Genetic diversity enhances gene discovery for bipolar disorder

Bipolar Disorder Working Group of the Psychiatric Genomics Consortium, 23andMe Research Team

Abstract

Bipolar disorder (BD) is a severe, highly heritable mental illness. The underlying mechanisms remain largely unknown. To gain greater insight, we performed the largest genome-wide association study (GWAS) meta-analyses of BD, combining clinical and community (biobank and self-report) samples of European, East Asian, African American and Latino ancestry. We detected 337 independent genome-wide significant variants mapped to 298 loci in the multi-ancestry meta-analysis, a 4-fold increase over previous findings, and a novel ancestral-specific locus in the East Asian cohort. Fine-mapping and integration of eQTL data implicated 47 credible genes in the etiology of BD. The genetic architecture of BD in community-based samples was more similar to BD type II than to BD type I, potentially reflecting a non-hospitalized, non-psychotic portion of the BD spectrum.

Introduction

Bipolar disorder (BD) is a mental disorder characterized by episodes of expansive or irritable mood and high energy[1]. BD is a major public health problem and a leading contributor to the global burden of disease[2]. BD is often associated with recurrent suicidality and impaired functional level and quality of life. Individuals affected by the disorder also have a high degree of comorbidity with other psychiatric conditions, such as anxiety and substance use, and other chronic medical conditions, such as metabolic syndrome, migraine and type 2 diabetes[2]. BD is a multifactorial disorder with a high heritability (60-80%)[3], and mostly affects people starting in early adulthood. Clinically, BD is often classified into two main subtypes, bipolar I disorder (BDI) and bipolar II disorder (BDII). BDI is described by manic episodes, which commonly alternate with depressive episodes, while BDII is characterized by the occurrence of at least one hypomanic and one depressive episode[4]. These subtypes each have a life-time risk in the population of around 1 %[5]. The disorder affects individuals irrespective of nationality, ethnicity, or socioeconomic status, and frequently is a chronic, remitting-relapsing condition[2]. Current treatments include lithium and mood stabilizers, together with psychosocial interventions[1,2].

To date, disease mechanisms remain mostly unknown, but advances in genetics and neuroimaging have begun to make inroads. The Psychiatric Genomics Consortium (PGC) Bipolar Disorder Working Group has spearheaded genetic discoveries in BD[6,7]. A genome-wide association study (GWAS) of 41,917 BD cases and 371,549 controls of European ancestry identified 64 genomic loci and highlighted calcium channel antagonists as potential targets for drug repurposing for BD treatment[8]. Brain imaging studies have shown widespread patterns of lower cortical thickness, subcortical volume and disrupted white matter integrity associated with BD, as well as mapping brain alterations caused by medications[9].

Here we present results from a much larger, multi-ancestry GWAS meta-analysis of 156,643 BD spectrum cases and 2,799,462 controls, combining clinical, biobank and self-report samples. We identified 337 independent genome-wide significant variants mapped to 298 loci that provide new insight into the genetic architecture, and highlight novel neurobiological mechanisms involved in BD, with the potential for the development of new treatments.

Results

Sample description

The studies included in these meta-analyses used a range of measures to define bipolar disorder (BD): structured clinical interviews (clinical), medical records (biobank), and self-report surveys (self-report). The clinical samples included 27,543 cases and 44,461 controls from 52 cohorts. The biobank samples totaled 40,405 cases and 823,249 controls from 23 cohorts. In addition, 90,088 self-report cases and 1,928,789 self-report controls from four cohorts were included. Thus, the total available samples for analyses included 158,036 cases across the BD spectrum and 2,796,499 controls. Details of the cohorts, including sample size, ancestry, and inclusion/exclusion criteria for cases and controls are provided in Supplementary Table 1 and the Supplementary Note.

Ancestry-specific GWAS meta-analyses

We first conducted GWAS meta-analyses stratified by ancestral groups. For each ancestral group, meta-analyses were performed with and without the inclusion of the self-report data. Table 1 provides a summary of the GWAS meta-analyses. The various meta-analyses included up to 131,969 cases and 2,322,416 controls (N_{eff} = 441,002) of European ancestry, 7,076 cases and 91,350 controls (N_{eff} = 23,484) of African American ancestry, 5,969 cases and 145,911 controls (N_{eff} = 22,610) of East Asian ancestry, and 13,022 cases and 236,822 controls (N_{eff} = 48,357) of Latino ancestry.

Table 1. Sample size (cases/controls), ascertainment type and discovery from ancestry-specific and multi-ancestry meta-analyses

Meta-analysis	Clinical (52 cohorts)	Biobank (23 cohorts)	Clinical + Biobank	Loci[a]	Indep GWS signals*	Self-report (4 cohorts)	Clinical + Biobank + Self-report	Loci[b]	Indep GWS signals*
European (69 cohorts)	27,196 / 43,792	32,091 / 737,230	59,287 / 781,022	88	94	72,682 / 1,541,394	131,969 / 2,322,416	229	261
African American (4 cohorts)	347 / 669	2,803 / 7,204	3,150 / 7,873	0	0	3,926 / 83,477	7,076 / 91,350	0	0
East Asian (4 cohorts)**	4,479 / 75,725	0 / 0	4,479 / 75,725	1	1	1,490 / 70,186	5,969 / 145,911	1	1
Latino (2 cohorts)	0 / 0	1,032 / 3,090	1,032 / 3,090	0	0	11,990 / 233,732	13,022 / 236,822	0	0
Multi-ancestry (79 cohorts)	27,543 / 44,461	40,405 / 823,249	67,948 / 867,710	105	116	90,088 / 1,928,789	158,036 / 2,796,499	298	337

[a] Loci from the meta-analysis of clinical and biobank cohorts

[b] Loci from the meta-analyses of all cohorts (clinical, biobank and self-report)

* number of independent genome-wide significant SNPs after conditional analysis

** the locus for the East Asian meta-analysis of clinical and biobank cohorts is different to that identified for the meta-analysis of all cohorts

The European ancestry meta-analysis that included all cohorts identified 261 independent genome-wide significant (GWS) variants ($P < 5 \times 10^{-8}$) in 229 loci (Figure 1, Supplementary Table 2). The LDSC intercept was 1.050 (s.e. = 0.016), and the attenuation ratio was 0.065 (s.e. = 0.015), indicating that 93.4% of the observed inflation in test-statistics was due to polygenic signal. In the meta-analysis excluding self-report data (59,287 cases and 781,022 controls; N_{eff} = 163,366), we identified 94 independent GWS variants ($P < 5 \times 10^{-8}$) in 88 loci (LDSC intercept=1.055 (s.e. = 0.012), attenuation ratio=0.097 (s.e. = 0.019)). Forty-eight of the 88 loci identified when excluding self-report data are included among the 229 loci identified when including the self-report data, and the remaining 40 loci had p-values < 3.15×10^{-3} (Supplementary Table 3). There is 100% concordance of effect direction for the 342 independent GWS variants identified in the European ancestry meta-analyses, with and without the inclusion of the self-report data (Supplementary Tables 2 and 3), suggesting that these variants are within the shared component between clinical, biobank and self-report estimated by trivariate MiXeR analysis (see Genetic architecture of BD and Figure 2). Analysis of the X chromosome in the subset of cohorts for which individual level data were available (33,399 cases and 53,467 controls) did not identify any GWS loci.

We identified one GWS locus in the East Asian ancestry meta-analysis that included all cohorts (Figure 1, Supplementary Table 4). The G allele of rs117130410 (4:105734758, build GRCh37) was associated with increased risk of BD (OR = 1.31, SE = 0.05, $P = 3.69 \times 10^{-8}$). The nearest protein coding genes to this lead SNP were *TET2* and *CXXC4*, approximately 300 kb up- and downstream respectively. While this variant had a frequency of 16% and 9% in East Asian BD cases and controls, respectively, it is monomorphic in non-Asian populations. We identified a second locus associated with BD in the East Asian ancestry meta-analysis that excluded the self-report data (4,479 cases and 75,725 controls; N_{eff} = 16,772). The C allele of rs174576 (11:61603510, build GRCh37) was associated with reduced risk of BD (OR = 0.86, standard error (SE) = 0.03, $P = 7.78 \times 10^{-9}$). This variant was not present in the self-report dataset and was therefore filtered out (<75% N_{eff}) in the East Asian meta-analysis that included the self-report data. This locus includes the *FADS1* and *FADS2* genes and has previously been associated with BD in the Japanese cohort included in this meta-analysis[10]. No additional independent significant signals were identified within either of these loci.

No variants were associated at genome-wide significance in the GWAS in samples of African American or Latino ancestry.

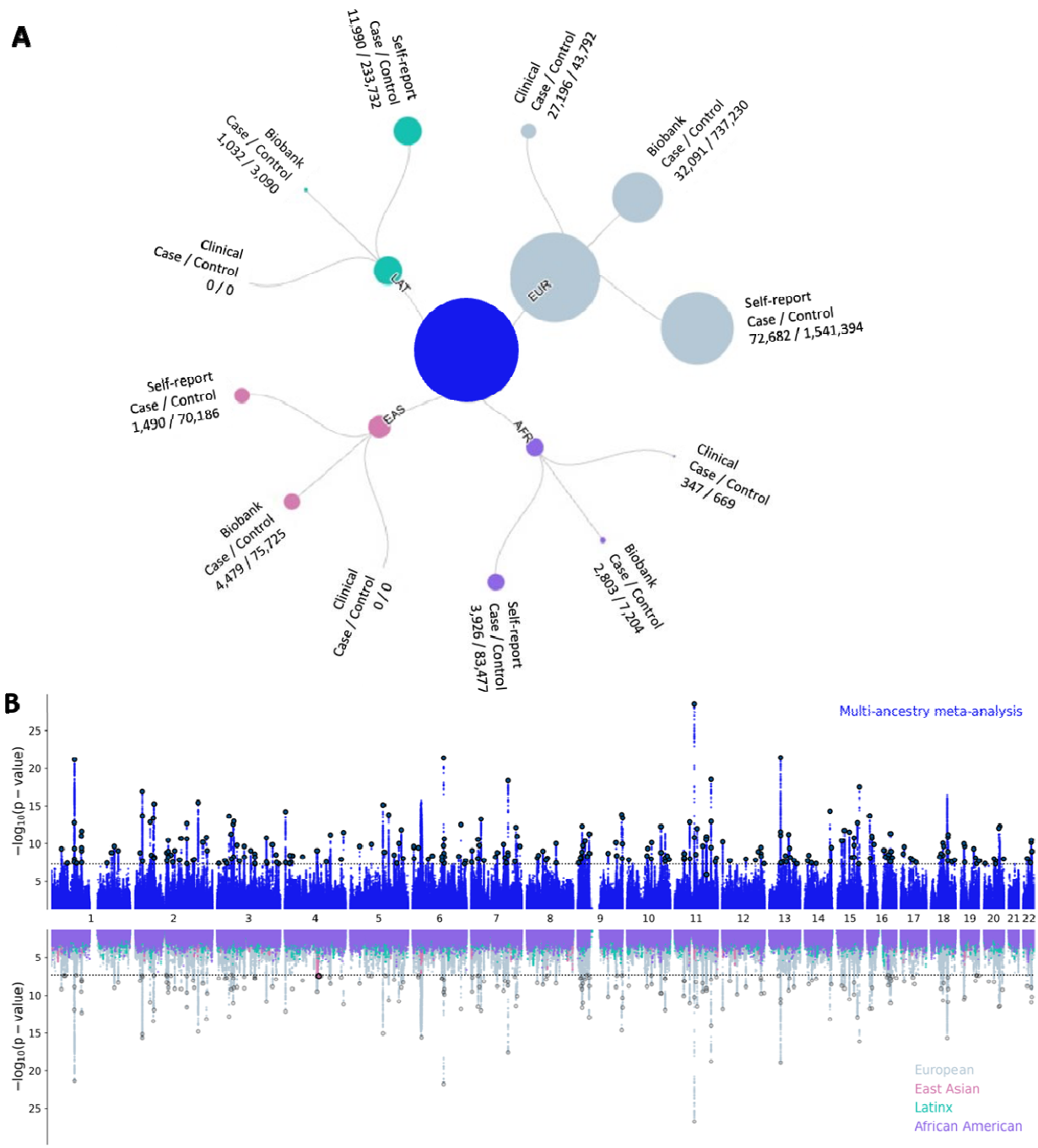

Figure 1. (A) Circular dendrogram of samples included in the ancestry-specific and multi-ancestry meta-analyses. The size of the circles corresponds to the sample size as a proportion of the full sample (n cases = 158,036, n controls = 2,796,499). (B) Miami plot for BD genome-wide meta-analyses, including all cohorts (clinical, biobank and self-report data). 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 genome-wide significance threshold ($P < 5 \times 10^{-8}$).

Multi-ancestry meta-analysis

We carried out two multi-ancestry meta-analyses of African American, East Asian, Latino and European samples, again with and without the inclusion of the self-report data. The meta-analysis excluding the self-report data comprised 74 cohorts totalling 67,948 cases and 867,710 controls (N_{eff} = 191,722). There was no evidence of residual population stratification (LDSC intercept = 1.050 (se = 0.012), attenuation ratio=0.086 (s.e. = 0.018)). We identified 116 independent GWS variants mapping to 105 loci (Table 1, Supplementary Table 5). The meta-analysis that included the four self-report cohorts (n cases = 158,036, n controls = 2,796,499, N_{eff} = 535,812) identified 337 independent GWS variants mapping to 298 loci (Figure 1, Table 1, Supplementary Table 6), with no evidence of residual population stratification (LDSC intercept = 1.052 (se = 0.016), attenuation ratio=0.071 (s.e. = 0.013)).

Fifty-nine of the 105 loci identified when excluding self-report data are also included in the 298 loci identified when including the self-report data, and the remaining 46 loci had top SNPs with p-values $< 5 \times 10^{-3}$ (Supplementary Tables 5 and 6). Moreover, these 46 loci include all 40 loci not identified in the European ancestry analysis that included the self report data. Overall, there is 100% concordance of effect direction for the 434 independent GWS variants identified in the muli-ancestry meta-analyses, with and without the inclusion of the self-report data (Supplementary Tables 5 and 6). As in the case of the analysis of European samples above, this complete concordance of direction for associated variants suggests that these variants fall within the shared component between clinical, biobank and self-report estimated by trivariate MiXeR analysis (see Genetic architecture of BD and Figure 2).

Genetic architecture of BD

In addition to the EUR meta-analyses described above, meta-analyses were also performed separately for clinical, biobank, and self-report data, and for BD subtypes in European ancestry cohorts (Supplementary Table 1). LDSC[68] was used to estimate the SNP-heritability (h^2_{SNP}) on the liability scale (assuming a lifetime BD prevalence of 1-2%) of BD from each of these meta-analyses, including all cohorts as well as sub-groups by ascertainment and BD subtype (Supplementary Table 7). At a population prevalence of 2%, BDI was the most heritable phenotype (h^2_{SNP} = 0.25 (s.e. = 0.01)), followed by BD ascertained from clinical cohorts (h^2_{SNP} = 0.22 (s.e. = 0.01)). Conversely, BD ascertained from biobank cohorts (h^2_{SNP} = 0.05 (s.e. = 0.003)) and self-report BD (h^2_{SNP} = 0.08 (s.e. = 0.003)) had the lowest heritability estimates. The same pattern is observed when assuming a population prevalence of 1%.

Moreover, we performed genetic correlation and trivariate MiXeR[11,12] analyses to further investigate the genetic architecture of BD based on ascertainment and subtype. Despite a high genetic correlation between BD ascertained from clinical and biobank samples (r_g = 0.95 (s.e. = 0.03)), biobank ascertained BD showed considerably higher genetic correlation with self-report BD(r_g = 0.79 (s.e. = 0.02)) than with BD ascertained from clinical samples (r_g = 0.47 (s.e. = 0.02)). We observed similar results when stratifying the BD cases into BDI and BDII subtypes, with a high genetic correlation observed between the subtypes (r_g = 0.88 (s.e. = 0.05)), but differing correlations with self-report BD (BDI r_g = 0.42 (s.e. = 0.02), BDII r_g = 0.76 (s.e. = 0.05)). MiXeR estimates show greatest polygenicity for BD ascertained via self-report, than from clinical and finally from biobank samples (Figure 2A, Supplementary Table 8).

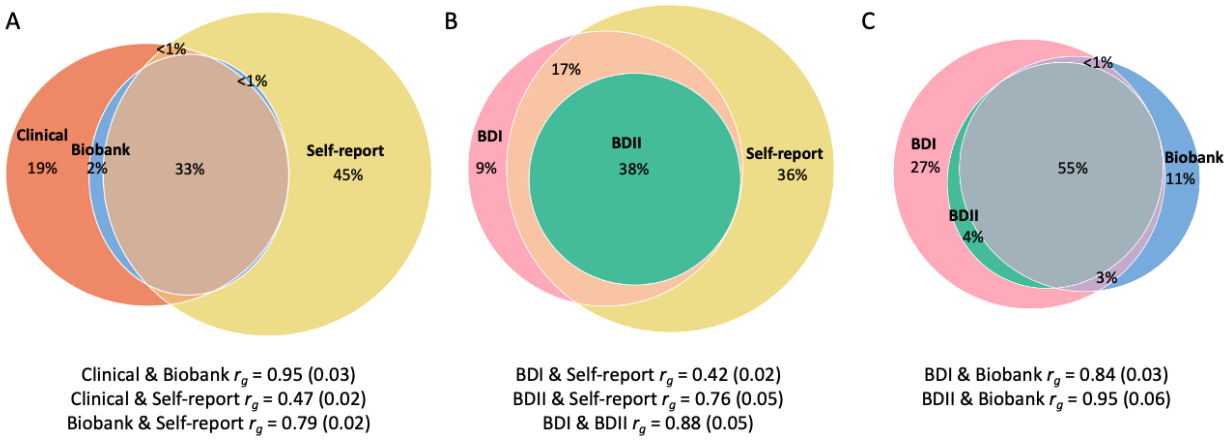

Figure 2. Trivariate MiXeR estimates for the genetic overlap of BD ascertained from (A) clinical, biobank and self-report samples, (B) bipolar I disorder (BDI), bipolar II disorder (BDII) and self-report BD, and (C) bipolar I disorder (BDI), bipolar II disorder (BDII) and BD ascertained from biobank 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.

Almost all variants estimated to influence BD in biobank samples were shared with BD ascertained from clinical samples. The majority of clinical and biobank BD-influencing variants were also shared with self-report BD, but trivariate MiXeR estimated that only 33% of all variants were shared among all three ascertainment types (Figure 2A). Moreover, when considering the BD subtypes, almost all variants estimated to influence BDII are shared with those influencing BDI (Figure 2B and 2C), and the majority of variants influencing BDI and BDII are shared with those estimated to influence both self-report (Figure 2B) and biobank BD (Figure 2C).

The median correlation of variant effects, estimated with MiXeR, in the shared components was also high across all comparator groups (biobank and self-report r_{g_shared} = 0.78 (s.e. = 0.02), biobank and clinical r_{g_shared} = 0.98 (s.e. = 0.01) and clinical and self-report r_{g_shared} = 0.73 (s.e. = 0.05)) (Figure 2A, Supplementary Table 8), supporting our decision to meta-analyze these three BD ascertainment types. Similarly, the median correlation in the shared component between BDI and BDII was also high (r_{g_shared} = 1.00 (s.e. = 0.004)). Interestingly, the median correlation of variant effects in the shared components between BDI and BDII and biobank and self-report BD was predominantly high (BDII and self-report r_{g_shared} = 1.00 (s.e. = 0.02), BDI and biobank r_{g_shared} = 1.00 (s.e. = 0.02) and BDII and biobank r_{g_shared} = 0.97 (s.e. = 0.01)), except for BDI and self-report BD (r_{g_shared} = 0.52 (s.e. = 0.06)) (Figure 2B and 2C, Supplementary Table 8).

Genome-wide genetic correlations (r_g) were also estimated between EUR BD GWASs (with and without self-report) and human diseases and traits via the Complex Traits Genetics Virtual Lab (CTG-VL—http://genoma.io) web platform [13]. After correction for multiple testing, BD showed significant positive r_g with seven out of 11 other psychiatric disorders tested, although the pattern of these correlations shifted depending on the inclusion of the BD self-report data (Supplementary Table 9). Specifically, major depressive disorder was more strongly correlated when self-report data was included (BD without self report r_g = 0.52 (s.e. = 0.03), BD with self-report r_g = 0.77 (s.e. = 0.02)) and schizophrenia was more strongly correlated when self-report data was excluded (BD without self report r_g = 0.65 (s.e. = 0.02), BD with self-report r_g = 0.54 (s.e. = 0.02)). In addition, BD was significantly correlated with 487 and 120 out of 1390 tested phenotypes, with and without inclusion of the self-report data, respectively (Supplementary Tables 10 and 11). The strongest correlations were with depression-related traits and substance dependence.

Evidence for similar genetic architectures across ancestries

We employed cross-population empirical Bayes (XPEB)[14], to assess the extent of overlapping genetic architecture across ancestries, and whether loci identified in the European meta-analysis display evidence for association with BD in the African American, East Asian or Latino ancestry samples. The output of XPEB includes a population-wide estimate of the degree of genetic architecture overlap ($\kappa 1$; ranging between 0 and 0.90) between the analyzed cohorts, with higher values indicating that more SNPs showing association evidence in the non-EUR GWAS are likely to show strong association in the EUR GWAS as well. The overlapping genetic architecture between the European and East Asian samples was estimated as $\kappa 1$=0.90, both with and without inclusion of self-report data. The estimate of overlapping genetic architecture was similarly high between European and Latino samples ($\kappa 1$=0.90). Estimates were considerably lower between European and African American samples $\kappa 1$=0.02 (without self-report) and $\kappa 1$=0.45 (with self-report). Applying XPEB to the East Asian and Latino samples, we identified 3 and 11 independent significant variants (FDR < 0.05) (Supplementary Table 12), respectively. No significant variants were identified when targeting the African American sample.

Sex-stratified results

Sex-stratified analyses were performed in European ancestry cohorts for which individual-level genotype and phenotype data were contributed and available. The analyses included 20,473 cases and 29,506 controls (N_{eff} = 43,578) in the female meta-analysis and 13,795 cases and 25,818 controls (N_{eff} = 31,394) in the male meta-analysis, all of European ancestry. The female meta-analysis identified 10 GWS loci. The LDSC intercept was 1.025 (s.e.□=□0.009). The male meta-analysis identified 4 GWS loci with an LDSC intercept of 1.005 (s.e.□=□0.009). For one locus, the same lead SNP rs9834970 (3:36856030, build GRCh37) reached genome-wide significance in both the male (OR T allele = 0.90, SE = 0.02, P = 6.6×10^{-10}) and female (OR T allele = 0.91, SE = 0.01, P = 5.7×10^{-12}) meta-analyses and has been consistently associated with bipolar disorder previously[7,8]. The remaining loci identified in the female and male meta-analyses were sex-specific (Supplementary Table 13), however all of these loci overlap with those identified in the unstratified European meta-analyses. MiXeR estimated greater polygenicity for BD in males, compared to females, with approximately 76% of all variants shared between BD in females and males (Supplementary Figure 2). The genetic correlation between BD in males and females was *rg* = 0.92 (s.e = 0.03, P = $3.08\text{x}10^{-167}$), which is significantly different from 1 (P = 0.008).

Polygenic risk scoring

Polygenic risk score (PRS) analyses were performed in European ancestry cohorts for which individual-level genotype and phenotype data were available, as well as one cohort of African American and two cohorts of East Asian ancestry, respectively. The samples for PRS analyses totalled 36,129 cases and 55,826 controls (Supplementary Tables 14-17). The PRS with the most explanatory power was based on risk alleles derived from the larger multi-ancestry GWAS rather than the European only GWAS. The median R^2 on the liability scale, based on the weighted median R^2 across cohorts (Figure 3A, Supplementary Tables 14-17) and assuming 1-2% of disease prevalence, was 0.076-0.090 (s.e. = 0.016-0.019). The median Area Under the Receiver Operating Characteristic Curve (AUC) was 0.668 (95% CI= 0.653-0.725) to 0.702 (CI = 0.668-0.737); therefore the liability explained remains insufficient for diagnostic prediction in the general population. The median R2 across 37 cohorts (12,429 cases/33,463 controls, N_{eff} = 14,626) was 0.107 – 0.126 (s.e = 0.022-0.026) in BDI cases (Figure 3B), assuming 1-2% of disease prevalence. PRS explained most variance in liability when using as the discovery sample the multi-ancestry GWAS without the self-report data (Figure 3C); these clinical and biobank cohorts likely included more severe cases than the self-report cohorts. The median R2 in BDII cases was higher when the self-report data were incorporated into the discovery sample (Figure 3B).

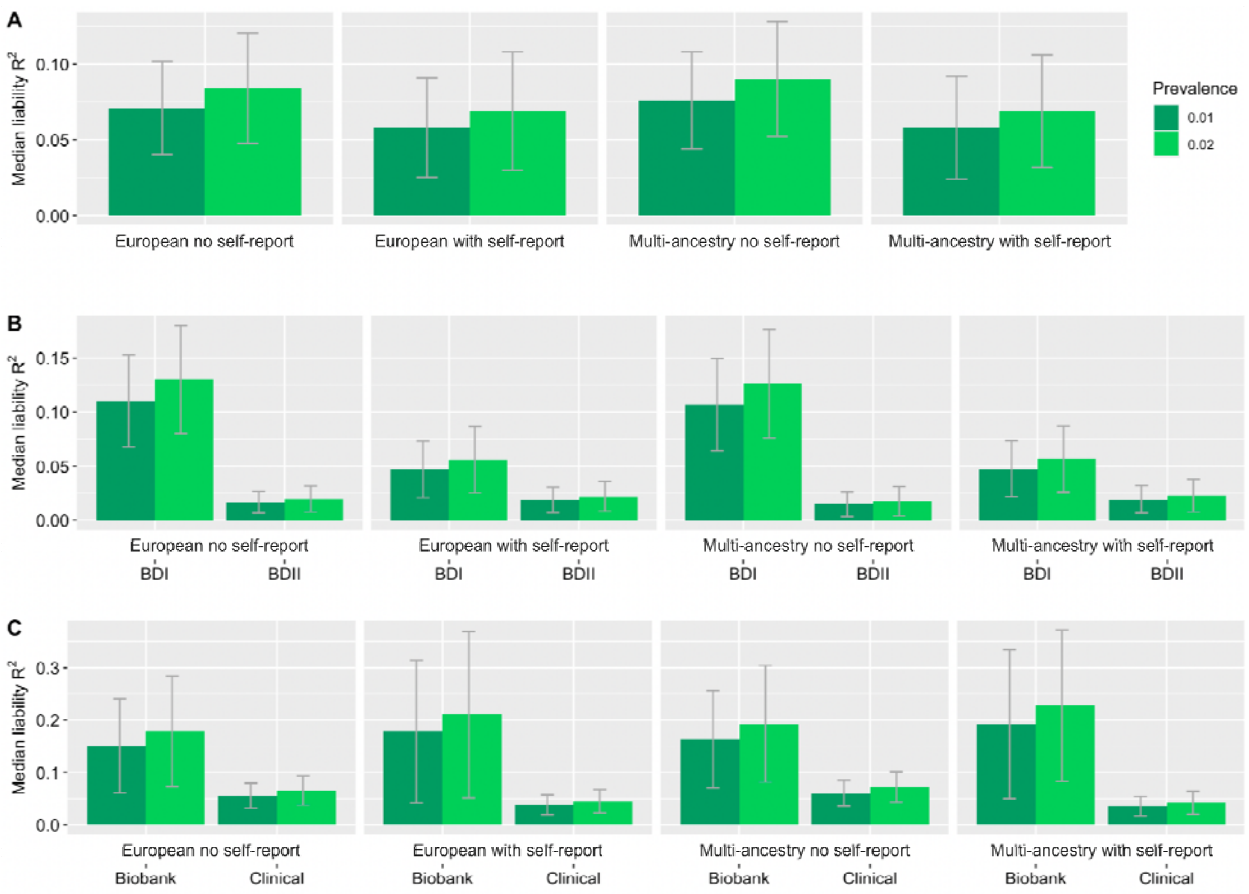

Figure 3. Phenotypic variance in bipolar disorder explained by polygenic risk scores
Variance explained is presented on the liability scale, assuming a 1-2% population prevalence of bipolar disorder. (A) The results shown are the weighted median R^2 values across 55 cohorts (40,992 cases/80,215 control, N_{eff} = 44,722), weighted by the effective *n* per cohort. Error bars show 95% confidence intervals (CI). (B) The results shown are the weighted median R^2 values across 37 cohorts (37 Bipolar disorder I (BDI) cohorts, 12,429 cases/33,463 controls, N_{eff} = 14,626 and 26 Bipolar disorder II (BDII cohorts, 2568 cases/25,987 controls, N_{eff} = 4054), weighted by the effective *n* per cohort. Error bars show 95% CI. (C) The results shown are the weighted median R^2 values across 55 cohorts (7 biobank /48 clinical)(13,159 cases/36,592 controls, N_{eff} = 17,178 and 27,833 cases/46,623 controls, N_{eff} = 29,543), weighted by the effective *n* per cohort. Error bars show 95% CI.

When the multi-ancestry PRS was applied to an African American target cohort (GAIN_AFR (clinical ascertainment); n cases = 347; n control = 669), assuming 2% disease prevalence, the inclusion of self-report data greatly increased the explained variance (no self-report R2 = 0.010, with self-report R2 = 0.023). A similar pattern was observed for the European ancestry PRS (no self-report R2 = 0.010, with self-report R2 = 0.022) (Supplementary Table 18). Similar to the European ancestry results above, PRS analysis of two clinically ascertained East Asian (Taiwanese (n cases = 967; n controls = 8,376) and Japanese (n cases = 2,964; n controls = 61,887)) cohorts revealed that the PRSs derived from GWAS excluding the self-report data

(Taiwan; EUR-PRS R2 = 0.069, Multi-PRS R2 = 0.075. Japan; EUR-PRS R2 = 0.027, Multi-PRS R2 = 0.025) performed better than those which included self-report data (Taiwan; EUR-PRS R2 = 0.026, Multi-PRS R2 = 0.036. Japan; EUR-PRS R2 = 0.015, Multi-PRS R2 = 0.015) (Supplementary Table 18).

Gene enrichment analyses

Genome-wide gene-based and gene-set analyses were performed for the multi-ancestry GWAS summary statistics, and for each ancestry group separately, both with and without self-report data, using MAGMA[16] (Supplementary Tables 19-29). For the multi-ancestry meta-analysis including self-report data we identified significant enrichment of BD associated variants in 430 genes (Supplementary Table 23) and 6 gene-sets (Supplementary table 29) related to the synapse and transcription factor activity. The association signal was enriched among genes expressed in the brain (Supplementary Tables 25 and 26), and single-cell enrichment analyses of brain cell types were largely consistent with findings from the previous PGC BD GWAS[8], indicating involvement of neuronal populations from different brain regions (Supplementary Figure 1). In human brain samples, signal enrichment was mostly observed in hippocampal pyramidal neurons and interneurons of the prefrontal cortex and hippocampus. Similar signals were also observed in murine brain samples (Supplementary Fig. 1). In addition to MAGMA enrichment analysis, we employed an exploratory analysis using GSA-MiXeR[17], which fits a model for the heritability of individual genes, accounting for linkage disequilibrium across variants, and allows the quantification of partitioned heritability and fold enrichment for small gene-sets. GSA-MiXeR highlighted enrichment of specific dopamine- and calcium-related biological processes and molecular functions, as well as GABAergic interneuron development, driven by the *DRD2* and *CACNA1B* genes, respectively (Supplementary Table 30).

Single-cell enrichment analysis in 914 cell types across 29 non-brain murine tissues, implemented in FUMA[18], identified significant enrichment in the enteroendocrine cells of the large intestine and delta cells of the pancreas (Supplementary Table 31).

When considering the gene targets of individual drugs, the targets of anticonvulsant pregabalin were significantly enriched (Supplementary Table 32). There was also significant enrichment in the targets of psycholeptics (specifically antipsychotics and anxiolytics) and beta-lactam antibacterials other than penicillins (Supplementary Table 33).

eQTL integration analyses

Bulk brain-derived eQTL data from PsychENCODE[19] and BrainMeta[20] were integrated with the GWAS results using TWAS as implemented in FUSION[21] and isoTWAS[22] as well as summary-data-based Mendelian randomization (SMR) with subsequent HEIDI (heterogeneity in dependent instruments) tests[23]. All analyses were performed for the multi-ancestry GWAS summary statistics, and for each ancestry group separately, both with and without self-report

data where possible (Supplementary Tables 34-40). In total, 71 genes were implicated by TWAS (Bonferroni P < 0.05, COLOC.PP4 > 0.8), 144 by isoTWAS (perm P < 0.05, PIP > 0.8) and 53 by SMR (Bonferroni P < 0.05, HEIDI > 0.01) for the multi-ancestry meta-analysis including self-report data (Supplementary Table 41). When considering the multi-ancestry meta-analysis excluding self-report data, 59, 152 and 42 genes were implicated by TWAS, isoTWAS and SMR, respectively (Supplementary Table 42). Results of the eQTL integration analyses based on ancestry-specific meta-analyses are provided in Supplementary Tables 34-40.

Fine-mapping results

Functional fine-mapping using Polyfun+SuSiE[24] was performed on the multi-ancestry, European and East Asian datasets, with and without the inclusion of the self-report data (Supplementary Tables 37-42). We also performed single-variant fine-mapping for both multi-ancestry datasets, i.e. excluding the use of an LD reference panel. The number of causal variants was set to five per locus and credible SNPs (part of a 95% credible set) were determined at two thresholds, PIP > 0.50 and PIP > 0.95. At a threshold of PIP > 0.50, we identified 64 and 83 credible SNPs for the European and multi-ancestry meta-analyses including self-report data, respectively. At the more stringent threshold of PIP > 0.95 we identified 17 and 22 credible SNPs. When the self-report data were excluded from the meta-analyses, we identified 21 and 33 credible SNPs (PIP > 0.5), and 4 and 6 credible SNPs (PIP > 0.95), from the European and multi-ancestry data, respectively. When comparing the number of SNPs within 95% credible sets, those derived from multi-ancestry meta-analyses (with or without self-report data) were smaller than those derived from corresponding European ancestry meta-analyses (Supplementary Figures 3-6). Considering 95% credible sets containing less than 10 SNPs, we identified 113 and 42 derived from multi-ancestry meta-analyses (with and without self-report data), while we identified only 70 and 33 from European ancestry meta-analyses (with and without self-report data), respectively. Finally, functional fine-mapping of the East Asian dataset including self-report data, highlighted rs174572 within the FADS2 locus as a SNP with a PIP above 0.50 and within a 95% credible set (Supplementary Tables 37).

Fine-mapped SNPs with a PIP > 0.5 were mapped to genes by performing variant annotation with Variant Effect Predictor (VEP) (GRCh37) Ensembl release 109[25]. This analysis identified 80 and 31 genes fine-mapped to loci from the multi-ancestry meta-analyses including and excluding self-report data, respectively. Similarly, we identified 62 and 17 fine-mapped genes from the European ancestry meta-analyses including and excluding self-report data, respectively (Supplementary Tables 39-42).

Converging evidence of common and rare variation

Within loci associated with BD in the multi-ancestry meta-analysis, genes (N = 80) implicated by fine-mapped variants were enriched for ultra rare (<=5 minor allele count) damaging missense and protein-truncating variants in both the BipEx[26] (Odds ratio (OR) = 1.16, 95% confidence

interval (CI) = 1.05 - 1.28, P = 0.002) and SCHEMA[27] data (OR = 1.21, 95% CI = 1.02 - 1.43, P = 0.024), respectively.

Identification of credible BD-associated genes

In order to gain insight into the genes that may underlie the identified BD-associated loci we integrated the results from the post-GWAS analyses described above. When considering the multi-ancestry meta-analysis including self-report data and the intersection of genes identified by the seven different gene-mapping approaches we identified a credible set of 47 genes identified by at least three of the described approaches (Supplementary Table 43). Within this subset, four genes (*DCC, PACSIN2, TMEM258, ZDHHC21*) were implicated by at least four approaches, three genes (*FURIN, MED24, TTC12*) by at least five, and one gene (*FADS1*) by six of the seven approaches (Supplementary Table 43).
Nine of the 47 credible genes have synaptic annotations in the SynGO database[28], six postsynaptic, three presynaptic and two genes not mapped to any specific compartment (Figure 4, Supplementary Table 44).

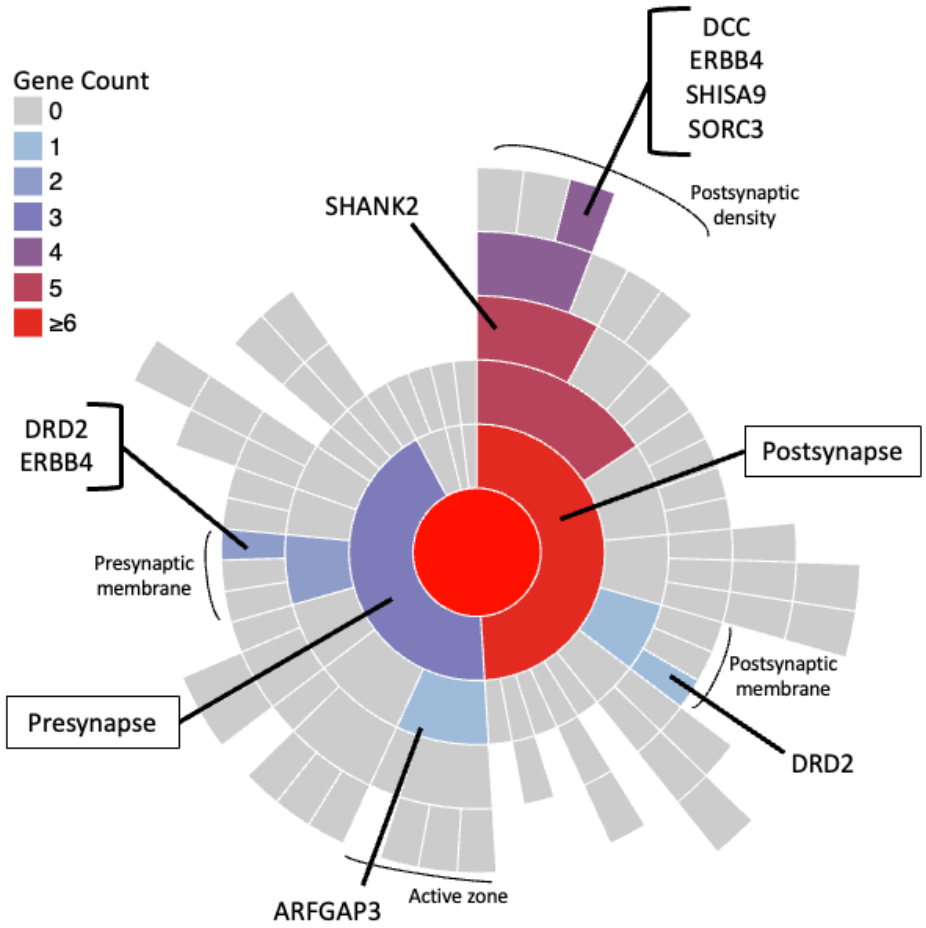

Figure 4. Sunburst plot depicting synaptic locations starting with the synapse (centre), pre- and postsynaptic locations in the first ring and child terms (that is, terms that are subsets of the adjacent inner ring) in subsequent rings. Locations enriched for credible BD genes, and the genes within these locations, are labeled. The number of genes in each term is indicated by the colour scheme in the legend. Credible genes (Supplementary table 44; N = 47), of which 9 are SynGO-annotated, 7 to cellular components.

Gene expression perturbation profiles of drugs versus credible gene expression

First, we used the DGIdb[29] to identify drug-gene interactions using the 47 credible genes as input and identified a total of 417 drugs. Of these 417 drugs, 168 were present in Connectivity Map (CMap)[30,31], a database of cellular signatures representing systematic perturbation with genetic and pharmacologic perturbagens. Of the 47 credible genes, TWAS z-values were available for 38 genes, of which 29 genes were also present in CMap[31]. For each drug, a correlation analysis between the drug-induced gene expression perturbations and the disorder-associated gene expression perturbations was performed, i.e., 168 correlation analyses each including 29 genes. Out of the 168 drugs, 17 showed nominally significant ($p < 0.05$) opposite gene expression perturbations in drug-induced expression versus bipolar-disorder-associated expression of the 29 corresponding genes. These drugs mainly include antipsychotics, anti-cancer drugs, and drugs with hormonal activity (Supplementary Table 45).

Discussion

We have performed the largest GWAS of BD, including diverse samples of European, East Asian, African American and Latino ancestry, resulting in an over four fold increase in the number of BD-associated loci. In a combined sample of 158,036 BD cases we identified 337 independent GWS variants mapping to 298 loci. Moreover, we investigated these common variant associations by applying an array of complementary post-GWAS methodologies to provide novel insights into the biological underpinnings and genetic architecture of BD.

Enrichment of the common variant associations from this multi-ancestry meta-analysis highlights the synapse as well as hippocampal pyramidal neurons and interneurons of the prefrontal cortex and hippocampus, when compared to other tissues and other brain cell types, respectively. Moreover, exploratory analyses using GSA-MiXeR[17] also suggest enrichment of dopamine- and calcium-related biological processes and molecular functions, as well as GABAergic interneuron development, highlighting interesting molecular mechanisms and pathways to consider as targets for drug-repurposing. Genes fine-mapped to associated loci were also shown to be enriched for ultra rare damaging missense and protein-truncating variation in the BipEx[26] and SCHEMA[27] datasets, respectively, highlighting convergence of common and rare variant signals as recently shown in schizophrenia[32].

In addition to enrichment analyses, we mapped genes to the 298 GWS loci using seven complementary approaches and identified a subset of 47 credible genes that were mapped to loci by at least three of these approaches. Thirteen of these are not the closest gene to the index SNP in the locus, highlighting the importance of probing other mechanisms of linking variants to genes. Moreover, examination of these 'credible' genes in the SynGo[28] database revealed nine to be annotated to pre- or postsynaptic locations and functions.

Multi-ancestry GWAS signals were also enriched in the gene targets of the anticonvulsant pregabalin, an active metabolite of gabapentin. A recent systematic review identified four double-blinded randomized control trials of gabapentin, not pregabalin, for BD which does not support the use of gabapentinoids in the management of BDI[33]. When drugs were grouped according to their anatomical therapeutic chemical classes the most significant enrichment was in the targets of psycholeptics (specifically antipsychotics and anxiolytics) as has been shown previously[8].

When considering ancestry-specific meta-analyses, we identified 221 and 88 loci associated with BD in EUR samples with the inclusion and exclusion of the self-report data, respectively. In the EAS meta-analysis we identified 2 BD-associated loci, one of which is novel and ancestry-specific (rs117130410, 4:105734758, build GRCh37). This lead SNP is present in various Asian populations with a minor allele frequency of approximately 5%, but is monomorphic in non-Asian populations. The nearest protein coding genes to this lead SNP included *TET2* and *CXXC4*, approximately 300 kb up- and downstream respectively. *TET2* is present as a single isoform in mammalian cells and its putative CXXC domain is encoded by the neighbouring gene *CXXC4*, and this gene is shown to affect neurogenesis in mice[34,35].

The multi-ancestry PRS shows the greatest improvement over the EUR PRS in an out-of-sample African American target cohort, achieving an over two fold greater increase in variance explained. More subtle improvements were seen when EAS and EUR target cohorts were analyzed. These results are in line with previous studies[36], and highlight the benefits of multi-ancestry representation in the PRS training data. Although these results represent a substantial improvement when compared to our previous findings[8] this BD PRS still lacks clinical utility at present[37].

In line with the PRS results, when investigating the similarity of genetic architecture for BD across ancestral groups using an empirical Bayes approach, we identified high genetic overlap between EUR and EAS and LAT samples ($\kappa 1$=0.90), but considerably lower overlap between EUR and AFR samples ($\kappa 1$=0.02 without self-report data and $\kappa 1$=0.45 with self-report data). Although differential linkage disequilibrium patterns between populations in the target and base GWAS may lead to lower estimates of genetic overlap ($\kappa 1$)[14], since overlap is estimated on the basis of the consistency in the occurrence of putative trait-associated loci in the target and base populations, increasing the sample size of the target GWAS should improve the accuracy of the overlap parameter by bringing more true trait-associated loci into a detectable range[14]. Indeed, the inclusion of self-report data greatly increased the effective sample size in the AFR target

GWAS (from N_{eff} = 8,486 to N_{eff} = 23,484) and resulted in a corresponding increase in genetic overlap (*κ1*).

We investigated the genetic architecture of BD subtypes and ascertainment in the EUR samples. BD within clinical and biobank samples was highly but imperfectly correlated (r_g = 0.95) and displayed varying correlations with self-report BD (clinical r_g = 0.47, biobank r_g = 0.79). Similarly, BDI and BDII subtypes are also highly but imperfectly correlated (r_g = 0.88), and vary in their relationship with the self-report BD (BDI r_g = 0.42, BDII r_g = 0.76). In line with these observations, the pattern of correlations between BD and other psychiatric disorders differed with the inclusion of self-report data in the meta-analysis. Schizophrenia was mostly highly correlated with BD without the inclusion of the self-report data in the BD meta-analysis (BD no self report r_g = 0.65, BD with self-report r_g = 0.54), while major depressive disorder was most strongly correlated with BD after the inclusion of the self-report data (BD no self report r_g = 0.52, BD with self-report r_g = 0.77). Moreover, trivariate MiXeR analyses revealed extensive genetic overlap between BD ascertainment types. The correlation of variant effects within the shared components were high between all ascertainment types (r_{g_shared} range 0.73-0.98). In addition, when considering BD subtypes, the correlation of variant effects within the shared components were high for BDII and both biobank and self-report samples (r_{g_shared} range 0.97-1.00), while BDI had a stronger median correlation of shared effects with BD in biobank samples (r_{g_shared} = 1.00) than with self-report BD (r_{g_shared} = 0.52). These results suggest that BD patients included in this self-report data likely reflect a non-hospitalised, non-psychotic part of the BD spectrum. The utilization of the different ascertainment types as described above allowed us to achieve large sample sizes for genetic discovery, especially in under-represented non-European ancestry cohorts. Still, this approach has some limitations, such as lower heritability and reduced specificity of associated loci to the trait of interest[38,39].

In conclusion, in this first large-scale multi-ancestry GWAS of the BD spectrum, we identified 298 significant BD-associated loci, from which we demonstrate convergence of common variant associations with rare variant signals and underscore 47 credible genes implicated in the pathobiology. Our findings highlight that increasing ancestral diversity in genetic studies of BD is crucial to improve discovery and to ensure equitable benefit from genetic discoveries across ancestry groups.

Methods

Sample description

Details of each of the cohorts, including sample size, ancestry, inclusion/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) Clinical, 2) biobank (real-world

diagnoses) and 3) self-report. The clinical samples included 51 cohorts of European ancestry (27,196 cases and 43,792 controls), 46 of which were included in previous PGC-BD GWAS publications[6–8], and one cohort of African American ancestry (347 cases, 669 controls). The biobank samples included 17 cohorts of European ancestry (32,091 cases and 737,230 controls), three cohorts of East Asian ancestry (4,479 cases and 75,725 controls), two cohorts of African American ancestry (2,803 cases and 7,204 controls), and one cohort of Latino ancestry (1032 cases and 3090 controls). Eleven of the European ancestry biobank cohorts were included in the previous PGC-BD GWAS[8].

Individual-level genotype and phenotype data were shared with the PGC for 53 'internal' cohorts (47 clinical cohorts and 6 biobank cohorts), while the remaining 22 'external' cohorts (9 clinical cohorts and 13 biobank cohorts) contributed summary statistics data.

Finally, we included summary statistics data from four self-report cohorts of European (72,682 cases and 1,541,394 controls), African American (3,926 cases and 83,477 controls), East Asian (1,490 cases and 70,186 controls) and Latino ancestry (11,990 cases and 233,732 controls) from 23andMe, Inc, in which individuals were classified as cases if they self-reported having received a clinical diagnosis or treatment for bipolar disorder in responses to web-based surveys ("Have you ever been diagnosed with, or treated for, bipolar disorder?"). The final multi-ancestry meta-analysis included up to 156,643 cases and 2,799,462 controls (N_{eff} = 535,738) with 82.3% of participants (proportion of effective sample size) of European ancestry, 4.4% of African American ancestry, 4.2% of East Asian ancestry and 9.1% of Latino ancestry.

Bipolar I disorder (BDI) and bipolar II disorder (BDII) subtypes were defined based on international consensus criteria (DSM-IV, DSM-V, ICD-9 or ICD-10), established using structured diagnostic instruments from assessments by trained interviewers, clinician-administered checklists or medical record review, as in the previous PGC BD GWAS[8]. In the external biobank cohorts, BD subtypes were defined using ICD codes where possible (Supplementary Note). The majority of new cohorts included in this study were external biobank cohorts where subtype definition 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[8] (Supplementary Table 1). Thus, the previous BDI and BDII GWAS summary statistics data were used for BDI and BDII analyses in this study.

Genotyping, quality control 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[40] 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 ($P > 1 \times 10^{-10}$ in psychiatric 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 (PCs) were

generated using genotyped SNPs in each cohort separately using EIGENSTRAT v6.1.4[41]. Genotype imputation was performed using the prephasing/imputation stepwise approach implemented in Eagle v2.3.5[42] and Minimac3[43] to the Haplotype Reference Consortium (HRC) reference panel v1.0[44]. 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. Full details related to the genotyping, quality control and imputation for these cohorts are provided in the Supplementary Note. Identical individuals between PGC processed cohorts and these external cohorts 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 (GWAS)

For internal PGC cohorts, GWASs were conducted within each cohort using an additive logistic regression model in PLINK v1.90[45], covarying for PCs 1–5 and any others as required, as previously described[8]. 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 European ancestry for which individual level data were available. For external cohorts, GWASs were conducted by the collaborating research teams using comparable procedures (Supplementary Note). To control test statistic inflation at SNPs with low minor allele frequency (MAF) in small cohorts, SNPs were retained only if cohort MAF was >1% and minor allele count was >10 in either cases or controls (whichever had smaller n).

Initially, meta-analysis of GWAS summary statistics was conducted using inverse-variance-weighted fixed-effects models in METAL (version 2011-03-25)[46] across cohorts within ancestral groups (both including and excluding the self-report data). A GWS locus was defined as the region around a SNP with $P < 5.0 \times 10^{-8}$ with linkage disequilibrium (LD) r2 > 0.1, within a 3000 kb window, based on the LD structure of the ancestry matched HRC reference panel v1.0[44], except Latino (EUR panel used). Multi-ancestry meta-analysis (both including and excluding the self-report data) was similarly performed by combining cohorts with diverse ancestry using inverse-variance-weighted fixed-effects models in METAL[46]. Given that >80% of the included participants were of European ancestry, the LD structure of the European HRC reference panel was used to define GWS loci.

For all meta-analyses, SNPs present in <75% of total effective sample size (N_{eff}) were removed from the meta-analysis results. In addition, we employed the DENTIST tool for summary data-based analyses, which leverages LD from a reference sample (ancestry matched HRC reference panel v1.0[44], except Latino 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 neighboring variants[47].

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 East Asian, European and multi-ancestry meta-analysis (both including and excluding the self-report data), with the East Asian and European HRC reference panels, respectively[48,49]. Regional association plots and forest plots of the index SNPs for all GWS loci from all meta-analyses are presented in Supplementary Datafiles 1-12.

Sex-stratified analysis

The sex-stratified analyses were conducted on 52 internal PGC cohorts of European ancestry following the same protocol for the overall genome-wide association study and meta-analysis. Each cohort was split into males and females, using PLINK v1.90 after they had undergone quality control and imputation as above. After splitting, sex-stratified quality control using the same standards was conducted as per the RICOPILI pipeline. Relatedness was calculated across cohorts and sexes and one of each pair of related individuals was excluded. PCs were generated separately on genotyped SNPs for each cohort and sex using EIGENSTRAT v6.1.4. PCs 1-5 and any others as required were included in the additive logistic regression model for each cohort and sex separately, conducted using PLINK v1.90. The sex-stratified GWASs underwent post-GWAS SNP filtering as described above. Meta-analysis for each sex was conducted using inverse-variance-weighted fixed-effects models in METAL (version 2011-03-25)[46] and using the European HRC LD reference panel v1.0. SNPs were then filtered again using N_{eff} and DENTIST as above.

SNPs reaching genome-wide significance level were considered to be sex-specific if they reached genome-wide significance in only one sex and were not in LD ($r2<0.6$) with any SNPs reaching genome-wide significance in the other sex.

Cross-population empirical Bayes method

We implemented the cross-population empirical Bayes method, XPEB[14], to assess whether loci identified in the European meta-analysis also display evidence for association in the African American, East Asian or Latino cohorts. XPEB was designed to boost signal in a target (small, usually non-European) population GWAS when the target and base (large, usually European) GWAS populations share genetic architecture, but not to generate false positives when a signal is only present in the base population GWAS. XPEB takes the p-values from summary statistics of two input GWAS, and reprioritizes variants in the target GWAS to compute local false discovery rate (FDR) values. We ran XPEB with each of the non-European ancestry GWAS meta-analyses as targets and the European ancestry meta-analysis as base (both with and without the self-report data). The output includes a population-wide estimate of the degree of genetic architecture overlap ($\kappa 1$; between 0 and 0.90), and FDR value for each individual SNP

in common between the base and target GWAS. SNPs with FDR < 0.05 were considered statistically significant and indicative of transferable loci.

Gene-Level and Gene-set Association Analysis

Gene-level and gene-set associations were performed using a SNP-wise mean model (±10kb window) implemented in MAGMA[16], based on p-values from the various meta-analyses described above. Bonferroni correction was used to control for multiple testing.
Identified credible genes were further assessed for enrichment in synaptic processes using the SynGO tool (https://www.syngoportal.org/) with default settings [28]. In addition, we performed exploratory gene-set analysis with GSA-MiXeR[17]. This method employs a competitive gene-set enrichment analysis, which applies stochastic gradient-based optimization for maximum likelihood estimation from GWAS summary statistics, to accurately model the contribution of more than 18,000 genes to SNP-heritability of complex polygenic traits, while controlling for polygenicity, MAF- and LD- dependent genetic architecture, and functional categories. This enables the robust estimation of gene-set fold enrichment alongside partitioned heritability.

Fine-mapping

We performed functional fine-mapping of GWS loci with or without the inclusion of the self-report datasets via Polyfun-SuSiE, using functional annotations of the baseline-LF2.2 UKB model [24] and LD estimates from the Haplotype Reference Consortium (HRC) European (N = 21,265) or East Asian reference panels (N = 541). 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 fall within the MHC locus (6:25119106-33854733, build GRCh37) due to the known complexity of the LD architecture in that region. GWS loci ranges with a LD r2 above 0.1 were used as fine-mapping ranges.
European and multi-ancestry datasets (both including and excluding the self-report data) were fine-mapped by using LD estimates from the HRC European reference panel, whilst the HRC East-Asian reference panel was used for fine-mapping the East Asian dataset. Fine-mapping was also performed without the usage of an LD reference panel for the multi-ancestry datasets, both including and excluding self-report data, and assuming one causal variant per fine-mapped region.
Fine-mapped SNPs with a PIP value above 0.5 were mapped to genes by performing variant annotation with Variant Effect Predictor (VEP) (GRCh37) Ensembl release 109[25].

Convergence of common and rare variant signal

Data from the Bipolar Exome (BipEx) consortium[26] were used to assess the convergence of common and rare variant signals, using a similar approach as previously used for schizophrenia[32]. Ultra rare variants (<=5 minor allele count) for damaging missense and

protein-truncating variants were considered. An enrichment of rare variants in Fine-mapped SNP genes in cases relative to controls were assessed using a Fisher's Exact Test. Given the genetic overlap between bipolar disorder and schizophrenia, we repeated the analysis in data from the Schizophrenia Exome Meta-analysis (SCHEMA) cohort[27].

eQTL integrative analysis

We conducted different eQTL integration analyses to identify genes associated with BD via their gene expression. All analyses were performed for the multi-ancestry GWAS summary statistics both with and without self-report data, as well as for summary statistics generated from each ancestry group separately (including European (EUR), African (AFR), East Asian (EAS) and Latin American (LAT) subsets with/without self-report, where possible. The HRC LD reference panel was used for each of the analyses, including individuals of European ancestry for the multi-ancestry GWAS, and the respective subsets of individuals of each superpopulation for the ancestry-specific analyses.

Transcriptome-wide association studies (TWAS) were performed using the FUSION software[21] with precomputed gene expression weights from the PsychENCODE dataset,[19] available online at http://resource.psychencode.org/. This data is derived from 1,321 postmortem human brain samples and comprises 14,750 genes with significant *cis*-SNP heritability. Monomorphic and rare (MAF < 0.01) variants were excluded from the HRC reference panels via filtering with Plink 2.[45] We also ran colocalization[50] tests on any gene-trait associations with a TWAS p-value less than 0.05 (--coloc_P 0.05 flag in FUSION). After excluding the 119 MHC locus genes due to complex LD structure and genes that were skipped by FUSION due to technical reasons such as an insufficient overlap between eQTLs and GWAS SNPs, TWAS p-values were Bonferroni-corrected for the number of genes included in the respective analysis. Since a large number of genes were skipped in the AFR-specific TWAS without self-report data, this analysis was omitted. Gene-trait associations with an adjusted p-value < 0.05 were filtered for COLOC posterior probability (PP4) >= 0.8 to obtain gene-trait associations with a shared causal variant between GWAS and eQTL effect.

We performed fine-mapping of TWAS results using FOCUS[51,52] to model the correlation among the TWAS signals and prioritize the most likely causal gene(s) in each region. For the ancestry-specific subsets, we used the independent LD block definitions provided in the FOCUS software (37:EUR for EUR, 37:EUR for LAT, 37:EAS for EAS, and 37:AFR for AFR). For the multi-ancestry summary statistics, we used the multi-ancestry LD blocks (37:EUR-EAS-AFR). We also applied the MA-FOCUS[51,52] software implementing the joint analysis of each of the ancestry-specific summary statistics, again using the multi-ancestry LD blocks and ancestry-matched LD reference panels from the HRC, though for the gene expression weights we used the same PsychENCODE database (EUR-specific) for all ancestries. We then kept all associations that were in the credible set and had a posterior inclusion probability (PIP) >= 0.8.

We also applied the recently developed isoform-level TWAS method (isoTWAS)[22], using precomputed weights per isoform derived from adult PsychENCODE data, as provided by the isoTWAS developers (https://zenodo.org/record/6795947#.Y8mi2-zMLBI). In total we tested the 7,530 genes, each with varying numbers of isoforms, that had positive heritability with a p-value < 0.05 within the PsychENCODE data. We then performed probabilistic fine-mapping of the significant associations, and filtered the resulting statistics to those transcripts in the credible sets with a permutation p-value less than 0.05 and PIP >= 0.8.

As additional eQTL integration analysis, summary data-based Mendelian randomization (SMR) (v1.3)[23] with subsequent heterogeneity in dependent instruments (HEIDI)[53] tests were performed. The BrainMeta eQTL weights[20] were chosen as primary eQTL dataset instead of the PsychENCODE weights due to the larger sample size of the BrainMeta study (2,865 brain cortex transcriptome samples) and a considerable sample overlap with the PsychENCODE study. Results could not be obtained for the African- and East Asian-specific analyses due to more than 5% of tested genetic variants with allele frequency differences > 0.2 between any pair of the input datasets (GWAS, eQTL, LD reference). Within each analysis, gene-trait associations were considered significant at a Bonferroni-adjusted SMR p-value < 0.05 and a non-significant HEIDI test (p-value >= 0.01).

Credible gene identification

We opted to provide a set of credible causal genes by integrating information from various gene-mapping strategies including, proximity, gene-level associations, GWS loci fine-mapping and integrative eQTL analyses. In other words, we compared the nearest gene to each GWS locus lead SNP, significant gene-level associations, genes harboring potentially causal SNPs with posterior inclusion probability (PIP > 0.50) and within a 95% credible set (CS) and genes prioritized through multiple levels of SMR, TWAS, isoTWAS and FOCUS analyses. The criteria for filtering genes from the different eQTL methods were: (i) TWAS adjusted p-values less than 0.05 and colocalization probability (COLOC.PP4) greater than 0.7, (ii) FOCUS posterior inclusion probability greater than 0.7 and within a credible set, (iii) isoTWAS permutation p-value less than 0.05, isoTWAS poster inclusion probability greater than 0.7 and within a credible set, (iv) SMR adjusted p-value less than 0.05 and HEIDI test p-value greater than 0.01.

Gene expression perturbation profiles of drugs versus credible gene expression

Drug-gene interactions were obtained from the Drug Gene Interaction Database (DGIdb, (https://www.dgidb.org/) v4.2.0[29]. For drugs interacting with one or more of the 47 credible genes, we retrieved drug-induced gene expression data (drug versus no drug) from the Connectivity Map (CMap) 2020[30,31], extracted from the Phase 2 data release of the Library of

Integrated Cellular Signatures (LINCS) (level5_beta_trt_cp_n720216x12328.gctx.gz available at https://clue.io/data/CMap2020#LINCS2020) using the cmapR package[54] in R version 4.3.1. As low drug concentrations in CMap have been shown to reduce the quality of the data[55], we selected the highest concentration per drug.

To evaluate if the drugs could change bipolar disorder-associated gene expression perturbations (whether they down-regulate genes up-regulated in bipolar disorder or vice versa), we calculated the Spearman correlation between the drug-induced perturbations and the bipolar disorder-associated perturbations (TWAS z-values) in the genes defined as credible. This was done for each drug interacting with one or more credible genes, where negative correlation coefficients indicate that the drug could counteract bipolar disorder-associated effects.

Polygenic risk scoring

We used PRS-CS-auto[56] to compute polygenic risk scores for all cohorts from the multi-ancestry GWAS meta-analysis, using a discovery GWAS where the target cohort was left out. Given that the majority of the individuals included in the meta-analysis were of European descent, we used the European LD reference panel based on UK BioBank data as provided by PRS-CS developers (https://github.com/getian107/PRScs). Raw scores were standardized to Z scores, and covariates including sex and PCs 1–5 and any others as required (as above for each cohort GWAS) were included in the logistic regression model, via the glm() function in R, with family=binomial and link=logit. The variance explained by PRS (R2) was first converted to Nagelkerke's pseudo-R2 via the fmsb package in R, then converted to the liability scale to account for proportion of cases in each cohort and the population prevalence of BD.[57] We provide R2 values for BD population prevalences ranging from 1-5%. The weighted average R2 values were then calculated using the N_{eff} for each cohort. The odds ratios for BD for individuals in the top quintile of PRS compared with those in the middle quintile were calculated for all internal PGC cohorts. Similarly, the area under the curve (AUC) statistic was calculated via the pROC package in R, 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 average AUC after all permutations is provided Supplementary Tables 14-18. The mean 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 ancestry PC covariates. In order to assess the gain in AUC due to the PRS itself, we subtracted the mean AUC of the model containing only the covariates from the full model, reported in Supplementary Tables 14-18 as AUC.

Cell type specific enrichment analyses

Single-cell enrichment analyses of brain cell types were performed according to Mullins et al. (2021). Briefly, from five publicly available single-cell RNA sequencing datasets derived from human [58,59] and murine [60–62] brain tissues, the 10% of genes with highest gene expression specificity per celltype were extracted. After MAGMA[16] gene analysis of the multi-ancestry GWAS summary statistics including self-report data using an annotation window of 35 kb upstream and 10 kb downstream of the gene boundaries and the 1000 Genomes phase 3 European reference panel, MAGMA gene-set analyses were conducted for all cell types in each dataset, respectively. Within each dataset, FDR-adjusted p-values below 0.05 were considered statistically significant.
In addition, we performed an exploratory single-cell enrichment analysis in 914 cell types across 29 non-brain murine tissues as implemented in FUMA[18]. Cell types with FDR-adjusted p-values below 0.05 were considered statistically significant.

Drug enrichment analyses

Gene-set analyses were performed restricted to genes targeted by drugs, assessing individual drugs and grouping drugs with similar actions [8,63]. Gene-level and gene-set analyses of the multi-ancestry GWAS summary statistics including self-report data were performed in MAGMA v1.10 [16], 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 v.4.2.0 [29]; the Psychoactive Drug Screening Database Ki DB [64]; ChEMBL v27 [65]; the Target Central Resource Database v6.7.0 [66]; and DSigDB v1.0 [67]; all downloaded in October 2020. Analyses were performed using competitive gene-set analyses in MAGMA. Results from the drug-set analysis were then grouped according to the Anatomical Therapeutic Chemical class of the drug [63]. Only drug classes with at least ten valid drug gene sets within them were analyzed. Drug-class analysis was performed using enrichment curves. All drug gene sets were ranked by their association in the drug-set analysis, and then for a given drug class an enrichment curve was drawn scoring a ‘hit’ if the drug gene set was within the class, or a ‘miss’ if it was outside the class. The area under the curve was calculated, and a P value for this was calculated using the Wilcoxon Mann–Whitney test comparing drug gene sets within the class to drug gene sets outside the class [63]. Multiple testing was controlled using a Bonferroni-corrected significance threshold of $P < 5.41 \times 10^{-5}$ (924 drug-sets with at ten valid drug gene sets) for drug-set analysis and $P < 5.49 \times 10^{-4}$ (91 drug classes with at ten valid drug gene sets) for drug-class analysis, respectively.

LD score regression

LDSC[68] was used to estimate the SNP-heritability (h^2_{SNP}) of BD from EUR GWAS summary statistics, including all cohorts as well as sub-groups by ascertainment and BD subtype. h^2_{SNP}

was converted to the liability scale using a lifetime BD prevalence of 1-2%. LDSC bivariate genetic correlations (r_g) were also estimated between EUR BD GWASs (with and without self-report) and 11 other psychiatric disorders, as well, as 1390 human diseases and traits via the Complex Traits Genetics Virtual Lab (CTG-VL—http://genoma.io) web platform [13]. Adjusting for the number of traits tests, the Bonferroni corrected p-value was p-value < 3.569×10^{-5}.
We calculated a chi-square statistic corresponding to the estimated r_g as $((r_g - 1)/\text{s.e.})^2$ when determining if the LDSC r_g was different from 1.

MiXeR

We applied causal mixture models (MiXeR)[11,12] to investigate the genetic architecture of BD, specifically the overlap between clinical, biobank and self-report samples, as well as BD subtypes. We employed trivariate MiXeR, which first computes univariate MiXeR analyses to estimate the polygenicity, discoverability and heritability of each trait. These were followed by bivariate MiXeR analyses to compute the proportion of shared variants between pairs of traits, and finally trivariate analyses to compute the proportion of shared variants between all three traits analyzed. We also determined the correlation of effect sizes of SNPs within the bivariate shared components.

References

1. McIntyre, R. S. *et al.* Bipolar disorders. *Lancet* **396**, 1841–1856 (2020).
2. Carvalho, A. F., Firth, J. & Vieta, E. Bipolar Disorder. *N. Engl. J. Med.* **383**, 58–66 (2020).
3. Lichtenstein, P. *et al.* Common genetic determinants of schizophrenia and bipolar disorder in Swedish families: a population-based study. *Lancet* **373**, 234–239 (2009).
4. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders 5th edn*. (American Psychiatric Association Publishing, 2013).
5. Merikangas, K. R. *et al.* Lifetime and 12-month prevalence of bipolar spectrum disorder in the National Comorbidity Survey replication. *Arch. Gen. Psychiatry* **64**, 543–552 (2007).
6. 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).

7. Stahl, E. A. *et al.* Genome-wide association study identifies 30 loci associated with bipolar disorder. *Nat. Genet.* **51**, 793–803 (2019).
8. 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).
9. 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).
10. 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).
11. Holland, D. *et al.* Beyond SNP heritability: Polygenicity and discoverability of phenotypes estimated with a univariate Gaussian mixture model. *PLoS Genet.* **16**, e1008612 (2020).
12. Frei, O. *et al.* Bivariate causal mixture model quantifies polygenic overlap between complex traits beyond genetic correlation. *Nat. Commun.* **10**, 2417 (2019).
13. Cuéllar-Partida, G. *et al.* Complex-Traits Genetics Virtual Lab: A community-driven web platform for post-GWAS analyses. *bioRxiv* 518027 (2019) doi:10.1101/518027.
14. Coram, M. A. *et al.* Leveraging Multi-ethnic Evidence for Mapping Complex Traits in Minority Populations: An Empirical Bayes Approach. *Am. J. Hum. Genet.* **96**, 740–752 (2015).
15. Goes, F. S. *et al.* Genome-wide association study of schizophrenia in Ashkenazi Jews. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* **168**, 649–659 (2015).
16. 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).
17. Frei, O. *et al.* Improved functional mapping with GSA-MiXeR implicates biologically specific gene-sets and estimates enrichment magnitude. *bioRxiv* (2022) doi:10.1101/2022.12.08.22283159.

18. 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).
19. Gandal, M. J. *et al.* Transcriptome-wide isoform-level dysregulation in ASD, schizophrenia, and bipolar disorder. *Science* **362**, (2018).
20. Qi, T. *et al.* Genetic control of RNA splicing and its distinct role in complex trait variation. *Nat. Genet.* **54**, 1355–1363 (2022).
21. Gusev, A. *et al.* Integrative approaches for large-scale transcriptome-wide association studies. *Nat. Genet.* **48**, 245–252 (2016).
22. Bhattacharya, A. *et al.* Isoform-level transcriptome-wide association uncovers extensive novel genetic risk mechanisms for neuropsychiatric disorders in the human brain. *medRxiv* 2022.08.23.22279134 (2022) doi:10.1101/2022.08.23.22279134.
23. Zhu, Z. *et al.* Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nat. Genet.* **48**, 481–487 (2016).
24. Weissbrod, O. *et al.* Functionally informed fine-mapping and polygenic localization of complex trait heritability. *Nat. Genet.* **52**, 1355–1363 (2020).
25. McLaren, W. *et al.* The Ensembl Variant Effect Predictor. *Genome Biol.* **17**, 122 (2016).
26. 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).
27. Singh, T. *et al.* Rare coding variants in ten genes confer substantial risk for schizophrenia. *Nature* **604**, 509–516 (2022).
28. Koopmans, F. *et al.* SynGO: An Evidence-Based, Expert-Curated Knowledge Base for the Synapse. *Neuron* **103**, 217–234.e4 (2019).
29. Freshour, S. L. *et al.* Integration of the Drug-Gene Interaction Database (DGIdb 4.0) with open crowdsource efforts. *Nucleic Acids Res.* **49**, D1144–D1151 (2021).
30. Lamb, J. *et al.* The Connectivity Map: using gene-expression signatures to connect small

molecules, genes, and disease. *Science* **313**, 1929–1935 (2006).

31. Subramanian, A. *et al.* A Next Generation Connectivity Map: L1000 Platform and the First 1,000,000 Profiles. *Cell* **171**, 1437–1452.e17 (2017).
32. Trubetskoy, V. *et al.* Mapping genomic loci implicates genes and synaptic biology in schizophrenia. *Nature* **604**, 502–508 (2022).
33. Hong, J. S. W. *et al.* Gabapentin and pregabalin in bipolar disorder, anxiety states, and insomnia: Systematic review, meta-analysis, and rationale. *Mol. Psychiatry* **27**, 1339–1349 (2022).
34. Dick, A. & Chen, A. The role of TET proteins in stress-induced neuroepigenetic and behavioural adaptations. *Neurobiol Stress* **15**, 100352 (2021).
35. Gontier, G. *et al.* Tet2 Rescues Age-Related Regenerative Decline and Enhances Cognitive Function in the Adult Mouse Brain. *Cell Rep.* **22**, 1974–1981 (2018).
36. Duncan, L. *et al.* Analysis of polygenic risk score usage and performance in diverse human populations. *Nat. Commun.* **10**, 3328 (2019).
37. Lewis, C. M. & Vassos, E. Polygenic risk scores: from research tools to clinical instruments. *Genome Med.* **12**, 44 (2020).
38. 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).
39. Cai, N. *et al.* Minimal phenotyping yields genome-wide association signals of low specificity for major depression. *Nat. Genet.* **52**, 437–447 (2020).
40. Lam, M. *et al.* RICOPILI: Rapid Imputation for COnsortias PIpeLIne. *Bioinformatics* **36**, 930–933 (2020).
41. Price, A. L. *et al.* Principal components analysis corrects for stratification in genome-wide association studies. *Nature Genetics* vol. 38 904–909 Preprint at https://doi.org/10.1038/ng1847 (2006).
42. Loh, P.-R. *et al.* Reference-based phasing using the Haplotype Reference Consortium

panel. *Nat. Genet.* **48**, 1443–1448 (2016).

43. Das, S. *et al.* Next-generation genotype imputation service and methods. *Nat. Genet.* **48**, 1284–1287 (2016).
44. McCarthy, S. *et al.* A reference panel of 64,976 haplotypes for genotype imputation. *Nat. Genet.* **48**, 1279–1283 (2016).
45. Chang, C. C. *et al.* Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**, 7 (2015).
46. Tian, B., Deng, Y., Cai, Y., Han, M. & Xu, G. Efficacy and safety of combination therapy with sodium-glucose cotransporter 2 inhibitors and renin-angiotensin system blockers in patients with type 2 diabetes: a systematic review and meta-analysis. *Nephrol. Dial. Transplant* **37**, 720–729 (2022).
47. Chen, W. *et al.* Improved analyses of GWAS summary statistics by reducing data heterogeneity and errors. *Nat. Commun.* **12**, 7117 (2021).
48. 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).
49. Yang, J. *et al.* Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nat. Genet.* **44**, 369–75, S1–3 (2012).
50. Giambartolomei, C. *et al.* Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. *PLoS Genet.* **10**, e1004383 (2014).
51. Mancuso, N. *et al.* Probabilistic fine-mapping of transcriptome-wide association studies. *Nat. Genet.* **51**, 675–682 (2019).
52. Lu, Z. *et al.* Multi-ancestry fine-mapping improves precision to identify causal genes in transcriptome-wide association studies. *Am. J. Hum. Genet.* **109**, 1388–1404 (2022).
53. Zhu, Z. *et al.* Causal associations between risk factors and common diseases inferred from GWAS summary data. *Nat. Commun.* **9**, 224 (2018).

54. Enache, O. M. *et al.* The GCTx format and cmap{Py, R, M, J} packages: resources for optimized storage and integrated traversal of annotated dense matrices. *Bioinformatics* **35**, 1427–1429 (2019).

55. Lim, N. & Pavlidis, P. Evaluation of connectivity map shows limited reproducibility in drug repositioning. *Sci. Rep.* **11**, 17624 (2021).

56. 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).

57. 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).

58. Habib, N. *et al.* Massively parallel single-nucleus RNA-seq with DroNc-seq. *Nat. Methods* **14**, 955–958 (2017).

59. 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).

60. Skene, N. G. *et al.* Genetic identification of brain cell types underlying schizophrenia. *Nat. Genet.* **50**, 825–833 (2018).

61. Molecular Architecture of the Mouse Nervous System. *Cell* **174**, 999–1014.e22 (2018).

62. Website. https://doi.org/10.1016/j.cell.2018.07.028 doi:10.1016/j.cell.2018.07.028.

63. 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).

64. 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).

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

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

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

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