sequencing data from adult post-mortem brain tissue, which highlighted specific clusters of interneurons derived from the caudal and medial ganglionic eminences and medium spiny neurons predominantly localized in the striatum. Medium spiny neurons are not enriched in depression using the same dataset[25]. Although interneurons derived from ganglionic eminences were also enriched in schizophrenia, stronger signals were observed for amygdala excitatory and hippocampal neurons[25].

A novel finding is that single-cell enrichment analysis of non-brain mouse tissues identified significant enrichment in the enteroendocrine cells of the large intestine and delta cells of the pancreas. Conditional analyses suggest that this enrichment is independent of overlapping genes between these cell types and those expressed in neurons. Stimulation of enteroendocrine cells by short-chain fatty acids promotes serotonin production in the colon, which leads to enhanced levels of serotonin in systemic circulation and in the brain, and is a proposed mechanism by which microbiota influence the gut–brain axis[38,39]. Of note, lithium treatment is shown to upregulate short-chain fatty acid-producing bacteria, highlighting a potential mechanism of action[40].

We mapped genes to the 80 putatively causal SNPs identified from fine-mapping based on seven complementary approaches

and identified a subset of 36 credible genes implicated by at least three of these approaches. The top credible gene, identified by six gene-mapping approaches, was *SP4*, which has also been implicated in schizophrenia through both rare[28] and common[41] variation. Moreover, we clustered the credible genes based on similar patterns of temporal variation in expression over the lifespan and found suggestive evidence for two clusters. Although within cluster gene expression was highly variable across the lifespan, the second cluster had a peak in expression during fetal development aligning with the neurodevelopmental hypothesis of mental disorders[42]. Genes prioritized through fine-mapping were shown to be enriched for ultra-rare damaging missense and protein-truncating variation in the BipEx[4] and SCHEMA[28] datasets, respectively, highlighting convergence of common and rare variant signals as recently shown in schizophrenia[41].

We identified differences in the genetic architecture of BD subtypes related to ascertainment. BD within clinical and community samples was highly but imperfectly correlated, with varying correlations with self-reported BD. The low genetic correlation and minimal genetic overlap between cases ascertained through clinical studies and cases with self-reported BD are driven by a greater proportion of BDI within the clinical and community samples. In line with these results, PRS derived from meta-analyses excluding the self-reported data performed better in clinical and BDI target samples, whereas the inclusion of self-reported data improved the PRS in community and BDII target samples. Moreover, the pattern of correlations between BD and other psychiatric disorders differed with the inclusion of self-reported data. Schizophrenia had the highest genetic correlation with BD without the inclusion of the self-reported data, whereas major depressive disorder was most strongly correlated with BD after the inclusion of the self-reported data. These results suggest that the self-reported samples may include a high proportion of people with BDII. Moreover, this is in line with recent findings in individuals diagnosed with BDII, which showed increasing PRS for depression and ADHD and decreasing PRS for BD over time[43]. However, BD in the outpatient setting may be overdiagnosed in people with conditions such as chronic depression or borderline personality disorder, highlighting a higher rate of comorbid disorders and potential for 'overdiagnosis' of BD within cohorts of this nature[44,45]. We showed that the differences in genetic architecture and phenotypic proportions of the clinical, community and self-reported cohorts with BD affected the replication of previous BD-associated loci. Previously associated loci that fell short of meeting genome-wide significance in the current study were genome-wide significant in the clinical samples and in the meta-analyses that excluded self-reported data, and all top SNPs (12,151 SNPs with $P < 1 \times 10^{-5}$) from the previous GWAS were consistent in direction of association in this multi-ancestry meta-analysis of all samples (Supplementary Table 9).

Investigation of the novel ancestral-specific association in the EAS ancestry meta-analysis in the GWAS Catalog[46] highlights overlaps with genome-wide significant loci for reduced sleep duration[47] and lower educational attainment[48], as well as a suggestive locus ($P < 2 \times 10^{-6}$) for the interaction between cognitive function and major depressive disorder[49]. These findings suggest a role for this genomic region in complex brain-related phenotypes.

The multi-ancestry PRS provided the greatest improvement over the EUR-PRS in two of the three EAS ancestry target cohorts (Korea and Taiwan). More subtle improvements were seen when the EUR target cohorts were analysed. Multi-ancestry training data provided little improvement in the AFR target cohort, which may be due to the genetic heterogeneity of this target cohort[50]. These results highlight the benefits of multi-ancestry representation in the PRS training data, in line with findings from other diseases[51]. The predictive power of this BD PRS shows a substantial improvement compared with previous findings[3]; however, this BD PRS alone still falls short of clinical utility[52].

One limitation is the lack of in-sample linkage disequilibrium estimates for all cohorts due to a lack of in-house raw genotype data for some cohorts. For instance, analysis of the MHC/C4 locus was not considered as the number of samples for which individual-level genotype data were accessible did not increase much since the previous analysis[3]. We used a EUR linkage disequilibrium reference panel to analyse the multi-ancestry meta-analyses[53] in which linkage disequilibrium patterns and interindividual heterogeneity within the ancestry groups are not fully captured. Another limitation is the inclusion of samples with minimal phenotyping. Although this allowed us to achieve large sample sizes, especially in under-represented non-EUR ancestry cohorts, and greatly increase the number of loci identified, minimally phenotyped samples have some shortcomings. For example, minimal phenotyping may result in low specificity association signals, as shown in major depression[54,55], and individuals in community-based biobanks may represent those less severely affected, as shown in schizophrenia[56].

In conclusion, in this large-scale multi-ancestry GWAS of BD, we identified 298 significant BD-associated loci, from which we have demonstrated convergence of common variant associations with rare variant signals and highlight 36 genes credibly implicated in the pathobiology of the disorder. We identified differences in the genetic architecture of BD based on ascertainment and subtype, suggesting that stratification by subtype will be important in BD genetics moving forwards. Several analyses implicate specific cell types in BD pathophysiology, including GABAergic interneurons and medium spiny neurons, as well as the enteroendocrine cells of the large intestine and delta cells of the pancreas. Enrichment of dopamine-related and calcium-related biological processes were also identified, further contributing to our understanding of the biological aetiology of BD.

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/s41586-024-08468-9.

1. Carvalho, A. F., Firth, J. & Vieta, E. Bipolar disorder. *N. Engl. J. Med.* **383**, 58–66 (2020).
2. Lichtenstein, P. et al. Common genetic determinants of schizophrenia and bipolar disorder in Swedish families: a population-based study. *Lancet* **373**, 234–239 (2009).
3. Mullins, N. et al. Genome-wide association study of more than 40,000 bipolar disorder cases provides new insights into the underlying biology. *Nat. Genet.* **53**, 817–829 (2021).
4. Palmer, D. S. et al. Exome sequencing in bipolar disorder identifies AKAP11 as a risk gene shared with schizophrenia. *Nat. Genet.* **54**, 541–547 (2022).
5. McIntyre, R. S. et al. Bipolar disorders. *Lancet* **396**, 1841–1856 (2020).
6. Aghababaie-Babaki, P. et al. Global, regional, and national burden and quality of care index (QCI) of bipolar disorder: a systematic analysis of the Global Burden of Disease Study 1990 to 2019. *Int. J. Soc. Psychiatry* **69**, 1958–1970 (2023).
7. Geddes, J. R. & Miklowitz, D. J. Treatment of bipolar disorder. *Lancet* **381**, 1672–1682 (2013).
8. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders (DSM-5®)* (American Psychiatric Association, 2013).
9. Reed, G. M. et al. Innovations and changes in the ICD-11 classification of mental, behavioural and neurodevelopmental disorders. *World Psychiatry* **18**, 3–19 (2019).
10. Psychiatric GWAS Consortium Bipolar Disorder Working Group. Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near ODZ4. *Nat. Genet.* **43**, 977–983 (2011).
11. Stahl, E. A. et al. Genome-wide association study identifies 30 loci associated with bipolar disorder. *Nat. Genet.* **51**, 793–803 (2019).
12. Ching, C. R. K. et al. What we learn about bipolar disorder from large-scale neuroimaging: findings and future directions from the ENIGMA Bipolar Disorder Working Group. *Hum. Brain Mapp.* **43**, 56–82 (2022).
13. Bulik-Sullivan, B. K. et al. LD score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* **47**, 291–295 (2015).
14. McGrath, J. J. et al. Age of onset and cumulative risk of mental disorders: a cross-national analysis of population surveys from 29 countries. *Lancet Psychiatry* **10**, 668–681 (2023).
15. Holland, D. et al. Beyond SNP heritability: polygenicity and discoverability of phenotypes estimated with a univariate Gaussian mixture model. *PLoS Genet.* **16**, e1008612 (2020).
16. Frei, O. et al. Bivariate causal mixture model quantifies polygenic overlap between complex traits beyond genetic correlation. *Nat. Commun.* **10**, 2417 (2019).
17. Merikangas, K. R. et al. Prevalence and correlates of bipolar spectrum disorder in the World Mental Health Survey Initiative. *Arch. Gen. Psychiatry* **68**, 241–251 (2011).
18. Ikeda, M. et al. A genome-wide association study identifies two novel susceptibility loci and trans population polygenicity associated with bipolar disorder. *Mol. Psychiatry* **23**, 639–647 (2018).

19. Frei, O. et al. Improved functional mapping of complex trait heritability with GSA-MiXeR implicates biologically specific gene sets. *Nat. Genet.* https://doi.org/10.1038/s41588-024-01771-1 (2024).
20. Cuéllar-Partida, G. et al. Complex-Traits Genetics Virtual Lab: a community-driven web platform for post-GWAS analyses. Preprint at *bioRxiv* https://doi.org/10.1101/518027 (2019).
21. 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).
22. 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).
23. Siletti, K. et al. Transcriptomic diversity of cell types across the adult human brain. *Science* **382**, eadd7046 (2023).
24. Finucane, H. K. et al. Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nat. Genet.* **47**, 1228–1235 (2015).
25. Yao, S. et al. Connecting genomic results for psychiatric disorders to human brain cell types and regions reveals convergence with functional connectivity. Preprint at *medRxiv* https://doi.org/10.1101/2024.01.18.24301478 (2024).
26. Weissbrod, O. et al. Functionally informed fine-mapping and polygenic localization of complex trait heritability. *Nat. Genet.* **52**, 1355–1363 (2020).
27. McLaren, W. et al. The Ensembl Variant Effect Predictor. *Genome Biol.* **17**, 122 (2016).
28. Singh, T. et al. Rare coding variants in ten genes confer substantial risk for schizophrenia. *Nature* **604**, 509–516 (2022).
29. Olfson, E. et al. Rare de novo damaging DNA variants are enriched in attention-deficit/hyperactivity disorder and implicate risk genes. *Nat. Commun.* **15**, 5870 (2024).
30. Gusev, A. et al. Integrative approaches for large-scale transcriptome-wide association studies. *Nat. Genet.* **48**, 245–252 (2016).
31. Bhattacharya, A. et al. Isoform-level transcriptome-wide association uncovers genetic risk mechanisms for neuropsychiatric disorders in the human brain. *Nat. Genet.* **55**, 2117–2128 (2023).
32. Fulco, C. P. et al. Activity-by-contact model of enhancer–promoter regulation from thousands of CRISPR perturbations. *Nat. Genet.* **51**, 1664–1669 (2019).
33. Nasser, J. et al. Genome-wide enhancer maps link risk variants to disease genes. *Nature* **593**, 238–243 (2021).
34. Koopmans, F. et al. SynGO: an evidence-based, expert-curated knowledge base for the synapse. *Neuron* **103**, 217–234.e4 (2019).
35. Kang, H. J. et al. Spatio-temporal transcriptome of the human brain. *Nature* **478**, 483–489 (2011).
36. Gaspar, H. A. & Breen, G. Drug enrichment and discovery from schizophrenia genome-wide association results: an analysis and visualisation approach. *Sci. Rep.* **7**, 12460 (2017).
37. Cannon, M. et al. DGIdb 5.0: rebuilding the drug-gene interaction database for precision medicine and drug discovery platforms. *Nucleic Acids Res.* **52**, D1227–D1235 (2024).
38. Reigstad, C. S. et al. Gut microbes promote colonic serotonin production through an effect of short-chain fatty acids on enterochromaffin cells. *FASEB J.* **29**, 1395–1403 (2015).
39. Ortega, M. A. et al. Microbiota–gut–brain axis mechanisms in the complex network of bipolar disorders: potential clinical implications and translational opportunities. *Mol. Psychiatry* **28**, 2645–2673 (2023).
40. Huang, S. et al. Lithium carbonate alleviates colon inflammation through modulating gut microbiota and T_{reg} cells in a GPR43-dependent manner. *Pharmacol. Res.* **175**, 105992 (2022).
41. Trubetskoy, V. et al. Mapping genomic loci implicates genes and synaptic biology in schizophrenia. *Nature* **604**, 502–508 (2022).
42. Kloiber, S. et al. Neurodevelopmental pathways in bipolar disorder. *Neurosci. Biobehav. Rev.* **112**, 213–226 (2020).
43. Jonsson, L., Song, J., Joas, E., Pålsson, E. & Landén, M. Polygenic scores for psychiatric disorders associate with year of first bipolar disorder diagnosis: a register-based study between 1972 and 2016. *Psychiatry Res.* **339**, 116081 (2024).
44. Zimmerman, M., Ruggero, C. J., Chelminski, I. & Young, D. Is bipolar disorder overdiagnosed? *J. Clin. Psychiatry* **69**, 935–940 (2008).
45. Zimmerman, M., Ruggero, C. J., Chelminski, I. & Young, D. Psychiatric diagnoses in patients previously overdiagnosed with bipolar disorder. *J. Clin. Psychiatry* **71**, 26–31 (2010).
46. Sollis, E. et al. The NHGRI-EBI GWAS Catalog: knowledgebase and deposition resource. *Nucleic Acids Res.* **51**, D977–D985 (2023).
47. Li, X. & Zhao, H. Automated feature extraction from population wearable device data identified novel loci associated with sleep and circadian rhythms. *PLoS Genet.* **16**, e1009089 (2020).
48. Kichaev, G. et al. Leveraging polygenic functional enrichment to improve GWAS power. *Am. J. Hum. Genet.* **104**, 65–75 (2019).
49. Thalamuthu, A. et al. Genome-wide interaction study with major depression identifies novel variants associated with cognitive function. *Mol. Psychiatry* **27**, 1111–1119 (2022).
50. Bryc, K., Durand, E. Y., Macpherson, J. M., Reich, D. & Mountain, J. L. The genetic ancestry of African Americans, Latinos, and European Americans across the United States. *Am. J. Hum. Genet.* **96**, 37–53 (2015).
51. Duncan, L. et al. Analysis of polygenic risk score usage and performance in diverse human populations. *Nat. Commun.* **10**, 3328 (2019).
52. Lewis, C. M. & Vassos, E. Polygenic risk scores: from research tools to clinical instruments. *Genome Med.* **12**, 44 (2020).
53. Yengo, L. et al. A saturated map of common genetic variants associated with human height. *Nature* **610**, 704–712 (2022).
54. Wray, N. R. et al. Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nat. Genet.* **50**, 668–681 (2018).
55. Cai, N. et al. Minimal phenotyping yields genome-wide association signals of low specificity for major depression. *Nat. Genet.* **52**, 437–447 (2020).
56. Legge, S. E. et al. Genetic and phenotypic features of schizophrenia in the UK Biobank. *JAMA Psychiatry* https://doi.org/10.1001/jamapsychiatry.2024.0200 (2024).

Kevin S. O'Connell[1,2]✉, Maria Koromina[3,4,5,276], Tracey van der Veen[6,276], Toni Boltz[7,276], Friederike S. David[8,9,276], Jessica Mei Kay Yang[10,276], Keng-Han Lin[11,276], Xin Wang[11,276], Jonathan R. I. Coleman[12,13,276], Brittany L. Mitchell[14,15,276], Caroline C. McGrouther[16,276], Aaditya V. Rangan[16,17,276], Penelope A. Lind[14,15,18,276], Elise Koch[1,2,276], Arvid Harder[19,276], Nadine Parker[1,2,276], Jaroslav Bendl[3,5,20,21,276], Kristina Adorjan[22,23,24], Esben Agerbo[25,26,27], Diego Albani[28], Silvia Alemany[29,30,31], Ney Alliey-Rodriguez[32,33], Thomas D. Als[25,34,35], Till F. M. Andlauer[36], Anastasia Antoniou[37], Helga Ask[38,39], Nicholas Bass[6], Michael Bauer[40], Eva C. Beins[8], Tim B. Bigdeli[41,42,43,44], Carsten Bøcker Pedersen[25,26,27], Marco P. Boks[45], Sigrid Børte[46,47,48], Rosa Bosch[29,49], Murielle Brum[50], Ben M. Brumpton[51], Nathalie Brunkhorst-Kanaan[50], Monika Budde[22], Jonas Bybjerg-Grauholm[25,52], William Byerley[53], Judit Cabana-Domínguez[29,30,31], Murray J. Cairns[54,55], Bernardo Carpiniello[56], Miquel Casas[49,57,58], Pablo Cervantes[59], Chris Chatzinakos[41,43], Hsi-Chung Chen[60,61], Tereza Clarence[3,5,20,21], Toni-Kim Clarke[62], Isabelle Claus[8], Brandon Coombes[63], Elizabeth C. Corfield[38,64,65], Cristiana Cruceanu[59,66], Alfredo Cuellar-Barboza[67,68], Piotr M. Czerski[69], Konstantinos Dafnas[37], Anders M. Dale[70], Nina Dalkner[71], Franziska Degenhardt[8,72], J. Raymond DePaulo[73], Srdjan Djurovic[74,75], Ole Kristian Drange[2,76], Valentina Escott-Price[10], Ayman H. Fanous[77,78,79], Frederike T. Fellendorf[71], I. Nicol Ferrier[80], Liz Forty[10], Josef Frank[81], Oleksandr Frei[1,47], Nelson B. Freimer[82,83], John F. Fullard[3,5,20,21], Julie Garnham[84], Ian R. Gizer[85], Scott D. Gordon[86], Katherine Gordon-Smith[87], Tiffany A. Greenwood[88], Jakob Grove[25,89,90,91], José Guzman-Parra[92], Tae Hyon Ha[93,94], Tim Hahn[95], Magnus Haraldsson[96,97], Martin Hautzinger[98], Alexandra Havdahl[38,39,64], Urs Heilbronner[22], Dennis Hellgren[19], Stefan Herms[8,99,100], Ian B. Hickie[101], Per Hoffmann[8,99,100], Peter A. Holmans[10], Ming-Chyi Huang[102], Masashi Ikeda[103,104], Stéphane Jamain[105], Jessica S. Johnson[3,5,106], Lina Jonsson[107], Janos L. Kalman[22,23], Yoichiro Kamatani[108,109], James L. Kennedy[110,111,112,113], Euitae Kim[93,94,114], Jaeyoung Kim[93,115], Sarah Kittel-Schneider[116,117], James A. Knowles[118], Manolis Kogevinas[119], Thorsten M. Kranz[50], Kristi Krebs[120], Steven A. Kushner[121], Catharina Lavebratt[122,123], Jacob Lawrence[124], Markus Leber[125], Heon-Jeong Lee[126], Calwing Liao[127,128], Susanne Lucae[129], Martin Lundberg[122,123], Donald J. MacIntyre[130], Wolfgang Maier[131], Adam X. Maihofer[88,132], Dolores Malaspina[3,5], Mirko Manchia[56,133], Eirini Maratou[134], Lina Martinsson[135,136], Manuel Mattheisen[25,34,35,117,137], Nathaniel W. McGregor[138], Melvin G. McInnis[139], James D. McKay[140], Helena Medeiros[141], Andreas Meyer-Lindenberg[142,143], Vincent Millischer[122,123,144,145], Derek W. Morris[146], Paraskevi Moutsatsou[134], Thomas W. Mühleisen[99,147], Claire O'Donovan[84], Catherine M. Olsen[148], Georgia Panagiotaropoulou[149], Sergi Papiol[22,23,29], Antonio F. Pardiñas[10], Hye Youn Park[93,94], Amy Perry[87], Andrea Pfennig[40], Claudia Pisanu[150], James B. Potash[73], Digby Quested[151,152], Mark H. Rapaport[153], Eline J. Regeer[154], John P. Rice[155], Margarita Rivera[156,157,158], Eva C. Schulte[8,22,131], Fanny Senner[22,23], Alexey Shadrin[1,2,159], Paul D. Shilling[88], Engilbert Sigurdsson[96,97], Lisa Sindermann[8], Lea Sirignano[81], Dan Siskind[160], Claire Slaney[84], Laura G. Sloofman[3,5], Olav B. Smeland[1,2], Daniel J. Smith[161], Janet L. Sobell[162], Maria Soler Artigas[29,30,31,163], Dan J. Stein[164], Frederike Stein[9], Mei-Hsin Su[165], Heejong Sung[166], Beata Świątkowska[167], Chikashi Terao[109], Markos Tesfaye[1,2,75], Martin Tesli[1,2,168], Thorgeir E. Thorgeirsson[169], Jackson G. Thorp[14], Claudio Toma[170,171,172], Leonardo Tondo[173], Paul A. Tooney[174], Shih-Jen Tsai[175,176], Evangelia Eirini Tsermpini[84], Marquis P. Vawter[177], Helmut Vedder[178], Annabel Vreeker[45,179,180], James T. R. Walters[10], Bendik S. Winsvold[48,181,182], Stephanie H. Witt[81], Hong-Hee Won[115,183], Robert Ye[127,128], Allan H. Young[184,185], Peter P. Zandi[73], Lea Zillich[81], 23andMe Research Team*, Estonian Biobank research team*, Genoplan Research Team*, HUNT All-In Psychiatry*, PGC-FG Single cell working group*, Genomic Psychiatry Cohort (GPC) Investigators* & Cooperative Studies Program (CSP) #572* & Million Veteran Program (MVP)*, Rolf Adolfsson[186], Martin Alda[84,187], Lars Alfredsson[188], Lena Backlund[122,123], Bernhard T. Baune[189,190,191], Frank Bellivier[192,193], Susanne Bengesser[71], Wade H. Berrettini[194], Joanna M. Biernacka[63,68], Michael Boehnke[195], Anders D. Børglum[25,89,90], Gerome Breen[12,13], Vaughan J. Carr[171], Stanley Catts[196], Sven Cichon[8,99,100,147], Aiden Corvin[197], Nicholas Craddock[10], Udo Dannlowski[95], Dimitris Dikeos[198], Bruno Etain[192,193], Panagiotis Ferentinos[12,37], Mark Frye[68], Janice M. Fullerton[170,199], Micha Gawlik[117], Elliot S. Gershon[32,200], Fernando S. Goes[73], Melissa J. Green[170,171], Maria Grigoroiu-Serbanescu[201], Joanna Hauser[202], Frans A. Henskens[203], Jens Hjerling-Leffler[204], David M. Hougaard[25,52], Kristian Hveem[51,205], Nakao Iwata[104], Ian Jones[10], Lisa A. Jones[87], René S. Kahn[3,45], John R. Kelsoe[88], Tilo Kircher[9], George Kirov[10], Po-Hsiu Kuo[60,206], Mikael Landén[19,107], Marion Leboyer[105], Qingqin S. Li[207,208], Jolanta Lissowska[209], Christine Lochner[210], Carmel Loughland[211], Jurjen J. Luykx[212,213], Nicholas G. Martin[86,214], Carol A. Mathews[215], Fermin Mayoral[92], Susan L. McElroy[216], Andrew M. McIntosh[130], Francis J. McMahon[166], Sarah E. Medland[14,214,217], Ingrid Melle[1,218], Lili Milani[120], Philip B. Mitchell[171], Gunnar Morken[219,220], Ole Mors[25,221], Preben Bo Mortensen[25,222], Bertram Müller-Myhsok[129,223,224], Richard M. Myers[225], Woojae Myung[93,94], Benjamin M. Neale[127,128,226], Caroline M. Nievergelt[88,132], Merete Nordentoft[25,227], Markus M. Nöthen[8], John I. Nurnberger[228], Michael C. O'Donovan[10], Ketil J. Oedegaard[229,230], Tomas Olsson[135,231], Michael J. Owen[10], Sara A. Paciga[232], Christos Pantelis[191,233,234], Carlos N. Pato[235], Michele T. Pato[235], George P. Patrinos[236,237,238,239], Joanna M. Pawlak[202], Josep Antoni Ramos-Quiroga[29,30,31,57], Andreas Reif[50], Eva Z. Reininghaus[71], Marta Ribasés[29,30,31,163], Marcella Rietschel[81], Stephan Ripke[127,128,149], Guy A. Rouleau[240,241], Panos Roussos[3,5,20,21,242], Takeo Saito[104], Ulrich Schall[243,244], Martin Schalling[122,123], Peter R. Schofield[170,199], Thomas G. Schulze[22,73,81,245,246], Laura J. Scott[195], Rodney J. Scott[247,248], Alessandro Serretti[249,250,251], Jordan W. Smoller[128,252,253], Alessio Squassina[150], Eli A. Stahl[3,5,226], Hreinn Stefansson[169], Kari Stefansson[169,254], Eystein Stordal[255,256], Fabian Streit[81,142,257], Patrick F. Sullivan[19,258,259], Gustavo Turecki[260], Arne E. Vaaler[261], Eduard Vieta[262], John B. Vincent[110], Irwin D. Waldman[263], Cynthia S. Weickert[170,171,264], Thomas W. Weickert[170,171,264], Thomas Werge[25,265,266,267], David C. Whiteman[148], John-Anker Zwart[47,48,181], Howard J. Edenberg[268,269], Andrew McQuillin[6,277], Andreas J. Forstner[8,147,270,277], Niamh Mullins[3,4,5,277], Arianna Di Florio[10,259,277], Roel A. Ophoff[7,82,83,277], Ole A. Andreassen[1,2,277]✉ & Bipolar Disorder Working Group of the Psychiatric Genomics Consortium*

[1]Division of Mental Health and Addiction, Oslo University Hospital, Oslo, Norway. [2]Center for Precision Psychiatry, University of Oslo, Oslo, Norway. [3]Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, NY, USA. [4]Charles Bronfman Institute for Personalized Medicine, Icahn School of Medicine at Mount Sinai, New York, NY, USA. [5]Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, NY, USA. [6]Division of Psychiatry, University College London, London, UK. [7]Department of Human Genetics, David Geffen School of Medicine, University of California Los Angeles, Los Angeles, CA, USA. [8]Institute of Human Genetics, University of Bonn, School of Medicine and University Hospital Bonn, Bonn, Germany. [9]Department of Psychiatry and Psychotherapy, University of Marburg, Marburg, Germany. [10]Centre for Neuropsychiatric Genetics and Genomics, Division of Psychological Medicine and Clinical Neurosciences, Cardiff University, Cardiff, UK. [11]23andMe Inc., Sunnyvale, CA, USA. [12]Social, Genetic and Developmental Psychiatry Centre, King's College London, London, UK. [13]NIHR Maudsley BRC, King's College London, London, UK. [14]Mental Health and Neuroscience, QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia. [15]School of Biomedical Sciences and Faculty of Medicine, The University of Queensland, Brisbane, Queensland, Australia. [16]New York University, New York, NY, USA. [17]Flatiron Institute, New York, NY, USA. [18]School of Biomedical Sciences, Queensland University of Technology, Brisbane, Queensland, Australia. [19]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden. [20]Center for Disease Neurogenomics, Icahn School of Medicine at Mount Sinai, New York, NY, USA. [21]Friedman Brain Institute, Icahn School of Medicine at Mount Sinai, New York, NY, USA. [22]Institute of Psychiatric Phenomics and Genomics (IPPG), LMU University Hospital, LMU Munich, Munich, Germany. [23]Department of Psychiatry and Psychotherapy, University Hospital, LMU Munich, Munich, Germany. [24]University Hospital of Psychiatry and Psychotherapy, University of Bern, Bern, Switzerland. [25]iPSYCH, The Lundbeck Foundation Initiative for Integrative Psychiatric Research, Aarhus, Denmark. [26]National Centre for Register-based Research, Aarhus University, Aarhus, Denmark. [27]Centre for Integrated Register-based Research, Aarhus University, Aarhus, Denmark. [28]Department of Neuroscience, Istituto Di Ricerche Farmacologiche Mario Negri IRCCS, Milano, Italy. [29]Instituto de Salud Carlos III, Biomedical Network Research Centre on Mental Health (CIBERSAM), Madrid, Spain. [30]Department of Psychiatry, Hospital Universitari Vall d'Hebron, Barcelona, Spain. [31]Psychiatric Genetics Unit, Group of Psychiatry Mental Health and Addictions, Vall d'Hebron Research Institut (VHIR), Universitat Autònoma de Barcelona, Barcelona, Spain. [32]Department of Psychiatry and Behavioral Neuroscience, University of Chicago, Chicago, IL, USA. [33]Northwestern University, Chicago, IL, USA. [34]iSEQ, Center for Integrative Sequencing, Aarhus University, Aarhus, Denmark. [35]Department of Biomedicine-Human Genetics, Aarhus University, Aarhus, Denmark. [36]Department of Neurology, Klinikum rechts der Isar, School of Medicine, Technical University of Munich, Munich, Germany. [37]National and Kapodistrian University of Athens, 2nd Department of Psychiatry, Attikon General Hospital, Athens, Greece. [38]PsychGen Centre for Genetic Epidemiology and Mental Health, Norwegian Institute of Public Health, Oslo, Norway. [39]PROMENTA Research Centre, Department of Psychology, University of Oslo, Oslo, Norway. [40]Department of Psychiatry and Psychotherapy, University Hospital Carl Gustav Carus, Technische Universität Dresden, Dresden, Germany. [41]Department of Psychiatry and Behavioral Sciences, SUNY Downstate Health Sciences University, Brooklyn, NY, USA. [42]VA NY Harbor Healthcare System, Brooklyn, NY, USA. [43]Institute for Genomics in Health, SUNY Downstate Health Sciences University, Brooklyn, NY, USA. [44]Department of Epidemiology and Biostatistics, School of Public Health, SUNY Downstate Health Sciences University, Brooklyn, NY, USA. [45]Psychiatry, Brain Center UMC Utrecht, Utrecht, The Netherlands. [46]Research and Communication Unit for Musculoskeletal Health, Division of Clinical Neuroscience, Oslo University Hospital, Ullevål, Oslo, Norway. [47]Institute of Clinical Medicine, University of Oslo, Oslo, Norway. [48]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. [49]Programa SJD MIND Escoles, Hospital Sant Joan de Déu, Institut de Recerca Sant Joan de Déu, Esplugues de Llobregat, Spain. [50]Department of Psychiatry, Psychosomatic Medicine and Psychotherapy, University Hospital Frankfurt, Frankfurt am Main, Germany. [51]K. G. Jebsen Center for Genetic Epidemiology, Department of Public Health and Nursing, Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology, Trondheim, Norway. [52]Center for Neonatal Screening, Department for Congenital Disorders, Statens Serum Institut, Copenhagen, Denmark. [53]Psychiatry, University of California San Francisco, San Francisco, CA, USA. [54]School of Biomedical Sciences and Pharmacy, The University of Newcastle, Callaghan, New South Wales, Australia. [55]Precision Medicine Research Program, Hunter Medical Research Institute, New Lambton, New South Wales, Australia. [56]Section of Psychiatry, Department of Medical Sciences and Public Health, University of Cagliari, Cagliari, Italy. [57]Department of Psychiatry and Forensic Medicine, Universitat Autònoma de Barcelona, Barcelona, Spain. [58]Fundació Privada d'Investigació Sant Pau (FISP), Barcelona, Spain. [59]Department of Psychiatry, Mood Disorders Program, McGill University Health Center, Montreal, Québec, Canada. [60]Department of Psychiatry, National Taiwan University Hospital, Taipei, Taiwan. [61]Department of Psychiatry, College of Medicine, National Taiwan University, Taipei, Taiwan. [62]Division of Psychiatry, University of Edinburgh, Edinburgh, UK. [63]Department of Quantitative Health Sciences Research, Mayo Clinic, Rochester, MN, USA. [64]Nic Waals Institute, Lovisenberg Diaconal Hospital, Oslo, Norway. [65]Department of Genetics and Bioinformatics, Norwegian Institute of Public Health, Oslo, Norway. [66]Department of Physiology and Pharmacology, Karolinska Institutet, Stockholm, Sweden. [67]Department of Psychiatry, Universidad Autonoma de Nuevo Leon, Monterrey, Mexico. [68]Department of Psychiatry and Psychology, Mayo Clinic, Rochester, MN, USA. [69]Department of Psychiatry, Laboratory of Psychiatric Genetics, Poznan University of Medical Sciences, Poznan, Poland. [70]Center for Multimodal Imaging and Genetics, Departments of Neurosciences, Radiology, and Psychiatry, University of California, San Diego, CA, USA. [71]Division of Psychiatry and Psychotherapeutic Medicine, Medical University of Graz, Graz, Austria. [72]Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, University Hospital Essen,

University of Duisburg-Essen, Duisburg, Germany. [73]Department of Psychiatry and Behavioral Sciences, Johns Hopkins University School of Medicine, Baltimore, MD, USA. [74]Department of Medical Genetics, Oslo University Hospital Ullevål, Oslo, Norway. [75]Department of Clinical Science, University of Bergen, Bergen, Norway. [76]Department of Psychiatry, Sørlandet Hospital, Kristiansand, Norway. [77]Department of Psychiatry, University of Arizona College of Medicine-Phoenix, Phoenix, AZ, USA. [78]Carl T. Hayden Veterans Affairs Medical Center, Phoenix, AZ, USA. [79]Banner-University Medical Center, Phoenix, AZ, USA. [80]Academic Psychiatry, Newcastle University, Newcastle upon Tyne, UK. [81]Department of Genetic Epidemiology in Psychiatry, Central Institute of Mental Health, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany. [82]Center for Neurobehavioral Genetics, Semel Institute for Neuroscience and Human Behavior, Los Angeles, CA, USA. [83]Department of Psychiatry and Biobehavioral Science, Semel Institute, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, CA, USA. [84]Department of Psychiatry, Dalhousie University, Halifax, Nova Scotia, Canada. [85]Department of Psychological Sciences, University of Missouri, Columbia, MO, USA. [86]Genetics and Computational Biology, QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia. [87]Psychological Medicine, University of Worcester, Worcester, UK. [88]Department of Psychiatry, University of California San Diego, La Jolla, CA, USA. [89]Department of Biomedicine and the iSEQ Center, Aarhus University, Aarhus, Denmark. [90]Center for Genomics and Personalized Medicine, CGPM, Aarhus, Denmark. [91]Bioinformatics Research Centre, Aarhus University, Aarhus, Denmark. [92]Mental Health Department, University Regional Hospital, Biomedicine Institute (IBIMA), Málaga, Spain. [93]Department of Neuropsychiatry, Seoul National University Bundang Hospital, Seongnam, Republic of Korea. [94]Department of Neuropsychiatry, Seoul National University College of Medicine, Seoul, Republic of Korea. [95]Institute for Translational Psychiatry, University of Münster, Münster, Germany. [96]Faculty of Medicine, Department of Psychiatry, School of Health Sciences, University of Iceland, Reykjavik, Iceland. [97]Landspitali University Hospital, Reykjavik, Iceland. [98]Department of Psychology, Eberhard Karls Universität Tübingen, Tubingen, Germany. [99]Department of Biomedicine, University of Basel, Basel, Switzerland. [100]Institute of Medical Genetics and Pathology, University Hospital Basel, Basel, Switzerland. [101]Brain and Mind Centre, The University of Sydney, Sydney, New South Wales, Australia. [102]Department of Psychiatry, Taipei City Psychiatric Center, Taipei City Hospital, Taipei, Taiwan. [103]Department of Psychiatry, Nagoya University Graduate School of Medicine, Nagoya, Japan. [104]Department of Psychiatry, Fujita Health University School of Medicine, Toyoake, Japan. [105]Université Paris Est Créteil, INSERM, IMRB, Translational Neuropsychiatry, Créteil, France. [106]Department of Psychiatry, UNC Chapel Hill School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA. [107]Institute of Neuroscience and Physiology, University of Gothenburg, Gothenburg, Sweden. [108]Laboratory of Complex Trait Genomics, Department of Computational Biology and Medical Sciences, Graduate School of Frontier Sciences, The University of Tokyo, Tokyo, Japan. [109]Laboratory for Statistical and Translational Genetics, RIKEN Center for Integrative Medical Sciences, Yokohama, Japan. [110]Campbell Family Mental Health Research Institute, Centre for Addiction and Mental Health, Toronto, Ontario, Canada. [111]Neurogenetics Section, Centre for Addiction and Mental Health, Toronto, Ontario, Canada. [112]Department of Psychiatry, University of Toronto, Toronto, Ontario, Canada. [113]Institute of Medical Sciences, University of Toronto, Toronto, Ontario, Canada. [114]Department of Brain and Cognitive Sciences, Seoul National University College of Natural Sciences, Seoul, Republic of Korea. [115]Samsung Advanced Institute for Health Sciences and Technology (SAIHST), Sungkyunkwan University, Samsung Medical Center, Seoul, Republic of Korea. [116]Department of Psychiatry and Neurobehavioral Science, University College Cork, Cork, Ireland. [117]Department of Psychiatry, Psychosomatics and Psychotherapy, Center of Mental Health, University Hospital Würzburg, Würzburg, Germany. [118]Human Genetics Institute of New Jersey, Rutgers University, Piscataway, NJ, USA. [119]ISGlobal, Barcelona, Spain. [120]Estonian Genome Centre, Institute of Genomics, University of Tartu, Tartu, Estonia. [121]Department of Psychiatry, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands. [122]Translational Psychiatry, Department of Molecular Medicine and Surgery, Karolinska Institutet, Stockholm, Sweden. [123]Center for Molecular Medicine, Karolinska University Hospital, Stockholm, Sweden. [124]Psychiatry, North East London NHS Foundation Trust, Ilford, UK. [125]Clinic for Psychiatry and Psychotherapy, University Hospital Cologne, Cologne, Germany. [126]Department of Psychiatry, Korea University College of Medicine, Seoul, Republic of Korea. [127]Analytic and Translational Genetics Unit, Massachusetts General Hospital, Boston, MA, USA. [128]Stanley Center for Psychiatric Research, Broad Institute, Cambridge, MA, USA. [129]Department of Translational Research in Psychiatry, Max Planck Institute of Psychiatry, Munich, Germany. [130]Division of Psychiatry, Centre for Clinical Brain Sciences, The University of Edinburgh, Edinburgh, UK. [131]Department of Psychiatry and Psychotherapy, University of Bonn, School of Medicine and University Hospital Bonn, Bonn, Germany. [132]Research/Psychiatry, Veterans Affairs San Diego Healthcare System, San Diego, CA, USA. [133]Unit of Clinical Psychiatry, University Hospital Agency of Cagliari, Cagliari, Italy. [134]National and Kapodistrian University of Athens, Medical School, Clinical Biochemistry Laboratory, Attikon General Hospital, Athens, Greece. [135]Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden. [136]Centre for Psychiatry Research, SLSO Region Stockholm, Stockholm, Sweden. [137]Department of Clinical Neuroscience, Centre for Psychiatry Research, Karolinska Institutet, Stockholm, Sweden. [138]Human and Systems Genetics Working Group, Department of Genetics, Stellenbosch University, Stellenbosch, South Africa. [139]Department of Psychiatry, University of Michigan, Ann Arbor, MI, USA. [140]Genetic Cancer Susceptibility Group, International Agency for Research on Cancer, Lyon, France. [141]Institute for Genomic Health, SUNY Downstate Medical Center College of Medicine, Brooklyn, NY, USA. [142]Department of Psychiatry and Psychotherapy, Central Institute of Mental Health, Medical Faculty Mannheim, University of Heidelberg, Mannheim, Germany. [143]German Centre for Mental Health (DZPG), partner site Mannheim-Heidelberg-Ulm, Mannheim, Germany. [144]Department of Psychiatry and Psychotherapy, Clinical Division of General Psychiatry, Medical University of Vienna, Vienna, Austria. [145]Comprehensive Center for Clinical Neurosciences and Mental Health, Medical University of Vienna, Vienna, Austria. [146]Centre for Neuroimaging and Cognitive Genomics (NICOG), School of Biological and Chemical Sciences, University of Galway, Galway, Ireland. [147]Institute of Neuroscience and Medicine (INM-1), Research Centre Jülich, Jülich, Germany. [148]Population Health, QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia. [149]Department of Psychiatry and Psychotherapy, Charité-Universitätsmedizin, Berlin, Germany. [150]Department of Biomedical Sciences, University of Cagliari, Cagliari, Italy. [151]Oxford Health NHS Foundation Trust, Warneford Hospital, Oxford, UK. [152]Department of Psychiatry, University of Oxford, Warneford Hospital, Oxford, UK. [153]Department of Psychiatry and Behavioral Sciences, Emory University School of Medicine, Atlanta, GA, USA. [154]Outpatient Clinic for Bipolar Disorder, Altrecht, Utrecht, The Netherlands. [155]Department of Psychiatry, Washington University in Saint Louis, Saint Louis, MO, USA. [156]Department of Biochemistry and Molecular Biology II, Faculty of Pharmacy, University of Granada, Granada, Spain. [157]Institute of Neurosciences 'Federico Olóriz', Biomedical Research Center (CIBM), University of Granada, Granada, Spain. [158]Instituto de Investigación Biosanitaria ibs.GRANADA, Granada, Spain. [159]KG Jebsen Centre for Neurodevelopmental disorders, University of Oslo, Oslo, Norway. [160]Faculty of Medicine, University of Queensland, Brisbane, Queensland, Australia. [161]Division of Psychiatry, Centre for Clinical Brain Sciences, University of Edinburgh, Edinburgh, UK. [162]Psychiatry and the Behavioral Sciences, University of Southern California, Los Angeles, CA, USA. [163]Department of Genetics, Microbiology, and Statistics, Faculty of Biology, Universitat de Barcelona, Barcelona, Spain. [164]SAMRC Unit on Risk and Resilience in Mental Disorders, Dept of Psychiatry and Neuroscience Institute, University of Cape Town, Cape Town, South Africa. [165]Virginia Institute for Psychiatric and Behavioral Genetics, Virginia Commonwealth University, Richmond, VA, USA. [166]Human Genetics Branch, Intramural Research Program, National Institute of Mental Health, Bethesda, MD, USA. [167]Department of Environmental Epidemiology, Nofer Institute of Occupational Medicine, Lodz, Poland. [168]Department of Mental Disorders, Norwegian Institute of Public Health, Oslo, Norway. [169]deCODE Genetics/Amgen, Reykjavik, Iceland. [170]Neuroscience Research Australia, Sydney, New South Wales, Australia. [171]Discipline of Psychiatry and Mental Health, School of Clinical Medicine, Faculty of Medicine and Health, University of New South Wales, Sydney, New South Wales, Australia. [172]Centro de Biología Molecular Severo Ochoa, Universidad Autónoma de Madrid and CSIC, Madrid, Spain. [173]Department of Psychiatry, Harvard Medical School, Boston, MA, USA. [174]School of Biomedical Science and Pharmacy, University of Newcastle, Newcastle, New South Wales, Australia. [175]Department of Psychiatry, Taipei Veterans General Hospital, Taipei, Taiwan. [176]Division of Psychiatry, National Yang Ming Chiao Tung University, Taipei, Taiwan. [177]Department of Psychiatry and Human Behavior, School of Medicine, University of California, Irvine, CA, USA. [178]Psychiatry, Psychiatrisches Zentrum Nordbaden, Wiesloch, Germany. [179]Department of Child and Adolescent Psychiatry/Psychology, Erasmus MC Sophia Children Hospital, Erasmus University, Rotterdam, The Netherlands. [180]Department of Psychology Education and Child Studies, Erasmus School of Social and Behavioral Sciences, Erasmus University Rotterdam, Rotterdam, The Netherlands. [181]Department of Research, Innovation and Education, Division of Clinical Neuroscience, Oslo University Hospital, Oslo, Norway. [182]Department of Neurology, Oslo University Hospital, Oslo, Norway. [183]Samsung Genome Institute, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Republic of Korea. [184]Department of Psychological Medicine, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK. [185]South London and Maudsley NHS Foundation Trust, Bethlem Royal Hospital, Kent, UK. [186]Department of Clinical Sciences, Psychiatry, Umeå University Medical Faculty, Umeå, Sweden. [187]National Institute of Mental Health, Klecany, Czech Republic. [188]Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden. [189]Department of Psychiatry, University of Münster, Münster, Germany. [190]Department of Psychiatry, Melbourne Medical School, The University of Melbourne, Melbourne, Victoria, Australia. [191]The Florey Institute of Neuroscience and Mental Health, The University of Melbourne, Parkville, Victoria, Australia. [192]Université Paris Cité, INSERM, Optimisation Thérapeutique en Neuropsychopharmacologie, UMRS-1144, Paris, France. [193]APHP Nord, DMU Neurosciences, GHU Saint Louis-Lariboisière-Fernand Widal, Département de Psychiatrie et de Médecine Addictologique, Paris, France. [194]Psychiatry, University of Pennsylvania, Philadelphia, PA, USA. [195]Center for Statistical Genetics and Department of Biostatistics, University of Michigan, Ann Arbor, MI, USA. [196]University of Queensland, Brisbane, Queensland, Australia. [197]Neuropsychiatric Genetics Research Group, Department of Psychiatry and Trinity Translational Medicine Institute, Trinity College Dublin, Dublin, Ireland. [198]National and Kapodistrian University of Athens, 1st Department of Psychiatry, Eginition Hospital, Athens, Greece. [199]School of Biomedical Sciences, Faculty of Medicine and Health, University of New South Wales, Sydney, New South Wales, Australia. [200]Department of Human Genetics, University of Chicago, Chicago, IL, USA. [201]Biometric Psychiatric Genetics Research Unit, Alexandru Obregia Clinical Psychiatric Hospital, Bucharest, Romania. [202]Department of Psychiatric Genetics, Poznan University of Medical Sciences, Poznan, Poland. [203]School of Medicine and Public Health, University of Newcastle, Newcastle, New South Wales, Australia. [204]Department of Medical Biochemistry and Biophysics, Karolinska Institutet, Stockholm, Sweden. [205]HUNT Research Center, Department of Public Health and Nursing, Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology, Trondheim, Norway. [206]Department of Public Health and Institute of Epidemiology and Preventive Medicine, College of Public Health, National Taiwan University, Taipei, Taiwan. [207]Neuroscience Therapeutic Area, Janssen Research and Development, Titusville, NJ, USA. [208]JRD Data Science, Janssen Research and Development,

Titusville, NJ, USA. [209]Cancer Epidemiology and Prevention, M. Sklodowska-Curie National Research Institute of Oncology, Warsaw, Poland. [210]SA MRC Unit on Risk and Resilience in Mental Disorders, Department of Psychiatry, Stellenbosch University, Stellenbosch, South Africa. [211]University of Newcastle, Newcastle, New South Wales, Australia. [212]Department of Psychiatry, Amsterdam University Medical Center, Amsterdam, The Netherlands. [213]Department of Psychiatry and Neuropsychology, School for Mental Health and Neuroscience, Maastricht University Medical Center, Maastricht, The Netherlands. [214]School of Psychology, The University of Queensland, Brisbane, Queensland, Australia. [215]Department of Psychiatry, University of Florida, Gainesville, FL, USA. [216]Research Institute, Lindner Center of HOPE, Mason, OH, USA. [217]School of Psychology and Counselling, Queensland University of Technology, Brisbane, Queensland, Australia. [218]Division of Mental Health and Addiction, University of Oslo, Institute of Clinical Medicine, Oslo, Norway. [219]Department of Mental Health, Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology (NTNU), Trondheim, Norway. [220]Psychiatry, St Olavs University Hospital, Trondheim, Norway. [221]Psychosis Research Unit, Aarhus University Hospital-Psychiatry, Risskov, Denmark. [222]NCRR and CIRRAU, Aarhus BSS, Aarhus University, Aarhus, Denmark. [223]Munich Cluster for Systems Neurology (SyNergy), Munich, Germany. [224]University of Liverpool, Liverpool, UK. [225]HudsonAlpha Institute for Biotechnology, Huntsville, AL, USA. [226]Medical and Population Genetics, Broad Institute, Cambridge, MA, USA. [227]Mental Health Services in the Capital Region of Denmark, Mental Health Center Copenhagen, University of Copenhagen, Copenhagen, Denmark. [228]Psychiatry, Indiana University School of Medicine, Indianapolis, IN, USA. [229]Division of Psychiatry, Haukeland Universitetssjukehus, Bergen, Norway. [230]Faculty of Medicine and Dentistry, University of Bergen, Bergen, Norway. [231]Center for Molecular Medicine, Stockholm, Sweden. [232]Human Genetics and Computational Biomedicine, Pfizer Global Research and Development, Groton, CT, USA. [233]Melbourne Neuropsychiatry Centre, Department of Psychiatry, The University of Melbourne, Melbourne, Victoria, Australia. [234]Monash Institute of Pharmaceutical Sciences (MIPS), Monash University, Parkville, Victoria, Australia. [235]Rutgers Health, Rutgers University, Piscataway, NJ, USA. [236]University of Patras, School of Health Sciences, Department of Pharmacy, Laboratory of Pharmacogenomics and Individualized Therapy, Patras, Greece. [237]Department of Genetics and Genomics, College of Medicine and Health Sciences, United Arab Emirates University, Al-Ain, United Arab Emirates. [238]Zayed Center for Health Sciences, United Arab Emirates University, Al-Ain, United Arab Emirates. [239]Department of Pathology, Faculty of Medicine and Health Sciences, Clinical Bioinformatics Unit, Erasmus University Medical Center Rotterdam, Rotterdam, The Netherlands. [240]Department of Neurology and Neurosurgery, Faculty of Medicine, McGill University, Montreal, Québec, Canada. [241]Montreal Neurological Institute and Hospital, McGill University, Montréal, Québec, Canada. [242]Center for Precision Medicine and Translational Therapeutics, James J. Peters VA Medical Center, Bronx, NY, USA. [243]Centre for Brain and Mental Health Research, The University of Newcastle, Newcastle, New South Wales, Australia. [244]Hunter Medical Research Institute, New Lambton Heights, New South Wales, Australia. [245]Department of Psychiatry and Psychotherapy, University Medical Center Göttingen, Göttingen, Germany. [246]Department of Psychiatry and Behavioral Sciences, SUNY Upstate Medical University, Syracuse, NY, USA. [247]The School of Biomedical Sciences and Pharmacy, Faculty of Medicine, Health and Wellbeing, University of Newcastle, Newcastle, New South Wales, Australia. [248]Cancer Detection and Therapies Program, Hunter Medical Research Institute, University of Newcastle, Newcastle, New South Wales, Australia. [249]Department of Medicine and Surgery, Kore University of Enna, Enna, Italy. [250]Department of Biomedical and Neuromotor Sciences, University of Bologna, Bologna, Italy. [251]Oasi Research Institute-IRCCS, Troina, Italy. [252]Department of Psychiatry, Massachusetts General Hospital, Boston, MA, USA. [253]Psychiatric and Neurodevelopmental Genetics Unit (PNGU), Massachusetts General Hospital, Boston, MA, USA. [254]Faculty of Medicine, University of Iceland, Reykjavik, Iceland. [255]Department of Psychiatry, Hospital Namsos, Namsos, Norway. [256]Department of Neuroscience, Norges Teknisk Naturvitenskapelige Universitet Fakultet for naturvitenskap og teknologi, Trondheim, Norway. [257]Hector Institute for Artificial Intelligence in Psychiatry, Central Institute of Mental Health, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany. [258]Department of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA. [259]Department of Psychiatry, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA. [260]Department of Psychiatry, McGill University, Montreal, Québec, Canada. [261]Department of Psychiatry, Sankt Olavs Hospital Universitetssykehuset i Trondheim, Trondheim, Norway. [262]Clinical Institute of Neuroscience, Hospital Clinic, University of Barcelona, IDIBAPS, CIBERSAM, Barcelona, Spain. [263]Department of Psychology, Emory University, Atlanta, GA, USA. [264]Department of Neuroscience, SUNY Upstate Medical University, Syracuse, NY, USA. [265]Institute of Biological Psychiatry, Mental Health Services, Copenhagen University Hospital, Copenhagen, Denmark. [266]Department of Clinical Medicine, University of Copenhagen, Copenhagen, Denmark. [267]Center for GeoGenetics, GLOBE Institute, University of Copenhagen, Copenhagen, Denmark. [268]Biochemistry and Molecular Biology, Indiana University School of Medicine, Indianapolis, IN, USA. [269]Department of Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, IN, USA. [270]Centre for Human Genetics, University of Marburg, Marburg, Germany. [276]These authors contributed equally: Maria Koromina, Tracey van der Veen, Toni Boltz, Friederike S. David, Jessica Mei Kay Yang, Keng-Han Lin, Xin Wang, Jonathan R. I. Coleman, Brittany L. Mitchell, Caroline C. McGrouther, Aaditya V. Rangan, Penelope A. Lind, Elise Koch, Arvid Harder, Nadine Parker, Jaroslav Bendl. [277]These authors jointly supervised this work: Andrew McQuillin, Andreas J. Forstner, Niamh Mullins, Arianna Di Florio, Roel A. Ophoff, Ole A. Andreassen. *Lists of authors and their affiliations appears online. **Lists of members and their affiliations appear in the Supplementary Information.

23andMe Research Team
Keng-Han Lin[11,276] & Xin Wang[11,276]

Estonian Biobank research team
Kristi Krebs[120] & Lili Milani[120]

Genoplan Research Team
Byung-Chul Lee[271,272], Ji-Woong Kim[271,272], Young Kee Lee[271,272], Joon Ho Kang[271,272], Myeong Jae Cheon[271,272] & Dong Jun Kim[271,272]

HUNT All-In Psychiatry
Bendik S. Winsvold[48,181,182], Eystein Stordal[255,256], Gunnar Morken[219,220], John-Anker Zwart[47,48,181], Ole Kristian Drange[2,76] & Sigrid Børte[46,47,48]

PGC-FG Single cell working group
Arvid Harder[19,276], Howard J. Edenberg[268,269], Nadine Parker[1,2,276], Eva C. Schulte[8,22,131] & Jens Hjerling-Leffler[204]

Genomic Psychiatry Cohort (GPC) Investigators
Michele T. Pato[235], Carlos N. Pato[235], Tim B. Bigdeli[41,42,43,44], Ayman H. Fanous[77,78,79], James A. Knowles[118], Mark H. Rapaport[153], Janet L. Sobell[162], Helena Medeiros[141], William Byerley[53], James L. Kennedy[110,111,112,113], Jordan W. Smoller[128,252,253], Patrick F. Sullivan[19,258,259], Tiffany A. Greenwood[88], Marquis P. Vawter[177] & Chris Chatzinakos[41,43]

Cooperative Studies Program (CSP) #572
Mihaela Aslan[273], Philip D. Harvey[274], Grant D. Huang[275], Tim B. Bigdeli[41,42,43,44], Ayman H. Fanous[77,78,79] & John R. Kelsoe[88]

Million Veteran Program (MVP)
Grant D. Huang[275]

[271]Genoplan RnD Division, Genoplan Korea, Seoul, Republic of Korea. [272]Genoplan RnD Division, Genoplan Japan, Fukuoka, Japan. [273]Clinical Epidemiology Research Center (CERC), VA Connecticut Healthcare System, West Haven, CT, USA. [274]Bruce W. Carter Miami Veterans Affairs (VA) Medical Center, Miami, FL, USA. [275]Office of Research and Development, Veterans Health Administration, Washington, DC, USA.

Bipolar Disorder Working Group of the Psychiatric Genomics Consortium
Kevin S. O'Connell[1,2], Maria Koromina[3,4,5,276], Tracey van der Veen[6], Toni Boltz[7,276], Friederike S. David[8,9,276], Jessica Mei Kay Yang[10,276], Keng-Han Lin[11,276], Xin Wang[11,276], Jonathan R. I. Coleman[12,13,276], Brittany L. Mitchell[14,15,276], Caroline C. McGrouther[16,276], Aaditya V. Rangan[16,17,276], Penelope A. Lind[14,15,18,276], Elise Koch[1,2,276], Arvid Harder[19,276], Nadine Parker[1,2,276], Jaroslav Bendl[3,5,20,21,276], Kristina Adorjan[22,23,24], Esben Agerbo[25,26,27], Diego Albani[28], Silvia Alemany[29,30,31], Ney Alliey-Rodriguez[32,33], Thomas D. Als[25,34,35], Till F. M. Andlauer[36], Anastasia Antoniou[37], Helga Ask[38,39], Nicholas Bass[6], Michael Bauer[40], Eva C. Beins[8], Tim B. Bigdeli[41,42,43,44], Carsten Bøcker Pedersen[25,26,27], Marco P. Boks[45], Sigrid Børte[46,47,48], Rosa Bosch[29,49], Murielle Brum[50], Ben M. Brumpton[51], Nathalie Brunkhorst-Kanaan[50], Monika Budde[22], Jonas Bybjerg-Grauholm[25,52], William Byerley[53], Judit Cabana-Domínguez[29,30,31], Murray J. Cairns[54,55], Bernardo Carpiniello[56], Miquel Casas[49,57,58], Pablo Cervantes[59], Chris Chatzinakos[41,43], Hsi-Chung Chen[60,61], Tereza Clarence[3,5,20,21], Toni-Kim Clarke[62], Isabelle Claus[8], Brandon Coombes[63], Elizabeth C. Corfield[38,64,65], Cristiana Cruceanu[59,66], Alfredo Cuellar-Barboza[67,68], Piotr M. Czerski[69], Konstantinos Dafnas[37], Anders M. Dale[70], Nina Dalkner[71], Franziska Degenhardt[8,72], J. Raymond DePaulo[73], Srdjan Djurovic[74,75], Ole Kristian Drange[2,76], Valentina Escott-Price[10], Ayman H. Fanous[77,78,79], Frederike T. Fellendorf[71], I. Nicol Ferrier[80], Liz Forty[10], Josef Frank[81], Oleksandr Frei[1,47], Nelson B. Freimer[82,83], John F. Fullard[3,5,20,21], Julie Garnham[84], Ian R. Gizer[85], Scott D. Gordon[86], Katherine Gordon-Smith[87], Tiffany A. Greenwood[88], Jakob Grove[25,89,90,91], José Guzman-Parra[92], Tae Hyon Ha[93,94], Tim Hahn[95], Magnus Haraldsson[96,97], Martin Hautzinger[98], Alexandra Havdahl[38,39,64], Urs Heilbronner[22], Dennis Hellgren[19], Stefan Herms[8,99,100], Ian B. Hickie[101], Per Hoffmann[8,99,100], Peter A. Holmans[10], Ming-Chyi Huang[102], Masashi Ikeda[103,104], Stéphane Jamain[105], Jessica S. Johnson[3,5,106], Lina Jonsson[107], Janos L. Kalman[22,23], Yoichiro Kamatani[108,109], James L. Kennedy[110,111,112,113], Euitae Kim[93,94,114], Jaeyoung Kim[93,115], Sarah Kittel-Schneider[116,117], James A. Knowles[118], Manolis Kogevinas[119], Thorsten M. Kranz[50], Kristi Krebs[120], Steven A. Kushner[121], Catharina Lavebratt[122,123], Jacob Lawrence[124], Markus Leber[125], Heon-Jeong Lee[126], Calwing Liao[127,128], Susanne Lucae[129], Martin Lundberg[122,123], Donald J. MacIntyre[130], Wolfgang Maier[131], Adam X. Maihofer[88,132], Dolores Malaspina[3,5], Mirko Manchia[56,133], Eirini Maratou[134], Lina Martinsson[135,136], Manuel Mattheisen[25,34,35,117,137], Nathaniel W. McGregor[138], Melvin G. McInnis[139], James D. McKay[140], Helena Medeiros[141], Andreas Meyer-Lindenberg[142,143], Vincent Millischer[122,123,144,145], Derek W. Morris[146], Paraskevi Moutsatsou[134], Thomas W. Mühleisen[99,147], Claire O'Donovan[84], Catherine M. Olsen[148], Georgia Panagiotaropoulou[149], Sergi Papiol[22,23,29], Antonio F. Pardiñas[10], Hye Youn Park[93,94], Amy Perry[87], Andrea Pfennig[40], Claudia Pisanu[150], James B. Potash[73], Digby Quested[151,152], Mark H. Rapaport[153], Eline J. Regeer[154], John P. Rice[155], Margarita Rivera[156,157,158], Eva C. Schulte[8,22,131], Fanny Senner[22,23], Alexey Shadrin[1,2,159], Paul D. Shilling[88], Engilbert Sigurdsson[96,97], Lisa Sindermann[8], Lea Sirignano[81], Dan Siskind[160], Claire Slaney[84], Laura G. Sloofman[3,5], Olav B. Smeland[1,2], Daniel J. Smith[161], Janet L. Sobell[162], Maria Soler Artigas[29,30,31,163], Dan J. Stein[164], Frederike Stein[9], Mei-Hsin Su[165], Heejong Sung[166], Beata Świątkowska[167], Chikashi Terao[109], Markos Tesfaye[1,2,75], Martin Tesli[1,2,168], Thorgeir E. Thorgeirsson[169], Jackson G. Thorp[14], Claudio Toma[170,171,172], Leonardo Tondo[173], Paul A. Tooney[174], Shih-Jen Tsai[175,176],

Evangelia Eirini Tsermpini[84], Marquis P. Vawter[177], Helmut Vedder[178], Annabel Vreeker[45,179,180], James T. R. Walters[10], Bendik S. Winsvold[48,181,182], Stephanie H. Witt[81], Hong-Hee Won[115,183], Robert Ye[127,128], Allan H. Young[184,185], Peter P. Zandi[73], Lea Zillich[81], Rolf Adolfsson[186], Martin Alda[84,187], Lars Alfredsson[188], Lena Backlund[122,123], Bernhard T. Baune[189,190,191], Frank Bellivier[192,193], Susanne Bengesser[71], Wade H. Berrettini[194], Joanna M. Biernacka[63,68], Michael Boehnke[195], Anders D. Børglum[25,89,90], Gerome Breen[12,13], Vaughan J. Carr[171], Stanley Catts[196], Sven Cichon[8,99,100,147], Aiden Corvin[197], Nicholas Craddock[10], Udo Dannlowski[95], Dimitris Dikeos[198], Bruno Etain[192,193], Panagiotis Ferentinos[12,37], Mark Frye[68], Janice M. Fullerton[170,199], Micha Gawlik[117], Elliot S. Gershon[32,200], Fernando S. Goes[73], Melissa J. Green[170,171], Maria Grigoroiu-Serbanescu[201], Joanna Hauser[202], Frans A. Henskens[203], Jens Hjerling-Leffler[204], David M. Hougaard[25,52], Kristian Hveem[51,205], Nakao Iwata[104], Ian Jones[10], Lisa A. Jones[87], René S. Kahn[3,45], John R. Kelsoe[88], Tilo Kircher[9], George Kirov[10], Po-Hsiu Kuo[60,206], Mikael Landén[19,107], Marion Leboyer[105], Qingqin S. Li[207,208], Jolanta Lissowska[209], Christine Lochner[210], Carmel Loughland[211], Jurjen J. Luykx[212,213], Nicholas G. Martin[86,214], Carol A. Mathews[215], Fermin Mayoral[92], Susan L. McElroy[216], Andrew M. McIntosh[130], Francis J. McMahon[166], Sarah E. Medland[14,214,217], Ingrid Melle[1,218], Lili Milani[120], Philip B. Mitchell[171], Gunnar Morken[219,220], Ole Mors[25,221], Preben Bo Mortensen[25,222], Bertram Müller-Myhsok[129,223,224], Richard M. Myers[225], Woojae Myung[93,94], Benjamin M. Neale[127,128,226], Caroline M. Nievergelt[88,132], Merete Nordentoft[25,227], Markus M. Nöthen[8], John I. Nurnberger[228], Michael C. O'Donovan[10], Ketil J. Oedegaard[229,230], Tomas Olsson[135,231], Michael J. Owen[10], Sara A. Paciga[232], Christos Pantelis[191,233,234], Carlos N. Pato[235], Michele T. Pato[235], George P. Patrinos[236,237,238,239], Joanna M. Pawlak[202], Josep Antoni Ramos-Quiroga[29,30,31,57], Andreas Reif[50], Eva Z. Reininghaus[71], Marta Ribasés[29,30,31,163], Marcella Rietschel[81], Stephan Ripke[127,128,149], Guy A. Rouleau[240,241], Panos Roussos[3,5,20,21,242], Takeo Saito[104], Ulrich Schall[243,244], Martin Schalling[122,123], Peter R. Schofield[170,199], Thomas G. Schulze[22,73,81,245,246], Laura J. Scott[195], Rodney J. Scott[247,248], Alessandro Serretti[249,250,251], Jordan W. Smoller[128,252,253], Alessio Squassina[150], Eli A. Stahl[3,5,226], Hreinn Stefansson[169], Kari Stefansson[169,254], Eystein Stordal[255,256], Fabian Streit[81,142,257], Patrick F. Sullivan[19,258,259], Gustavo Turecki[260], Arne E. Vaaler[261], Eduard Vieta[262], John B. Vincent[110], Irwin D. Waldman[263], Cynthia S. Weickert[170,171,264], Thomas W. Weickert[170,171,264], Thomas Werge[25,265,266,267], David C. Whiteman[148], John-Anker Zwart[47,48,181], Howard J. Edenberg[268,269], Andrew McQuillin[6,277], Andreas J. Forstner[8,147,270,277], Niamh Mullins[3,4,5,277], Arianna Di Florio[10,259,277], Roel A. Ophoff[7,82,83,277] & Ole A. Andreassen[1,2,277]

Methods

Sample description

Details of each of the cohorts, including sample size, ancestry, inclusion and exclusion criteria for cases and controls as well as citations, are provided in Supplementary Table 1 and the Supplementary Note. We included three types of samples: (1) samples in which participants were assessed using semi-structured or structured interviews (clinical), (2) samples in which participants were assessed using medical records, registries and questionnaire data (community), and (3) samples in which participants self-report a diagnosis of BD (self-reported). The clinical samples included 55 cohorts, 46 of which were included in previous PGC-BD GWAS publications[3,10,11]. The community samples included 20 cohorts, 11 of which were included in the previous PGC-BD GWAS[3]. Finally, we included four self-reported cohorts from 23andMe, in which individuals were classified as cases if they self-reported having received a clinical diagnosis or treatment for BD in responses to web-based surveys ("Have you ever been diagnosed with, or treated for, bipolar disorder?").

Individual-level genotype and phenotype data were shared with the PGC for 53 'internal' cohorts, whereas the remaining 26 'external' cohorts contributed summary statistic data.

The final multi-ancestry meta-analysis included up to 158,036 cases and 2,796,499 controls. The total n_{eff}, equivalent to an equal number of cases and controls in each cohort ($4 \times n_{cases} \times n_{controls}/(n_{cases} + n_{controls})$, was 535,720 with 82.3% of participants (proportion of n_{eff}) of EUR ancestry, 4.4% of AFR ancestry, 4.2% of EAS ancestry and 9.1% of LAT ancestry.

The majority of new cohorts included in this study were external community cohorts in which subtype definitions were more difficult to determine, and as such, the total number of BDI and BDII subtype cases does not differ remarkably from the previous PGC-BD GWAS[3] (Supplementary Table 1). Thus, the previous BDI (25,060 cases and 449,978 controls) and BDII (6,781 cases and 364,075 controls) GWAS summary statistic data were used for BDI and BDII analyses in this study.

Genotyping and imputation

Technical quality control was performed separately on each cohort for which individual-level data were provided separately according to standards developed by the PGC[57], including SNP missingness < 0.05 (before sample removal), subject missingness < 0.02, autosomal heterozygosity deviation ($F_{het} < 0.2$), SNP missingness < 0.02 (after sample removal), difference in SNP missingness between cases and controls < 0.02, SNP Hardy–Weinberg equilibrium ($P > 1 \times 10^{-10}$ in BD cases and $P > 1 \times 10^{-6}$ in controls), and mismatches between pedigree and genetically determined sex based on the F statistic of X chromosome homozygosity (female $F < 0.2$ and male $F > 0.8$). In addition, relatedness was calculated across cohorts using identity by descent, and one of each pair of related individuals (pi_hat > 0.2) was excluded, prioritizing exclusion of individuals related to the most others, controls over cases, and individuals from larger cohorts. Principal components were generated using genotyped SNPs in each cohort separately using EIGENSTRAT (v6.1.4; https://www.hsph.harvard.edu/alkes-price/software/)[58]. Genotype imputation was performed using the prephasing/imputation stepwise approach implemented in Eagle (v2.3.5; https://alkesgroup.broadinstitute.org/Eagle/)[59] and Minimac3 (https://genome.sph.umich.edu/wiki/Minimac3)[60] to the Haplotype Reference Consortium (HRC) reference panel (v1.0)[61]. Data on the X chromosome were also available for all 53 internal cohorts, and these were imputed to the HRC reference panel in males and females separately. The remaining 22 external cohorts were processed by the contributing collaborative teams using comparable procedures. Identical individuals between PGC-processed cohorts and external cohorts with suspected sample overlap were detected using genotype-based checksums (https://personal.broadinstitute.org/sripke/share_links/zpXkV8INxUg9bayDpLToG4g58TMtjN_PGC_SCZ_w3.0718d.76) and removed from the PGC cohorts.

Genome-wide association study

For internal cohorts, GWAS were conducted within each cohort using an additive logistic regression model in PLINK (v1.90; https://www.cog-genomics.org/plink2/)[62], covarying for the first five principal components and any others as required, as previously described[3]. Analyses of the X chromosome were performed in males and females separately, with males scored 0 or 2 and females scored 0, 1 or 2. X chromosome analyses were performed only in individuals of EUR ancestry for which individual-level data were available. For external cohorts, GWAS were conducted by the collaborating research teams using comparable procedures. To control test statistic inflation at SNPs with low minor allele frequency (MAF) in small cohorts, SNPs were retained only if cohort MAF was more than 1% and minor allele count was more than 10 in either cases or controls (whichever had smaller n).

Initially, meta-analysis of GWAS summary statistics was conducted using inverse-variance-weighted fixed-effect models in METAL (v2011-03-25; https://genome.sph.umich.edu/wiki/METAL_Documentation)[63] across cohorts within ancestral groups. A genome-wide significant locus was defined as the region around a SNP with $P < 5.0 \times 10^{-8}$ with linkage disequilibrium $R^2 > 0.1$, within a 3,000-kb window, based on the linkage disequilibrium structure of the ancestry-matched HRC reference panel (v1.0)[61], except LAT (EUR panel used). Multi-ancestry meta-analysis was similarly performed by combining cohorts with diverse ancestry using inverse-variance-weighted fixed-effect models in METAL[63]. Given that more than 80% of the included participants were of EUR ancestry, the linkage disequilibrium structure of the EUR HRC reference panel was used to define genome-wide significant loci.

For all meta-analyses, SNPs present in less than 75% of total effective sample size (n_{eff}) were removed from the meta-analysis results. In addition, we used the DENTIST tool (https://github.com/Yves-CHEN/DENTIST) for summary data-based analyses, which leverages linkage disequilibrium from a reference sample (ancestry-matched HRC reference panel (v1.0)[61], except LAT and multi-ancestry for which the EUR panel was used) to detect and filter out problematic variants by testing the difference between the observed z-score of a variant and a predicted z-score from the neighbouring variants[64].

To identify independent association signals ($P < 5 \times 10^{-8}$), the GCTA forward selection and backward elimination process (command 'cojo-slct') was applied using the summary statistics from the EAS, EUR and multi-ancestry meta-analysis (both including and excluding the self-report data), with the EAS and EUR HRC reference panels, respectively[65,66].

The genetic correlation between meta-analyses based on all new cohorts (118,284 cases and 2,448,096 controls) and EUR cohorts from our previous PGC-BD GWAS[3] was $r_g = 0.64$ (s.e. = 0.02), and $r_g = 0.91$ (s.e. = 0.04) when excluding self-reported cohorts. Concordance of the direction of associations in the present GWAS with associations in the previously published BD data were evaluated as previously described[67].

Heritability and genetic correlation

LDSC (https://github.com/bulik/ldsc)[13] was used to estimate the SNP-h^2 of BD from EUR GWAS summary statistics, including all cohorts as well as subgroups by ascertainment and BD subtype. Popcorn was used to estimate SNP-h^2 of BD from non-EUR GWAS summary statistics[68]. SNP-h^2 was converted to the liability scale using a lifetime BD prevalence of 2%. LDSC bivariate genetic correlations were also estimated between EUR BD GWASs (with and without self-report data) and 11 other psychiatric disorders as well, as 1,390 human diseases and traits via the Complex Traits Genetics Virtual Lab (https://vl.genoma.io) web platform[20]. Adjusting for the number of traits tested, the Bonferroni-corrected $P < 3.569 \times 10^{-5}$. Cross-ancestry bivariate genetic correlations were estimated using Popcorn (https://github.com/brielin/Popcorn)[68].

Differences in r_g between phenotype pairs were tested as a deviation from 0 using the block jackknife approach implemented in LDSC[69].

The results of the clinical and community cohort meta-analyses were conditioned on genetic risks for BDI and BDII, to account for differences in proportion of the BD subtypes within these cohorts. Conditioning was conducted using multitrait-based conditional and joint analysis using GWAS summary data (mtCOJO; https://yanglab.westlake.edu.cn/software/gcta/#mtCOJO)[70], implemented in GCTA[65]. mtCOJO is robust to sample overlap between the GWAS of the exposure and outcome. The conditioned summary statistics were evaluated for genetic correlation with self-reported BD using LDSC.

MiXeR

We applied causal mixture models (MiXeR; https://github.com/precimed/mixer)[15,16,71] to investigate the genetic architecture of BD, specifically the overlap between clinical, community and self-report samples, as well as BD subtypes. We first computed univariate analyses to estimate the polygenicity, discoverability and heritability of each trait. These were followed by bivariate analyses to compute the number of shared trait-influencing variants between pairs of traits, and finally trivariate analyses to compute the proportion of shared variants between all three traits analysed. We also determined the correlation of effect sizes of SNPs within the bivariate shared components. For trivariate MiXeR analyses, model optimization procedures were repeated 20 times (20 runs) to obtain the means and standard errors of model parameters. Estimated parameters from the 'run' with the smallest deviation from the median overlap pattern were then selected and reported.

Polygenic association with BD

We used PRS-CS-auto[21] to compute PRSs in target cohorts, using a discovery GWAS in which the target cohort was left out. Given that the majority of the individuals included in the meta-analysis were of EUR descent, we used the EUR linkage disequilibrium reference panel based on UK BioBank data as provided by PRS-CS developers (https://github.com/getian107/PRScs). Raw scores were standardised to *z*-scores, and covariates including sex, the first five principal components and any others as required (as above for each cohort GWAS) were included in the logistic regression model, via the glm() function in R[72], with family=binomial and link=logit. The variance explained by PRS (R^2) was first converted to Nagelkerke's pseudo-R^2 via the fmsb package in R (https://cran.r-project.org/web/packages/fmsb/index.html), and then converted to the liability scale to account for the proportion of cases in each cohort and the population prevalence of BD[73]. We have provided R^2 values for BD assuming a population prevalence of 2%, based on a recent multinational survey[14]. The weighted average R^2 values were then calculated using the n_{eff} for each cohort. PRS-specific medians and their confidence intervals were computed using non-parametric bootstrap replicates (10,000 resamples with replacement). The odds ratios for BD for individuals in the top quintile of PRS compared with those in the middle quintile were calculated for all cohorts. Similarly, the area under the curve (AUC) statistic was calculated via the pROC package in R (https://cran.r-project.org/web/packages/pROC/index.html), for which we performed a training and testing procedure by taking 80% of the individuals in a given cohort on which to train the model, and tested the predictability in the remaining 20% of individuals. Ten random samplings of training and testing sets were performed in all cohorts, and the median AUC after all permutations is provided Supplementary Tables 16–22. The median confidence intervals for the AUC were similarly averaged across the ten random permutations. These AUC statistics were calculated based on the logistic regression model that includes the standardized PRS as a predictor and principal component covariates. To assess the gain in AUC due to the PRS itself, we subtracted the median AUC of the model containing only the covariates from the full model, reported in Supplementary Tables 16–22 as AUC.

Gene and gene set association analysis

Gene-level, gene set and tissue set associations were performed using a SNP-wise mean model (±10-kb window) implemented in MAGMA (https://ctg.cncr.nl/software/magma)[22]. Bonferroni correction was used to control for multiple testing. In addition, we performed gene set analysis with GSA-MiXeR (https://github.com/precimed/gsa-mixer)[19], which quantifies partitioned heritability attributed to $n = 10{,}475$ gene sets from the Gene Ontology[74] and SynGO[34] databases, alongside their fold enrichment with respect to a baseline model. The GSA-MiXeR full model incorporates 18,201 protein-coding genes, using a joint model to estimate heritability attributed to each gene based on GWAS summary statistics and HRC[59] reference panel to account for linkage disequilibrium between variants. The baseline model in GSA-MiXeR accounted for a set of 75 functional annotations[75], as well as accounting for MAF-dependent and linkage disequilibrium-dependent genetic architecture. The heritability model in GSA-MiXeR was estimated using Adam (method for stochastic gradient-based optimization of the likelihood function)[76]. Standard errors of fitted parameters were estimated from the observed Fisher's information matrix (the negative Hessian matrix of the log-likelihood function).

Identified credible genes were further assessed for enrichment in synaptic processes using the SynGO tool (v1.2; https://www.syngoportal.org/) with default settings[34].

Cell-type-specific enrichment analyses

Single-cell enrichment analyses of brain cell types were performed according to Mullins et al.[3]. In brief, from five publicly available single-cell RNA sequencing datasets derived from human[77,78] and mouse[79–81] brain tissues, 10% of genes with the highest gene expression specificity per cell type were extracted. After MAGMA[22] gene analysis of the multi-ancestry GWAS summary statistics including self-reported data using an annotation window of 35 kb upstream and 10 kb downstream of the gene boundaries and the 1,000 Genomes phase 3 EUR reference panel, MAGMA gene set analyses were conducted for all cell types in each dataset, respectively. Within each dataset, false discovery rate (FDR)-adjusted $P < 0.05$ was considered statistically significant.

In addition, we performed an exploratory single-cell enrichment analysis in 914 cell types across 29 non-brain mouse tissues as implemented in FUMA[82]. Cell types with FDR-adjusted $P < 0.05$ were considered statistically significant. Moreover, to determine that identified enrichment was not due to overlapping genes with neuronal cell types, we performed cross-dataset conditional analyses of significantly enriched cell types with mouse brain tissue.

Single-nucleus RNA sequencing enrichment

We used the Human Brain Atlas single-nucleus RNA sequencing dataset[23] consisting of 3.369 million nuclei sequenced using single-nucleus RNA sequencing. The nuclei were from adult post-mortem donors, and the dissections focused on 106 anatomical locations within 10 brain regions. Following quality control, the nuclear gene expression patterns allowed the identification of a hierarchy of cell types that were organized into 31 superclusters and 461 clusters. In the current paper, we used the same naming system for the cell types and the brain regions as in Siletti et al.[23]. We estimated SNP-h^2 enrichment for the top decile of expression proportion genes (approximately 1,300 genes) in each of the 31 superclusters and 461 clusters, respectively, using stratified LDSC[24], as previously described[25]. We used FDR correction (FDR < 0.05) to account for multiple comparisons.

Fine-mapping

We performed functional fine-mapping of genome-wide significant loci via Polyfun-SuSiE[26], using functional annotations of the baseline-LF2.2 UKB model and linkage disequilibrium estimates from the HRC EUR ($n = 21{,}265$) reference panel. The maximum number of causal variants

per fine-mapped region was adjusted accordingly based on the results from the conditional analysis. We excluded loci that fell within the MHC locus (chromosome 6: 28000000–34000000, build GRCh37) due to the known complexity of the linkage disequilibrium architecture in that region. Genome-wide significant locus ranges with a linkage disequilibrium $R^2 > 0.1$ were used as fine-mapping ranges. Putatively causal SNPs (PIP > 0.50 and part of 95% credible set) were mapped to genes by performing variant annotation with VEP (GRCh37) Ensembl release 75 (https://www.ensembl.org/info/docs/tools/vep/index.html)[27].

Convergence of common and rare variation

Data from the BipEx consortium[4] (13,933 cases of BD and 14,422 controls) were used to assess the convergence of common and rare variant signals, using a similar approach as previously used for schizophrenia[41]. This dataset includes approximately 8,200 individuals with BDI and 3,400 individuals with BDII, whereas the remainder of the sample lack BD subtype information. Ultra-rare variants (5 or less minor allele count) for damaging missense (missense badness, PolyPhen-2 and regional constraint score of more than 3) and protein-truncating variants (including transcript ablation, splice acceptor variants, splice donor variants, stop gained and frameshift variants) were considered. An enrichment of rare variants in genes prioritized through fine-mapping in cases relative to controls were assessed using a Fisher's exact test. Given the genetic overlap between BD and schizophrenia, we repeated the analysis in data from the Schizophrenia Exome Meta-analysis (SCHEMA) cohort (24,248 schizophrenia cases and 97,322 controls)[28]. Using the same approach as taken in the SCHEMA[28] and BipEx[4] papers, background genes included all genes surveyed in each sequencing study, respectively. As a sensitivity analysis, we further evaluated the enrichment of synonymous variants in the credible genes in cases of BD of the BipEx cohort and found no enrichment (OR = 0.96, 95% CI = 0.935–0.985).

QTL integrative analysis

We conducted different QTL integration analyses to elucidate molecular mechanisms by which variants associated with BD might be linked to the phenotype. Summary data-based Mendelian randomization (SMR; v1.3; https://yanglab.westlake.edu.cn/software/smr/)[83] with subsequent heterogeneity in dependent instruments (HEIDI)[70] tests were performed for eQTLs, sQTLs and mQTLs. Data on eQTLs and sQTLs were obtained from the BrainMeta study (v2; $n = 2{,}865$)[84], whereas data on mQTLs were obtained from the Brain-mMeta study (v1; $n = 1{,}160$)[85]. Putatively causal SNPs identified from fine-mapping, as outlined above, were used as the QTL instruments for the SMR analyses. Using the BD GWAS and QTL summary statistics, each putative causal SNP was analysed as the target SNP for probes within a 2-Mb window on either side using the --extract-target-snp-probe option in SMR. The EUR HRC linkage disequilibrium reference panel was used for the analyses of the multi-ancestry meta-analysis. A Bonferroni correction was applied for 2,021 tests, that is, SNP-QTL probe combinations, in the eQTL analysis ($P_{SMR} < 2.47 \times 10^{-5}$), 6,755 tests in the sQTL analysis ($P_{SMR} < 7.40 \times 10^{-6}$) and 2,222 tests in the mQTL analysis ($P_{SMR} < 2.25 \times 10^{-5}$). The significance threshold for the HEIDI test was $P_{HEIDI} \geq 0.01$. Additional eQTL integration analyses were conducted using TWAS (http://gusevlab.org/projects/fusion/), FOCUS and isoTWAS (https://github.com/bhattacharya-a-bt/isotwas). Details related to these analyses are provided in the Supplementary Note.

Enhancer–promoter gene interactions

To investigate enhancer–promoter interactions influenced by BD GWAS variants, we utilized cell-type-specific enhancer–promoter maps from a multi-omics dataset, which included joint single-nucleus ATAC–single nucleus RNA sequencing and cell-specific Hi-C data from developing brains. We used the activity-by-contact (ABC) model[32,33] for this analysis. Following the authors' guidelines, we excluded enhancer–promoter interactions that (1) had an ABC score below 0.015, (2) involved ubiquitously expressed genes or genes on the Y chromosome, or (3) included genes not expressed in major brain cell types. Focusing on the BD GWAS, we selected only those enhancer–promoter links that overlapped genome-wide significant SNPs (with peaks extended by 100 bp on both sides to increase overlap) or their linkage disequilibrium buddies ($R^2 \geq 0.8$). This selection process yielded 11,023 enhancer–promoter links. We then overlapped these putative disease-relevant variants with enhancer–promoter links to prioritize causal genes. To avoid multiple associations for a single variant, we applied the ABC-Max approach[33], retaining only the enhancer–promoter links with the highest ABC score for each peak.

Credible gene identification

We have provided a set of credible genes by integrating information from various gene-mapping strategies, using a similar approach previously described[86] (Extended Data Fig. 7 and Supplementary Table 31). First, genes identified through fine-mapping, and QTL (eQTL, mQTL and sQTL) analyses using SMR and proximity (nearest gene within 10 kb) to fine-mapped putatively causal SNPs were included. The identified set of 116 genes were then further assessed based on gene-level associations (MAGMA)[22], additional integrative eQTL analyses[30,31] and enhancer–promoter gene interactions[32,33]. The criteria for filtering genes from the different eQTL methods were: (1) SMR adjusted $P < 0.05$ and HEIDI test $P > 0.01$, (2) TWAS adjusted $P < 0.05$ and colocalization probability (COLOC.PP4) > 0.7, (3) FOCUS posterior inclusion probability > 0.7 and within a credible set, and (4) isoTWAS permutation $P < 0.05$, isoTWAS poster inclusion probability > 0.7 and within a credible set (Extended Data Fig. 7). Genes annotated by at least one of these eQTL approaches were confirmed as having eQTL evidence (Supplementary Table 31). Thus, seven approaches were considered by which loci could be mapped to genes, including eQTL evidence (eQTL or TWAS or FOCUS or isoTWAS), mQTL, sQTL, VEP, proximity, MAGMA and enhancer–promoter interactions.

Temporal clustering of credible genes

Lifespan gene expression from the Human Brain Transcriptome project (www.hbatlas.org)[35] was used to cluster the list of credible genes based on their temporal variation. The gene expression and associated metadata were acquired from the Gene Expression Omnibus (GEO accession GSE25219). The data consist of 57 donors 5.7 weeks post-conception to 82 years of age with samples extracted across regions of the brain. Before filtering gene expression for the list of credible genes, gene symbols of both credible genes and the gene expression dataset were harmonized using the 'limma' package in R (https://www.bioconductor.org/packages/release/bioc/html/limma.html), which updates any synonymous gene symbols to the latest Entrez symbol. Gene expression was available for 34 of the 36 credible genes. Within a given brain region, the expression of each gene was then mean centred and scaled. Outliers in gene expression more than 4 standard deviations from the mean were removed. To generate a single gene expression profile for each gene across the lifespan, at a given age, the mean gene expression for a given gene was taken across brain regions, and in some cases across donors. This resulted in a matrix in which each gene had a single expression value for each age across the lifespan. This gene expression-by-age matrix was then used to cluster the credible genes by the lifespan expression profiles using the R package 'TMixClust' (https://www.bioconductor.org/packages/release/bioc/html/TMixClust.html). This method used mixed-effect models with non-parametric smoothing splines to capture and cluster non-linear variation in temporal gene expression. We tested $K = 2$ to $K = 10$ clusters performing 50 clustering runs to analyse stability. The clustering solution with the highest likelihood (that is, the global optimum using an expectation maximization technique) is selected as the most stable solution across the 50 runs for each of the trials testing 2–10 clusters. We compared the average silhouette width across the $K = 2$ to $K = 10$ clusters and selected that

with the maximum value as the optimal number of clusters. The highest average silhouette width was 0.24 for two clusters, whereas the lowest was 0.17 for four clusters. Overall, evidence was suggestive for a two-cluster solution for the temporal expression of credible genes.

Drug enrichment analyses

Gene set analyses were performed restricted to genes targeted by drugs, assessing individual drugs and grouping drugs with similar actions as previously described[3,36]. Gene-level and gene set analyses of the multi-ancestry GWAS summary statistics including self-reported data were performed in MAGMA (v1.10)[22], as outlined above for cell-type-specific enrichment.

Gene sets were defined comprising the targets of each drug in the Drug Gene Interaction database DGIdb (v5.0.6)[37]; the Psychoactive Drug Screening Database Ki DB[87]; ChEMBL (v27)[88]; the Target Central Resource Database (v6.7.0)[89]; and DSigDB (v1.0)[90]; all downloaded in October 2020. Multiple testing was controlled using a Bonferroni-corrected significance threshold of $P < 5.41 \times 10^{-5}$ (924 drug sets with at least 10 valid drug gene sets) for drug set analysis and $P < 5.49 \times 10^{-4}$ (91 drug classes) for drug-class analysis, respectively.

We also assessed whether any of the 36 credible genes were classified as druggable in the OpenTargets platform (https://genetics.opentargets.org/).

In addition, gene set analyses were also performed to test the enrichment of drug–gene interactions on only credible genes as described above. Moreover, we investigated whether any lithium target genes, as well as their interaction partners, were among the 36 credible genes using the latest version of the human protein interactome[91]. We calculated the S_{AB} between credible genes and lithium target genes, in which a significant overlapping network neighbourhood would be indicative of functional similarity[92].

Reporting summary

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

Data availability

Genome-wide association summary statistics for these analyses are available at https://www.med.unc.edu/pgc/download-results/. The full GWAS summary statistics for the 23andMe datasets will be made available through 23andMe to qualified researchers under an agreement with 23andMe that protects the privacy of the 23andMe participants. Please visit https://research.23andme.com/collaborate/#dataset-access for more information and to apply to access the data. After applying with 23andMe, the full summary statistics including all analysed SNPs and samples in the GWAS meta-analyses will be accessible to the approved researchers. Genotype data are available for a subset of cohorts, including dbGAP accession numbers and/or restrictions, as described in the 'Cohort descriptions' section of the supplementary materials.

Code availability

No custom code was developed for this study. All software and tools used for the analyses presented are publicly available and referenced within the respective sections in the Methods of the article.

57. Lam, M. et al. RICOPILI: rapid imputation for consortias pipeline. *Bioinformatics* **36**, 930–933 (2020).
58. Price, A. L. et al. Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.* https://doi.org/10.1038/ng1847 (2006).
59. Loh, P.-R. et al. Reference-based phasing using the Haplotype Reference Consortium panel. *Nat. Genet.* **48**, 1443–1448 (2016).
60. Das, S. et al. Next-generation genotype imputation service and methods. *Nat. Genet.* **48**, 1284–1287 (2016).
61. McCarthy, S. et al. A reference panel of 64,976 haplotypes for genotype imputation. *Nat. Genet.* **48**, 1279–1283 (2016).
62. Chang, C. C. et al. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**, 7 (2015).
63. Willer, C. J., Li, Y. & Abecasis, G. R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190–2191 (2010).
64. Chen, W. et al. Improved analyses of GWAS summary statistics by reducing data heterogeneity and errors. *Nat. Commun.* **12**, 7117 (2021).
65. Yang, J., Lee, S. H., Goddard, M. E. & Visscher, P. M. GCTA: a tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* **88**, 76–82 (2011).
66. Yang, J. et al. Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nat. Genet.* **44**, 369–375 (2012).
67. Demontis, D. et al. Genome-wide analyses of ADHD identify 27 risk loci, refine the genetic architecture and implicate several cognitive domains. *Nat. Genet.* **55**, 198–208 (2023).
68. Brown, B. C., Asian Genetic Epidemiology Network Type 2 Diabetes Consortium, Ye, C. J., Price, A. L. & Zaitlen, N. Transethnic genetic-correlation estimates from summary statistics. *Am. J. Hum. Genet.* **99**, 76–88 (2016).
69. Hübel, C. et al. Genomics of body fat percentage may contribute to sex bias in anorexia nervosa. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* **180**, 428–438 (2019).
70. Zhu, Z. et al. Causal associations between risk factors and common diseases inferred from GWAS summary data. *Nat. Commun.* **9**, 224 (2018).
71. Shadrin, A. A. et al. Dissecting the genetic overlap between three complex phenotypes with trivariate MiXeR. Preprint at *medRxiv* https://doi.org/10.1101/2024.02.23.24303236 (2024).
72. R Core Team. R: A Language and Environment for Statistical Computing. *R Foundation* http://www.R-project.org/ (2013).
73. 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).
74. Ashburner, M. et al. Gene Ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* **25**, 25–29 (2000).
75. Gazal, S. et al. Linkage disequilibrium-dependent architecture of human complex traits shows action of negative selection. *Nat. Genet.* **49**, 1421–1427 (2017).
76. Kingma, D. P. & Ba, J. Adam: a method for stochastic optimization. Preprint at https://arxiv.org/abs/1412.6980 (2014).
77. Habib, N. et al. Massively parallel single-nucleus RNA-seq with DroNc-seq. *Nat. Methods* **14**, 955–958 (2017).
78. Lake, B. B. et al. Integrative single-cell analysis of transcriptional and epigenetic states in the human adult brain. *Nat. Biotechnol.* **36**, 70–80 (2017).
79. Skene, N. G. et al. Genetic identification of brain cell types underlying schizophrenia. *Nat. Genet.* **50**, 825–833 (2018).
80. Zeisel, A. et al. Molecular architecture of the mouse nervous system. *Cell* **174**, 999–1014.e22 (2018).
81. Saunders, A. et al. Molecular diversity and specializations among the cells of the adult mouse brain. *Cell* **174**, 1015–1030 (2018).
82. Watanabe, K., Umićević Mirkov, M., de Leeuw, C. A., van den Heuvel, M. P. & Posthuma, D. Genetic mapping of cell type specificity for complex traits. *Nat. Commun.* **10**, 3222 (2019).
83. Zhu, Z. et al. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nat. Genet.* **48**, 481–487 (2016).
84. Qi, T. et al. Genetic control of RNA splicing and its distinct role in complex trait variation. *Nat. Genet.* **54**, 1355–1363 (2022).
85. Qi, T. et al. Identifying gene targets for brain-related traits using transcriptomic and methylomic data from blood. *Nat. Commun.* **9**, 2282 (2018).
86. Shrine, N. et al. Multi-ancestry genome-wide association analyses improve resolution of genes and pathways influencing lung function and chronic obstructive pulmonary disease risk. *Nat. Genet.* **55**, 410–422 (2023).
87. Roth, B. L., Lopez, E., Patel, S. & Kroeze, W. K. The multiplicity of serotonin receptors: uselessly diverse molecules or an embarrassment of riches? *Neuroscientist* **6**, 252–262 (2000).
88. Mendez, D. et al. ChEMBL: towards direct deposition of bioassay data. *Nucleic Acids Res.* **47**, D930–D940 (2019).
89. Sheils, T. K. et al. TCRD and Pharos 2021: mining the human proteome for disease biology. *Nucleic Acids Res.* **49**, D1334–D1346 (2021).
90. Yoo, M. et al. DSigDB: drug signatures database for gene set analysis. *Bioinformatics* **31**, 3069–3071 (2015).
91. Morselli Gysi, D. & Barabási, A.-L. Noncoding RNAs improve the predictive power of network medicine. *Proc. Natl Acad. Sci. USA* **120**, e2301342120 (2023).
92. Menche, J. et al. Disease networks. Uncovering disease–disease relationships through the incomplete interactome. *Science* **347**, 1257601 (2015).

Acknowledgements We thank the participants who donated their time, life experiences and DNA to this research, the clinical and scientific teams that worked with them, and the investigators who comprise the PGC. The PGC has received major funding from the US National Institute of Mental Health (PGC4: R01 MH124839, PGC3: U01 MH109528, PGC2: U01 MH094421 and PGC1: U01 MH085520). Statistical analyses were carried out on the NL Genetic Cluster Computer (http://www.geneticcluster.org) hosted by SURFsara. The content is solely the responsibility of the authors and does not necessarily represent the official views of the US National Institutes of Health. Individual and cohort-specific funding acknowledgements are detailed in the Supplementary Information.

Author contributions The management group for this paper was led by O.A.A. The management group comprised a subset of authors responsible for the study design, conduct, primary and final interpretation, and included A.M., A.J.F., N.M., A.D.F., R.A.O., H.J.E., K.S.O. and O.A.A; this group was also responsible for primary drafting and editing of the manuscript. The analytical team, led by K.S.O., was responsible for the main analyses presented in the paper, and included M. Koromina, T.v.d.V., T.B., F.S.D., J.M.K.Y., K.-H.L., X.W., J.R.I.C., B.L.M., C.C.M, A.V.R., P.A.L., E. Koch, A. Harder, N.P., J.B. and K.S.O. Imputation, quality control and GWAS were conducted by K.S.O., M. Koromina, B.L.M., K.-H.L., X.W. and J.M.K.Y. Heritability and genetic

correlation analyses were performed by K.S.O. MiXeR was done by K.S.O. and A. Shadrin. Polygenic association was conducted by T.v.d.V., T.B., B.L.M. and P.A.L. Gene and gene set analyses were done by K.S.O., C.C.M. and A.V.R. Cell-type-specific analyses were performed by F.S.D. Single-nucleus RNA sequencing enrichment was done by A. Harder and J.H.-L. Fine-mapping was conducted by M. Koromina. Rare variant analyses were performed by C. Liao. QTL integrative analyses were done by M. Koromina, T.B. and F.S.D. Enhancer–promoter interactions were analysed by J.B. Credible gene prioritization was performed by K.S.O. and M. Koromina. Temporal clustering was done by N.P. Drug enrichment was conducted by J.R.I.C. and E. Koch. Clinical assessments were performed by A.A., A.C., A.C.-B., A.D.B., A.D.F., A.E.V., A.H.Y., A. Havdahl, A.M., A.M.M., A. Perry, A. Pfennig, A.R., A. Serretti, A.V., B. Carpiniello, B.E., B.S., B.T.B., C.A.M., C. Lavebratt, C. Loughland, C.N.P., C.O., C.S., D.D., D.H., D.J.M., D. J. Smith, D.M., D.Q., E.C.S., E.E.T., E.J.R., E. Kim, E. Sigurdsson, E.S.G., E. Stordal, E.V., E.Z.R., F. Senner, F.S.G., F. Stein, F. Streit, F.T.F., G.B., G.K., G.M., H.A., H.-J.L., H.M., H.V., H.Y.P., I.D.W., I.J., I.M., I.R.G., J.A.R.-Q., J.B.P., J.B.V., J.G.-P., J. Garnham, J. Grove, J.H., J.I.N., J. L. Kalman, J. L. Kennedy, J.L.S., J. Lawrence, J. Lissowska, J.M.P., J.P.R., J.R.D., J.W.S., K.A., K.D., K.G.-S., K.J.O., L.A.J., L.B., L.F., L. Martinsson, L. Sirignano, L.T., L.Z., M.A., M. Bauer, M. Brum, M. Budde, M.C., M.C.O., M.F., M.G., M.G.-S., M.G.M., M.H.R., M. Haraldsson, M. Hautzinger, M.I., M.J.G., M.J.O., M. Kogevinas, M. Landén, M. Lundberg, M. Manchia, M. Mattheisen, M.P.B., M.P.V., M. Rietschel, M. Tesfaye, M.T.P., M. Tesli, N.A.-R., N.B., N.B.-K., N.C., N.D., N.G.M., N.I., O.A.A., O.B.S., O.K.D., O.M., P. B. Mitchell, P. B. Mortensen, P.C., P.F., P.M.C., R.A., R.B., R.S.K., S.A.K., S. Bengeser, S.K.-S., S.L., S.L.M., S.P., T.G.S., T.H., T.H.H., T.K., T.M.K., T.O., T.S., T.W., T.W.W., U.D., U.H., V.M., W.B., W. Maier and W. Myung. Data processing and analyses were performed by A.C., A.D.B., A.F.P., A. Harder, A.J.F., A.M.D., A. Shadrin, A.V.R., A.X.M., B. Coombes, B.L.M., B.M.-M., B.M.B., B.S.W., C.B.P., C. Cruceanu, C.C.M., C. Chatzinakos, C. Liao, C.M.N., C.S.W., C. Terao, C. Toma, D.A., D.M.H., D.W.M., E.A., E.A.S., E.C.B., E.C.C., E. Koch, E.M., E.S.G., F.D., F.J.M., F.S.D., G.A.R., G.B., G.P., G.T., H.-C.C., H. Stefansson, H. Sung, H.-H.W., I.C., J.B., J.C.-D., J.D.M., J.F., J.F.F., J.G.T., J. Grove, J.H.-L., J.K., J.M.B., J.M.F., J.M.K.Y., J.R.I.C., J.S.J., J.T.R.W., K.K., K.-H.L., K.S.O., L.G.S., L.J., L. Milani, L. Sindermann, M.-C.H., M.I., M.J.C., M. Koromina, M. Leber, M.M.N., M. Mattheisen, M. Ribasés, M. Rivera, M.S.A., M.S., M. Tesfaye, N.B.F., N.I., N.M., N.P., N.W.M., O.B.S., O.F., O.K.D., P.A.H., P.A.L., P.A.T., P.D.S., P.F.S., P.H., P.-H.K., P.M., P.P.Z., P.R., Q.S.L., R.J.S., R.M.M., R.Y., S.A., S. Børte, S. Cichon, S.D., S.D.G., S.E.M., S.H., S.H.W., S.J., S.R., S.-J.T., T.A.G., T.B., T.B.B., T.C., T.D.A., T.E.T., T.F.M.A., T.O., T.S., T.v.d.V., T.W., T.W.M., V.E.-P., W. Myung, X.W. and Y.K. Funding was obtained by A.C., A.D.B., A.H.Y., A.M.M., B.E., B.M.N., B.T.B., C.N.P., C. Pantelis, C.S.W., C. Terao, D. J. Stein, D.M., D.S., E.S.G., F.B., F.J.M., G.A.R., G.B., G.P.P., G.T., H.J.E., I.B.H., I.J., I.M., I.N.F., J.A.K., J.B.P., J.B.V., J.I.N., J.M.B., J.M.F., J.R.D., J.W.S., K.H., L.A., L.A.J., L.B., M.A., M. Boehnke, M.C.O., M.F., M.G.-S., M.H.R., M.I., M.J.G., M.J.O., M. Leboyer, M. Landén, M.M.N., M.N., M. Rietschel, M.S., M.T.P., N.C., N.G.M., N.I., O.A.A., O.M., P.A.T., P. B. Mitchell, P. B. Mortensen, P.P.Z., P.R.S., R.A.O., R.J.S., R.M.M., S.E.M., S.J., S.L., T.B.B., T.G.S., T.O., T.S., T.W., T.W.W., W.H.B. and Y.K. Recruitment and genotyping were performed by A.C., A.D.B., A.D.F., A.E.V., A.H.F., A.J.F., A.M.M., A.M., A.R., A. Serretti, A. Squassina, B. Carpiniello, B.-C.L., B.E., B.M.-M., B.M.N., B.T.B., C.A.M., C.B.P., C. Lochner, C.M.N., C.M.O., C.N.P., C. Pantelis, C. Pisanu, C.S.W., D.C.W., D.D., D.J.K., E.A., E.S.G., E. Stordal, E.V., E.Z.R., F.A.H., F.B., F.J.M., F.M., F.S.G., F. Stein, G.A.R., G.B., G.D.H., G.M., G.P.P., G.T., H.J.E., H. Stefansson, H.-H.W., I.B.H., I.D.W., I.J., J.A.R.-Q., J.B.V., J.H., J.H.K., J.H.-L., J.I.N., J.J.L., J. Lissowska, J.M.B., J.M.F., J.M.P., J.R.D., J.R.K., J.-W.K., J.W.S., J.-A.Z., K.H., K.J.O., K.S., L.A., L.A.J., L.J.S., L. Milani, L.T., M.A., M. Aslan, M.C.O., M.F., M.G., M.G.-S., M.I., M.J.C., M.J.G., M.J.O., M. Leboyer, M. Landén, M.M.N., M. Manchia, M.N., M. Ribasés, M. Rietschel, M.S., M.T.P., N.C., N.G.M., N.I., O.A.A., O.M., P.A.L., P. B. Mitchell, P. B. Mortensen, P.D.H., P.F., P.-H.K., P.R., P.R.S., Q.S.L., R.A., R.A.O., R.S.K., S.A.P., S. Bengesser, S. Cichon, S. Catts, S.E.M., S.L.M., S.R., T.G.S., T.H., T.K., T.S., T.W., T.W.W., U.D., U.S., V.J.C., W.H.B., W. Myung and Y.K.L. Numerous authors beyond the initial writing group contributed to data interpretation and provided edits, comments and suggestions to the paper. All authors reviewed the manuscript critically for important intellectual content and approved the final version of the manuscript for publication. The Chair of the PGC is P.F.S. The Bipolar Disorder Working Group of the PGC is led by O.A.A.

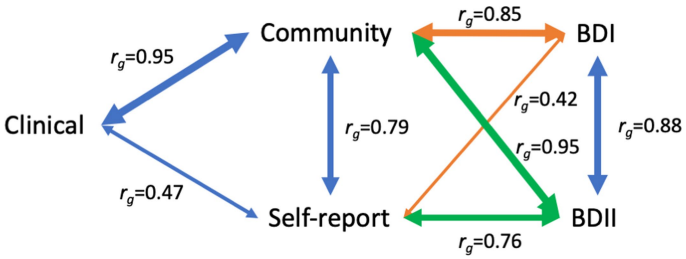
Competing interests T.E.T., H. Stefansson and K.S. are employed by deCODE Genetics/Amgen. E.A.S. is an employee of Regeneron Genetics Center and owns stocks of Regeneron Pharmaceutical. K.-H.L. and X.W. are employed by 23andMe. Multiple additional authors work for pharmaceutical or biotechnology companies in a manner directly analogous to academic co-authors and collaborators. A.H.Y. has given paid lectures and served on advisory boards relating to drugs used in affective and related disorders for several companies (AstraZeneca, Eli Lilly, Lundbeck, Sunovion, Servier, Livanova, Janssen, Allergan, Bionomics and Sumitomo Dainippon Pharma), was Lead Investigator for the Embolden study (AstraZeneca), BCI Neuroplasticity study and Aripiprazole Mania study, and is an investigator for Janssen, Lundbeck, Livanova and Compass. J.I.N. is an investigator for Janssen. P.F.S. is on the advisory committee and a shareholder of Neumora Therapeutics. G.B. reports consultancy and speaker fees from Eli Lilly and Illumina, and grant funding from Eli Lilly. M. Landén has received speaker fees from Lundbeck. O.A.A. has served as a speaker for Janssen, Lundbeck and Sunovion, and as a consultant for Cortechs.ai. A.M.D. is a founder of and holds equity interest in CorTechs Labs and serves on its scientific advisory board; is a member of the scientific advisory board of Human Longevity and the Mohn Medical Imaging and Visualization Center; and has received research funding from General Electric Healthcare. E.V. has received grants and served as a consultant, advisor or CME speaker for the following entities: AB-Biotics, Abbott, Allergan, Angelini, AstraZeneca, Bristol Myers Squibb, Dainippon Sumitomo Pharma, Farmindustria, Ferrer, Forest Research Institute, Gedeon Richter, GlaxoSmithKline, Janssen, Lundbeck, Otsuka, Pfizer, Roche, SAGE, Sanofi-Aventis, Servier, Shire, Sunovion, Takeda, the Brain and Behaviour Foundation, the Catalan Government (AGAUR and PERIS), the Spanish Ministry of Science, Innovation, and Universities (AES and CIBERSAM), the Seventh European Framework Programme and Horizon 2020 and the Stanley Medical Research Institute. S.K.-S. received author's, speaker's and consultant honoraria from Janssen, Medice Arzneimittel Pütter GmbH and Takeda outside of the current work. A. Serretti is or has been a consultant and/or speaker for: Abbott, AbbVie, Angelini, AstraZeneca, Clinical Data, Boheringer, Bristol Myers Squibb, Eli Lilly, GlaxoSmithKline, Innovapharma, Italfarmaco, Janssen, Lundbeck, Naurex, Pfizer, Polifarma, Sanofi and Servier. J.R.D. has served as an unpaid consultant to Myriad-Neuroscience (formerly Assurex Health) in 2017 and 2019, and owns stock in CVS Health. B.M.N. is a member of the scientific advisory board at Deep Genomics, and consultant for Camp4 Therapeutics, Takeda Pharmaceutical and Biogen. B.-C.L., J.-W.K., Y.K.L., J.H.K., M. J. Cheon and D.J.K. are employed by Genoplan. I.B.H. is the Co-Director of Health and Policy at the Brain and Mind Centre (BMC) University of Sydney. The BMC operates an early-intervention youth services at Camperdown under contract to Headspace. I.B.H. is also the Chief Scientific Advisor to, and a 3.2% equity shareholder in, InnoWell. InnoWell was formed by the University of Sydney (45% equity) and PwC (Australia; 45% equity) to deliver the $30 M (AUD) Australian Government-funded Project Synergy (2017–2020; a 3-year program for the transformation of mental health services) and to lead transformation of mental health services internationally through the use of innovative technologies. M.J.O. and M.C.O. have received funding from Takeda Pharmaceuticals and Akrivia Health outside the scope of the current work. P. B. Mitchell. has received remuneration from Janssen (Australia) and Sanofi (Hangzhou) for lectures or advisory board membership. J.A.R.-Q. was on the speakers' bureau and/or acted as consultant for Biogen, Idorsia, Casen-Recordati, Janssen-Cilag, Novartis, Takeda, Bial, Sincrolab, Neuraxpharm, Novartis, BMS, Medice, Rubió, Uriach, Technofarma and Raffo in the past 3 years; has also received travel awards (airplane tickets and hotel) for taking part in psychiatric meetings from Idorsia, Janssen-Cilag, Rubió, Takeda, Bial and Medice; and the Department of Psychiatry chaired by J.A.R.-Q. received unrestricted educational and research support from the following companies in the past 3 years: Exeltis, Idorsia, Janssen-Cilag, Neuraxpharm, Oryzon, Roche, Probitas and Rubió. All other authors declare no competing interests.

Additional information

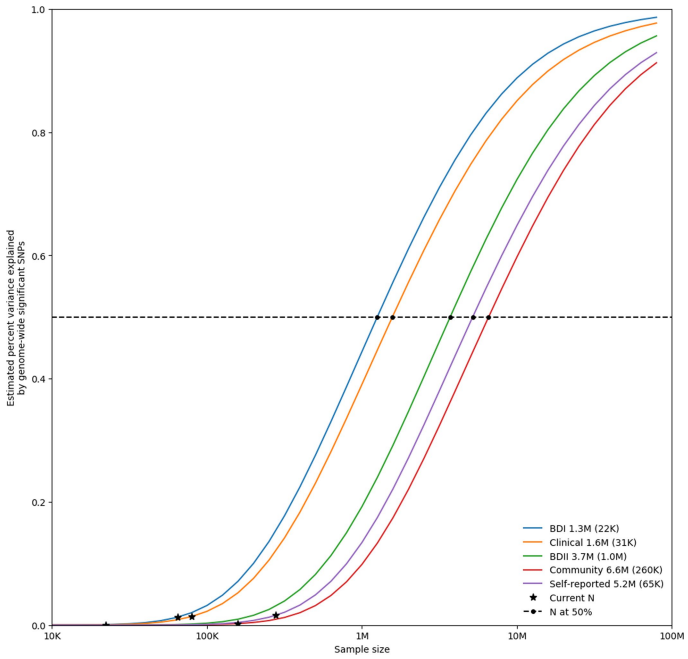
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41586-024-08468-9.

Correspondence and requests for materials should be addressed to Kevin S. O'Connell or Ole A. Andreassen.

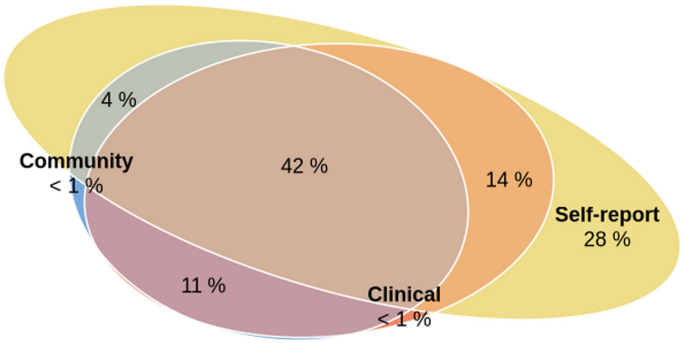
Peer review information *Nature* thanks Ditte Demontis, Veera Rajagopal and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.



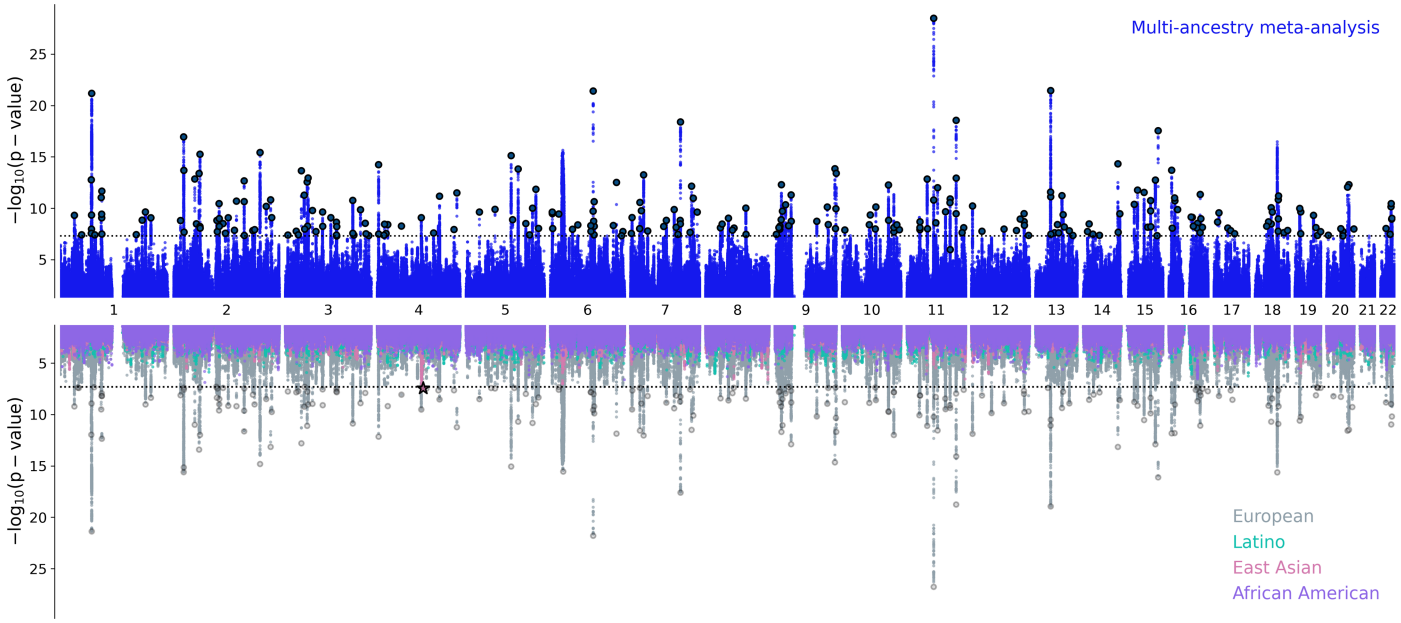
Extended Data Fig. 1 | Network diagram of the genetic correlations between BD ascertained from Clinical, Community and Self-report samples, as well as BD-subtypes (BDI and BDII). The line widths are proportional to the strength of the correlations between pairs. BDI: bipolar disorder I, BDII: bipolar disorder II.



Extended Data Fig. 2 | Univariate MiXeR estimates of the required effective sample size needed to capture 50% of the genetic variance (horizontal dashed line) associated with each BD ascertainment and subtype. N and Sample size refer to the effective sample size. The estimated effective sample size (and standard errors) are given in the legend alongside each trait name.

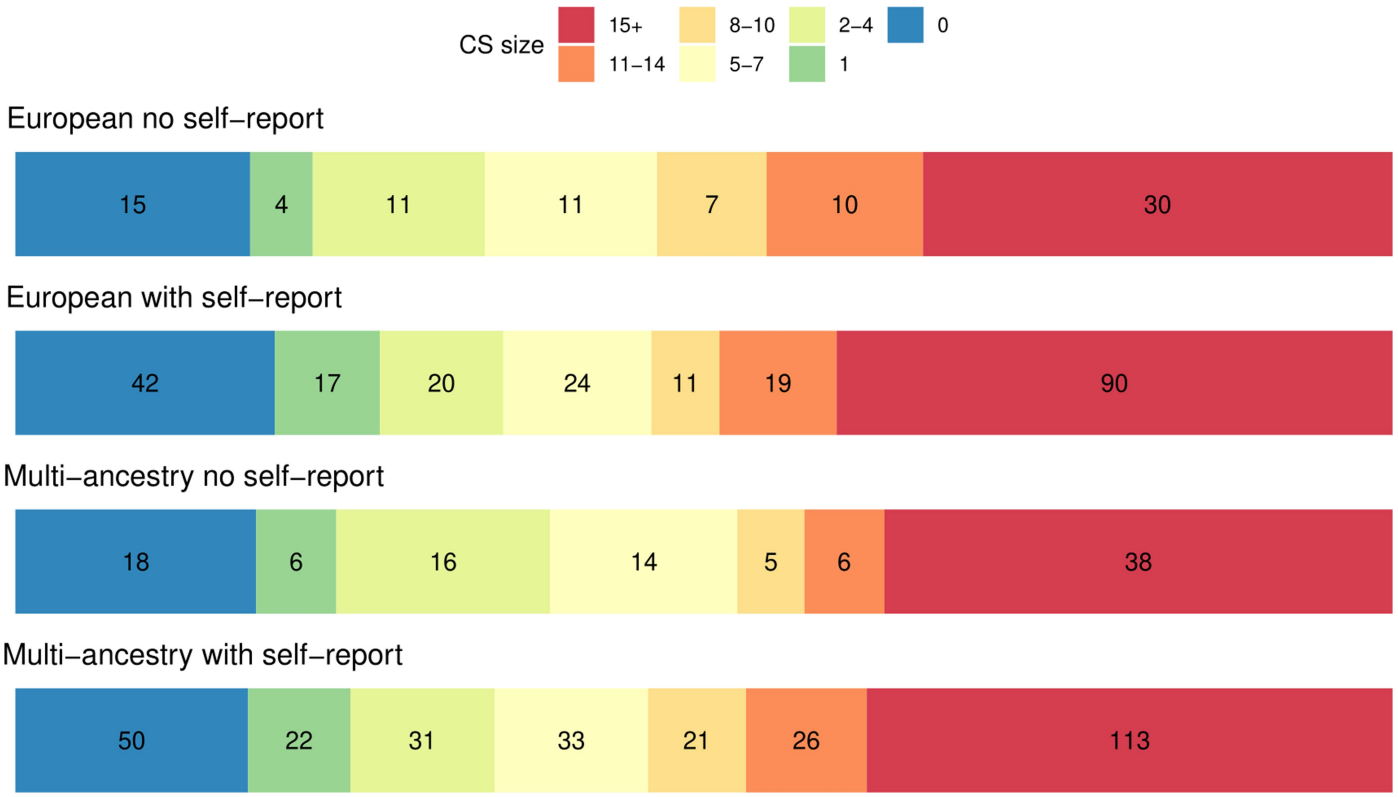


Extended Data Fig. 3 | Trivariate MiXeR estimates for the genetic overlap of BD from Clinical, Community and Self-report samples. The percentages show the proportion of trait-influencing variants within each section of the Venn diagram relative to the sum of all trait-influencing variants across all samples. The size of the circles reflects the polygenicity of each trait.

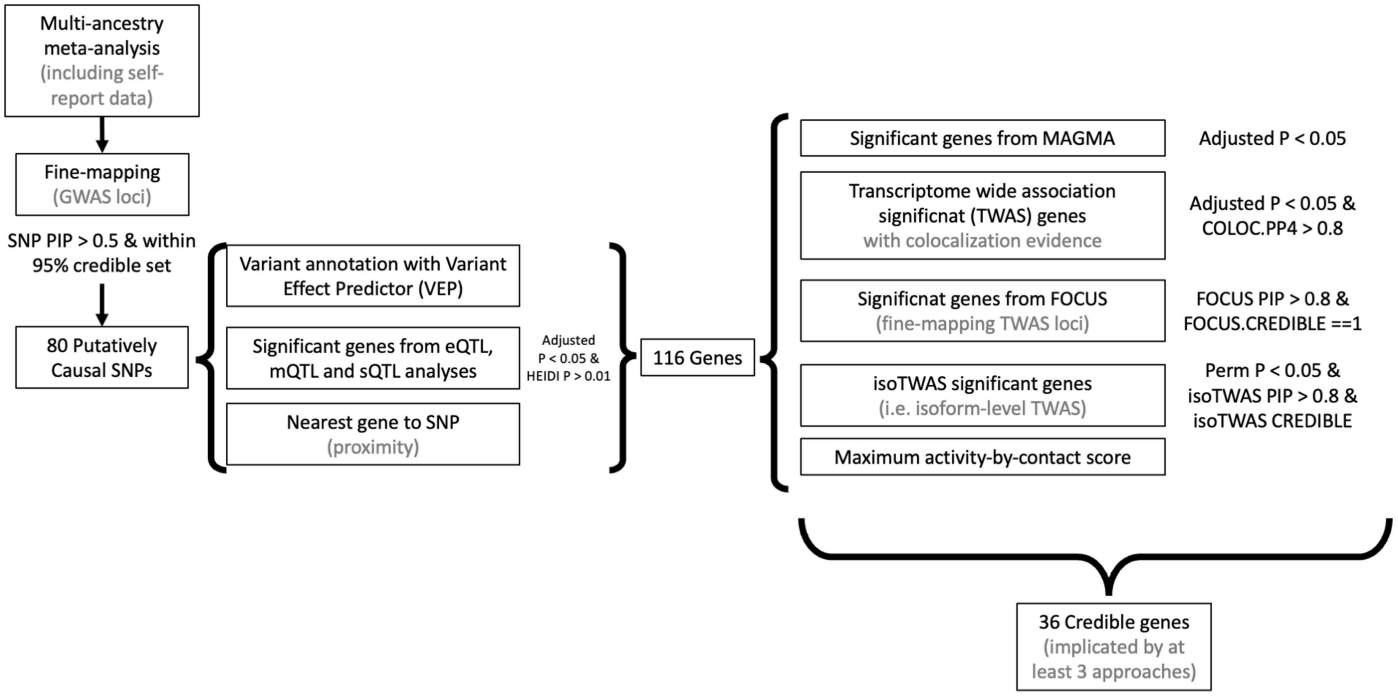


Extended Data Fig. 4 | Miami plot for BD genome-wide meta-analyses, including all cohorts. Upper panel: the multi-ancestry meta-analysis identified 298 genome-wide significant (GWS) loci. Lower panel: porcupine plot showing the results of the Latino (0 GWS loci), African American (0 GWS loci), East Asian (1 GWS locus) and European (229 GWS loci) meta-analyses. The x-axes show genomic position (chromosomes 1–22), and the y axes show statistical significance as −log10[p-value]. P-values are two-sided and based on an inverse-variance-weighted fixed-effects meta-analysis. The dashed black lines show the GWS threshold ($P < 5 \times 10^{-8}$). The star indicates the position of the East Asian GWS locus (rs117130410, 4:105734758, build GRCh37).

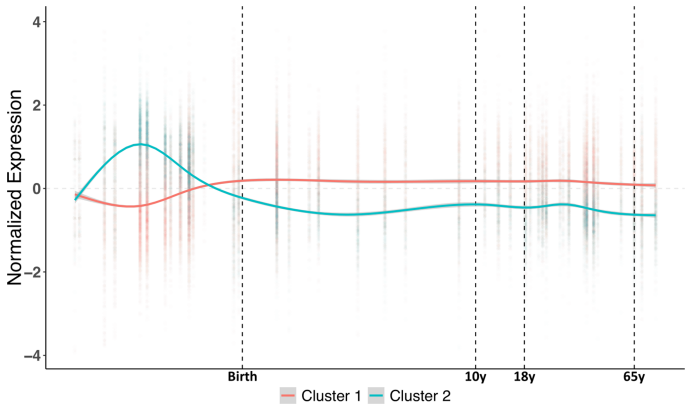
Extended Data Fig. 5 | Cluster-level SNP-heritability enrichment for bipolar disorder. The t-distributed stochastic neighbor embedding (tSNE) plot (left) (from Siletti et al.[23]) is coloured by the enrichment z-score. Grey indicates non-significantly enriched superclusters (FDR > 0.05). The barplot (right) shows the top 35 significantly enriched clusters. The numbers in parentheses on the y-axis indicate the cell type clusters as defined in Siletti et al.[23].



Extended Data Fig. 6 | Number of SNPs within the smallest 95% credible sets (CS) from meta-analysis of European and multi-ancestry meta-analyses when excluding and including self-report data. Colours represent CS of varying size, with blue CS containing 0 SNPs and red CS containing 15+ SNPs. All fine-mapped SNPs regardless of their PIP were used to assess the size of the 95% credible sets.



Extended Data Fig. 7 | Methods and criteria for credible gene identification.



Extended Data Fig. 8 | Clustering of patterns of temporal variation in expression of 34 credible genes. Cluster 1 (n = 21 genes) shows reduced prenatal gene expression, with gene expression peaking at birth and remaining stable over the life-course. Cluster 2 (n = 13 genes) includes genes with a peak gene expression during fetal development with a drop-off in expression before birth. Genes within each cluster are described in Supplementary Table 31. To illustrate the variability in gene expression within each cluster we plot each donor expression value in each sampled brain region for the 34 credible genes as individual points. Smoothing splines used to illustrate the age trajectory for each cluster is based on generalized additive models with the predicted 95% confidence interval in grey. We use age in days to plot the variation in gene expression with the x-axis on a log2 scale and labels for birth, 10, 18, and 65 years of age as reference points.

nature portfolio

Corresponding author(s): Kevin S. O'Connell and Ole A. Andreassen

Last updated by author(s): Oct 4, 2024

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	No software was used to collect the data in this study
Data analysis	Genotype QC, imputation, association analysis, meta-analysis: RICOPILI (https://sites.google.com/a/broadinstitute.org/ricopili/) EIGENSTRAT v6.1.4 (https://www.hsph.harvard.edu/alkes-price/software/) Eagle v2.3.5 (https://alkesgroup.broadinstitute.org/Eagle/) Minimac3 (https://genome.sph.umich.edu/wiki/Minimac3) PLINK v1.90 (https://www.cog-genomics.org/plink2/) METAL (version 2011-03-25) (https://genome.sph.umich.edu/wiki/METAL_Documentation) DENTIST (https://github.com/Yves-CHEN/DENTIST) LDSC (https://github.com/bulik/ldsc) Popcorn (https://github.com/brielin/Popcorn) Downstream analysis: Complex Traits Genetics Virtual Lab web platform (CTG-VL; https://vl.genoma.io) multitrait-based conditional and joint analysis using GWAS summary data (mtCOJO) (https://yanglab.westlake.edu.cn/software/gcta/#mtCOJO) MiXeR v1.3 (https://github.com/precimed/mixer) PRS-CS (https://github.com/getian107/PRScs) GSA-MiXeR (https://github.com/precimed/gsa-mixer) MAGMA (https://ctg.cncr.nl/software/magma)

SynGO tool v1.2 (https://www.syngoportal.org/)
Variant Effect Predictor (VEP) (GRCh37) Ensembl release 75 (https://www.ensembl.org/info/docs/tools/vep/index.html)
Summary data-based Mendelian randomization (SMR) (v1.3) (https://yanglab.westlake.edu.cn/software/smr/)
TWAS and FOCUS (http://gusevlab.org/projects/fusion/)
isoTWAS (https://github.com/bhattacharya-a-bt/isotwas)
OpenTargets platform (https://genetics.opentargets.org/)

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

Genome-wide association summary statistics for these analyses are available at https://www.med.unc.edu/pgc/download-results/. The full GWAS summary statistics for the 23andMe datasets will be made available through 23andMe to qualified researchers under an agreement with 23andMe that protects the privacy of the 23andMe participants. Please visit https://research.23andme.com/collaborate/#dataset-access for more information and to apply to access the data. After applying with 23andMe, the full summary statistics including all analysed SNPs and samples in the GWAS meta-analyses will be accessible to the approved researchers. Individual-level data are accessible through collaborative analysis proposals to the Bipolar Disorder working group of the PGC (https://pgc.unc.edu/for-researchers/data-access-committee/data-access-information/). This study included some publicly available datasets accessed through dbGAP (PGC bundle phs001254.v1.p1).

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	Sex was included as a covariate in the analyses. Sex-based analyses were not considered in this manuscript due to the amount of data already being presented. Theses analyses are being considered for a separate manuscript in the future that can focus specifically on this aspect.
Reporting on race, ethnicity, or other socially relevant groupings	Participants were grouped into European, East Asian, Latino and African American ancestral groups. This was based on participant/recruiter reported information and confirmed with genetic information.
Population characteristics	The first five genetic principal components, and any others as required, were included as covariates in the GWAS of each cohort.
Recruitment	Participants were recruited using different approaches. These are outlined for each cohort in the supplementary materials (cohort descriptions) of the manuscript.
Ethics oversight	These are listed for each cohort in the supplementary materials.

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	The final multi-ancestry meta-analysis included up to 158,036 cases and 2,796,499 controls aggregated from 79 cohorts. No sample size calculations were performed. This sample size is hypothesized to be insufficient to explain the full extent of genetic liability to bipolar disorder however, it is the largest study of its kind to date.
Data exclusions	Technical quality control was performed separately on each cohort for which individual-level data were provided separately according to standards developed by the PGC including; SNP missingness<0.05 (before sample removal), subject missingness<0.02, autosomal heterozygosity deviation (Fhet<0.2), SNP missingness<0.02 (after sample removal), difference in SNP missingness between cases and controls<0.02, SNP Hardy–Weinberg equilibrium (P>1×10−10 in BD cases and P>1×10−6 in controls), and mismatches between pedigree and genetically-determined sex based on the F statistic of X chromosome homozygosity (female F < 0.2 and male F >0.8). In addition, relatedness

was calculated across cohorts using identity by descent and one of each pair of related individuals (pi_hat>0.2) was excluded, prioritising exclusion of individuals related to the most others, controls over cases, and individuals from larger cohorts.

These exclusions follow standard guidelines in the field.

Replication

All available cohorts of bipolar disorder cases and controls were included in the primary multi-ancestry meta-analysis and therefore we do not perform replication for significant loci identified from these analyses.
We did test replication of previously identified loci associated with bipolar disorder from European ancestry. These results are detailed in the manuscript. Briefly, 31/64 previous loci met genome-wide significant in the present analysis containing all samples, and of the 33 that did not, 25 met genome-wide significant in either the Clinical samples or in the meta-analysis that excluded Self-reported data. Moreover, the direction of association for all top SNPs (12,151 SNPs with $p < 1 \times 10–5$) from the previous GWAS was consistent with the direction of association in this multi-ancestry meta-analysis of all samples

Randomization

This is an observational, genetic epidemiological study, and as such no randomization was performed.

Blinding

No disease-status blinding was used in recruitment or analysis. In principle, participants were recruited blind to their genotype and we do not expect to observe bias in association between genotype and disease-status.

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

NA

Novel plant genotypes

NA

Authentication

NA