Contribution of common genetic variants to disease status and symptom dimensions in affective and psychotic disorders

Doctoral thesis

to obtain a doctorate (PhD)

from the Faculty of Medicine

of the University of Bonn

Friederike Sophie David

from Goch, Germany

2025

Written with authorization of

the Faculty of Medicine of the University of Bonn

First reviewer: Prof. Dr. med. Andreas J. Forstner

Second reviewer: Prof. Dr. med. Dipl. Phys. Peter M. Krawitz

Day of oral examination: 24.03.2025

From the Institute of Human Genetics

This work is dedicated to anyone struggling with affective or psychotic disorders.

And to my mother, Eva David, for her unshakeable faith in my ability to succeed.

Table of Contents

List of Abbreviations

BD	Bipolar disorder
DSM-5	Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition
GWAS	Genome-wide association study
GWS	Genome-wide significant
h^2_{SNP}	SNP-based heritability
IHME	Institute for Health Metrics and Evaluation
MD	Major depression
MDD	Major depressive disorder
PGC	Psychiatric Genomics Consortium
PRS	Polygenic risk score
RDoC	Research Domain Criteria
SCZ	Schizophrenia
SNP	Single nucleotide polymorphism
SSD	Schizophrenia spectrum disorder
YLDs	Years lived with disability

1 Abstract

Affective and psychotic disorders, such as major depressive disorder (MDD), bipolar disorder (BD), and schizophrenia spectrum disorders (SSD), represent complex psychiatric conditions with a moderate to high heritability. Throughout the last decade, genome-wide association studies (GWAS) have demonstrated the association of many common genetic variants with disease risk. However, the pathophysiological mechanisms of affective and psychotic disorders are still incompletely understood and it is expected that many more disease-associated genetic loci await identification. Moreover, while the different affective and psychotic disorders are considered distinct entities by current diagnostic systems, they exhibit notable phenotypic overlaps and substantial genetic correlations. This suggests that etiological processes may be partially shared between diagnostic groups. Against this backdrop, the three studies included in this thesis were conducted to improve our understanding of the role of common genetic variation in affective and psychotic disorders. In particular, in the first and second study, the contribution of common genetic variants to symptom dimensions of acute and lifetime psychopathology observed across MDD, BD, and SSD was examined. In the third study, the largest GWAS meta-analysis of BD to date was conducted, which revealed novel disease-associated loci and provided insights into the underlying pathobiology via a plethora of GWAS downstream analyses. Altogether, the results of this research expand our knowledge on the complex relationships of common genetic variants with disease status and symptom dimensions within and across affective and psychotic disorders.

2 Introduction

According to the Global Burden of Disease 2021 Study (Institute for Health Metrics and Evaluation (IHME), 2024), mental disorders are among the top ten leading causes of health loss worldwide. Relative to the total number of years lived with disability (YLDs) across all causes in 2021, major depressive disorder (MDD) accounted for 5.1 %, schizophrenia (SCZ) for 1.7 %, and bipolar disorder (BD) for 0.9 % of total YLDs (Institute for Health Metrics and Evaluation (IHME), 2024). The high relevance of mental disorders can be explained, amongst others, by their frequent occurrence, with reported lifetime prevalences of around 14 % for MDD, around 2 % for BD, and around 0.5 % for SCZ (Kessler *et al.*, 2012; McGrath *et al.*, 2008; Merikangas *et al.*, 2011).

MDD, BD, and SCZ represent types of affective and psychotic disorders, for which diagnostic criteria are provided by classification systems such as the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5, American Psychiatric Association, 2013). As affective disorders, MDD and BD are characterized by severe disturbances in mood that cause substantial impairments in an individual's occupational or social functioning. In MDD, individuals experience one or more major depressive episodes, marked by symptoms such as low mood, loss of interest, and changes of appetite or sleep (Marx *et al.*, 2023). The main types of BD are characterized by episodes of mania, which typically alternate with major depressive episodes, (bipolar I disorder) or by hypomania in combination with at least one major depressive episode (bipolar II disorder). During manic episodes, individuals experience an abnormally elevated or irritable mood, with symptoms such as inflated self-esteem, reduced need for sleep, distractibility and excessive activity (Vieta *et al.*, 2018). In SCZ, symptoms of psychosis are prevailing, with delusions, hallucinations, disorganized thinking, disorganized or abnormal motor behavior, and negative symptoms as the five key symptom domains (Kahn *et al.*, 2015).

The etiology of MDD, BD, and SCZ is multifactorial, with a contribution of both environmental and genetic factors. Twin and family studies have estimated a heritability, i.e., proportion of phenotypic variance explained by genetic factors, of around 37 % for MDD (Sullivan *et al.*, 2000), 60 % to 85 % for BD (Johansson *et al.*, 2019; McGuffin *et al.*,

2003), and 73 % to 90 % for SCZ (Sullivan *et al.*, 2003). The underlying genetic architecture of these disorders is complex and highly polygenic. Genetic variants across the allelic spectrum have been implicated (Kendall *et al.*, 2021; Sullivan *et al.*, 2012). Recent large-scale exome sequencing studies have substantiated the role of rare variants in MDD, BD, and SCZ (Palmer *et al.*, 2022; Singh *et al.*, 2022; Tian *et al.*, 2024). An in-depth discussion of rare-variant studies is, however, beyond the scope of this introduction, as the focus of this dissertation lies on common genetic variation.

Common genetic variants, often defined based on a minor allele frequency above 1 % in the general population, and in particular single nucleotide polymorphisms (SNPs), have been shown to play an important role in affective and psychotic disorders by means of genome-wide association studies (GWAS). As the detection of small effect sizes requires large samples to obtain sufficient statistical power, a major breakthrough in understanding the contribution of common variants to disease risk could only be achieved through large-scale GWAS meta-analyses conducted by international consortia such as the Psychiatric Genomics Consortium (PGC, Sullivan, 2010; Sullivan *et al.*, 2018). The most recent GWAS meta-analyses of affective and psychotic disorders led by the PGC, including the GWAS meta-analysis of BD that is part of this dissertation, are listed in Table 1.

***Table 1**: Most recent GWAS meta-analyses of affective and psychotic disorders conducted by the PGC*

Disorder	Reference	# Cases	# Controls	# GWS SNPs
MD	Adams *et al.* (2025)	688,808	4,364,225	697
BD	O'Connell *et al.* (2025)	158,036	2,796,499	337
SCZ	Trubetskoy *et al.* (2022)	76,755	243,649	342

BD, bipolar disorder; GWAS, genome-wide association study; GWS SNPs, independent genome-wide significant (i.e., association $p < 5 \times 10^{-8}$) single nucleotide polymorphisms; MD, major depression; PGC, Psychiatric Genomics Consortium; SCZ, schizophrenia; #, number of.

Most disease-associated common variants are thought to act in an additive manner. This enables the calculation of polygenic risk scores (PRS) based on GWAS findings as weighted sums of risk alleles carried by an individual, reflecting part of the susceptibility to disease (Lewis, Vassos, 2020). PRS are a widely used research tool for studying the association of genetic risk conferred by common variants with various phenotypic and biological measures. A clinical utility of PRS in psychiatry has not been reached yet, owing

to a limited amount of variance explained as well as a low specificity and sensitivity (Andlauer, Nöthen, 2020).

While the individual affective and psychotic disorders are considered distinct entities in current nosology, they exhibit overlaps in clinical presentation and etiological factors. E.g., psychotic features are not only a hallmark of SCZ, but can also occur in severe forms of MDD or BD. In schizoaffective disorder, which is considered a schizophrenia spectrum disorder (SSD), symptoms of SCZ are observed in combination with a mood episode. On the genetic level, several loci with pleiotropic effects on multiple disorders have been identified (Cross-Disorder Group of the Psychiatric Genomics Consortium, 2019), and substantial pairwise genetic correlations between the disorders have been estimated, amounting to 68 % between BD and SCZ, 36 % between MDD and SCZ, and 34 % between MDD and BD (Grotzinger *et al.*, 2022). Moreover, phenotypic heterogeneity exists within diagnostic groups, e.g., there are more than 10,000 unique ways to meet the diagnostic criteria for MDD according to DSM-5 (Cai *et al.*, 2020). The issue of unclear boundaries and a limited alignment of symptom-based diagnostic categories with biomedical research findings has fueled transdiagnostic and dimensional perspectives in psychiatric research (e.g., Insel *et al.*, 2010; Kotov *et al.*, 2017), which are ingrained in the first two studies included in this dissertation.

2.1 Aims

Despite a large body of research investigating different aspects of affective and psychotic disorders, their genetic underpinnings remain incompletely understood. Therefore, the overarching objective of this doctoral thesis was to enhance the understanding of the complex relationships of common genetic variants with disease status and symptom dimensions within and across affective and psychotic disorders. To this end, three studies were conducted. The first and second study included in this dissertation were based on data of the German FOR2107 consortium (Kircher *et al.*, 2019), including individuals with a diagnosis of MDD, BD, or SSD. The focus of these studies was on the contribution of common genetic variants to psychopathological symptom dimensions observed across the diagnostic boundaries of MDD, BD, and SSD, using PRS analyses and exploratory GWAS. In particular, the first study (David *et al.*, 2023) represents a genetic follow-up on

the factor model of acute psychopathology previously published by Stein *et al.* (2020). In the second study (Krug *et al.*, 2024), a new factor model of lifetime psychopathology was described and the brain morphometric and genetic correlates of the three symptom dimensions were examined. The latest GWAS meta-analysis of BD by the PGC (O'Connell *et al.*, 2025), as mentioned above, constitutes the third study of this dissertation. The aim of this study was to identify novel associations of common variants with BD disease status by leveraging a considerable increase in sample size compared to the previous GWAS meta-analysis (Mullins *et al.*, 2021). Using a plethora of GWAS downstream analyses, the study further aimed to enhance and refine our understanding of the genetic architecture and biological underpinnings of BD. The findings of all three studies included in this dissertation were published in peer-reviewed journals and are presented in the following chapter.

2.2 References

Adams MJ, Streit F, Meng X, Awasthi S, Adey BN, … McIntosh AM. Trans-ancestry genome-wide study of depression identifies 697 associations implicating cell types and pharmacotherapies. *Cell* 2025; **188**: 640-652.e9

American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition. Arlington, VA: American Psychiatric Association, 2013

Andlauer TFM, Nöthen M. Polygenic scores for psychiatric disease: from research tool to clinical application. *Med. Genet.* 2020; **32**: 39–45

Cai N, Choi KW, Fried EI. Reviewing the genetics of heterogeneity in depression: operationalizations, manifestations and etiologies. *Hum Mol Genet* 2020; **29**: R10-R18

Cross-Disorder Group of the Psychiatric Genomics Consortium. Genomic Relationships, Novel Loci, and Pleiotropic Mechanisms across Eight Psychiatric Disorders. *Cell* 2019; **179**: 1469-1482.e11

David FS, Stein F, Andlauer TF, Streit F, Witt SH, … Forstner AJ. Genetic contributions to transdiagnostic symptom dimensions in patients with major depressive disorder, bipolar disorder, and schizophrenia spectrum disorders. *Schizophr Res* 2023; **252**: 161–171

Grotzinger AD, Mallard TT, Akingbuwa WA, Ip HF, Adams MJ, … Nivard MG. Genetic architecture of 11 major psychiatric disorders at biobehavioral, functional genomic and molecular genetic levels of analysis. *Nat Genet* 2022; **54:** 548–559

Insel T, Cuthbert B, Garvey M, Heinssen R, Pine DS, … Wang P. Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. *Am J Psychiatry* 2010; **167**: 748–751

Institute for Health Metrics and Evaluation (IHME). Global Burden of Disease 2021: Findings from the GBD 2021 Study. Seattle, WA, 2024

Johansson V, Kuja-Halkola R, Cannon TD, Hultman CM, Hedman AM. A population-based heritability estimate of bipolar disorder - In a Swedish twin sample. *Psychiatry Res* 2019; **278**: 180–187

Kahn RS, Sommer IE, Murray RM, Meyer-Lindenberg A, Weinberger DR, … Insel TR. Schizophrenia. *Nat Rev Dis Primers* 2015; **1**: 15067

Kendall KM, van Assche E, Andlauer TFM, Choi KW, Luykx JJ, … Lu Y. The genetic basis of major depression. *Psychol Med* 2021; **51**: 2217–2230

Kessler RC, Petukhova M, Sampson NA, Zaslavsky AM, Wittchen H-U. Twelve-month and lifetime prevalence and lifetime morbid risk of anxiety and mood disorders in the United States. *Int J Methods Psychiatr Res* 2012; **21**: 169–184

Kircher T, Wöhr M, Nenadic I, Schwarting R, Schratt G, … Dannlowski U. Neurobiology of the major psychoses: a translational perspective on brain structure and function-the FOR2107 consortium. *Eur Arch Psychiatry Clin Neurosci* 2019; **269**: 949–962

Kotov R, Krueger RF, Watson D, Achenbach TM, Althoff RR, … Zimmerman M. The Hierarchical Taxonomy of Psychopathology (HiTOP): A dimensional alternative to traditional nosologies. *J Abnorm Psychol* 2017; **126**: 454–477

Krug A, Stein F, David FS, Schmitt S, Brosch K, … Forstner AJ. Factor analysis of lifetime psychopathology and its brain morphometric and genetic correlates in a transdiagnostic sample. *Transl Psychiatry* 2024; **14**: 235

Lewis CM, Vassos E. Polygenic risk scores: from research tools to clinical instruments. *Genome Med* 2020; **12**: 44

Marx W, Penninx BWJH, Solmi M, Furukawa TA, Firth J, … Berk M. Major depressive disorder. *Nat Rev Dis Primers* 2023; **9**: 44

McGrath J, Saha S, Chant D, Welham J. Schizophrenia: a concise overview of incidence, prevalence, and mortality. *Epidemiol Rev* 2008; **30**: 67–76

McGuffin P, Rijsdijk F, Andrew M, Sham P, Katz R, Cardno A. The heritability of bipolar affective disorder and the genetic relationship to unipolar depression. *Arch Gen Psychiatry* 2003; **60**: 497–502

Merikangas KR, Jin R, He J-P, Kessler RC, Lee S, … Zarkov Z. Prevalence and correlates of bipolar spectrum disorder in the world mental health survey initiative. *Arch Gen Psychiatry* 2011; **68**: 241–251

Mullins N, Forstner AJ, O'Connell KS, Coombes B, Coleman JRI, … Andreassen OA. Genome-wide association study of more than 40,000 bipolar disorder cases provides new insights into the underlying biology. *Nat Genet* 2021; **53**: 817–829

O'Connell KS, Koromina M, van der Veen T, Boltz T, David FS, … Andreassen OA. Genomics yields biological and phenotypic insights into bipolar disorder. *Nature* 2025; **639**: 968–975

Palmer DS, Howrigan DP, Chapman SB, Adolfsson R, Bass N, … Neale BM. Exome sequencing in bipolar disorder identifies AKAP11 as a risk gene shared with schizophrenia. *Nat Genet* 2022; **54**: 541–547

Singh T, Poterba T, Curtis D, Akil H, Al Eissa M, … Daly MJ. Rare coding variants in ten genes confer substantial risk for schizophrenia. *Nature* 2022; **604**: 509–516

Stein F, Lemmer G, Schmitt S, Brosch K, Meller T, … Krug A. Factor analyses of multidimensional symptoms in a large group of patients with major depressive disorder, bipolar disorder, schizoaffective disorder and schizophrenia. *Schizophr Res* 2020; **218**: 38–47

Sullivan PF, Neale MC, Kendler KS. Genetic epidemiology of major depression: review and meta-analysis. *Am J Psychiatry* 2000; **157**: 1552–1562

Sullivan PF, Kendler KS, Neale MC. Schizophrenia as a complex trait: evidence from a meta-analysis of twin studies. *Arch Gen Psychiatry* 2003; **60**: 1187–1192

Sullivan PF. The psychiatric GWAS consortium: big science comes to psychiatry. *Neuron* 2010; **68**: 182–186

Sullivan PF, Daly MJ, O'Donovan M. Genetic architectures of psychiatric disorders: the emerging picture and its implications. *Nat Rev Genet* 2012; **13**: 537–551

Sullivan PF, Agrawal A, Bulik CM, Andreassen OA, Børglum AD, … O'Donovan MC. Psychiatric Genomics: An Update and an Agenda. *Am J Psychiatry* 2018; **175**: 15–27

Tian R, Ge T, Kweon H, Rocha DB, Lam M, … Chen C-Y. Whole-exome sequencing in UK Biobank reveals rare genetic architecture for depression. *Nat Commun* 2024; **15**: 1755

Trubetskoy V, Pardiñas AF, Qi T, Panagiotaropoulou G, Awasthi S, … O'Donovan MC. Mapping genomic loci implicates genes and synaptic biology in schizophrenia. *Nature* 2022; **604**: 502–508

Vieta E, Berk M, Schulze TG, Carvalho AF, Suppes T, … Grande I. Bipolar disorders. *Nat Rev Dis Primers* 2018; **4**: 18008

3 Publications

3.1 Genetic contributions to transdiagnostic symptom dimensions in patients with major depressive disorder, bipolar disorder, and schizophrenia spectrum disorders

This study was authored by David FS, Stein F, Andlauer TF, Streit F, Witt SH, Herms S, Hoffmann P, Heilmann-Heimbach S, Opel N, Repple J, Jansen A, Nenadić I, Papiol S, Heilbronner U, Kalman JL, Schaupp SK, Senner F, Schulte EC, Falkai PG, Schulze TG, Dannlowski U, Kircher T, Rietschel M, Nöthen MM*, Krug A*, and Forstner AJ*.

*shared last authors

In this study, we conducted a genetic follow-up investigation of five previously published symptom dimensions of acute psychopathology present in affective and psychotic disorders. As the first author, I played a major role in the planning of the scientific work and data collection, conducted the data analyses, evaluation, and interpretation, and was responsible for the primary drafting and revision of the manuscript.

The study was published in *Schizophrenia Research* on January 16, 2023:

David FS, Stein F, Andlauer TF, Streit F, Witt SH, … Forstner AJ. Genetic contributions to transdiagnostic symptom dimensions in patients with major depressive disorder, bipolar disorder, and schizophrenia spectrum disorders. *Schizophr Res* 2023; **252**: 161–171

DOI: https://doi.org/10.1016/j.schres.2023.01.002

Schizophrenia Research 252 (2023) 161–171

Contents lists available at ScienceDirect

Schizophrenia Research

journal homepage: www.elsevier.com/locate/schres

Genetic contributions to transdiagnostic symptom dimensions in patients with major depressive disorder, bipolar disorder, and schizophrenia spectrum disorders

Friederike S. David [a], Frederike Stein [b,c], Till F.M. Andlauer [d], Fabian Streit [e], Stephanie H. Witt [e,f], Stefan Herms [a,g], Per Hoffmann [a,g], Stefanie Heilmann-Heimbach [a], Nils Opel [h], Jonathan Repple [h], Andreas Jansen [b,c,i], Igor Nenadić [b,c], Sergi Papiol [j,k], Urs Heilbronner [k], Janos L. Kalman [j,k], Sabrina K. Schaupp [k], Fanny Senner [j,k], Eva C. Schulte [j,k], Peter G. Falkai [j], Thomas G. Schulze [k,l], Udo Dannlowski [h], Tilo Kircher [b,c], Marcella Rietschel [e], Markus M. Nöthen [a,1], Axel Krug [b,m,1], Andreas J. Forstner [a,n,o,*,1]

[a] Institute of Human Genetics, University of Bonn, School of Medicine & University Hospital Bonn, Bonn, Germany
[b] Department of Psychiatry and Psychotherapy, University of Marburg, Marburg, Germany
[c] Center for Mind, Brain and Behavior, University of Marburg, Marburg, Germany
[d] Department of Neurology, Klinikum rechts der Isar, School of Medicine, Technical University of Munich, Munich, Germany
[e] Department of Genetic Epidemiology in Psychiatry, Central Institute of Mental Health, Medical Faculty Mannheim/Heidelberg University, Mannheim, Germany
[f] Center for Innovative Psychiatry and Psychotherapy Research, Central Institute of Mental Health, Medical Faculty Mannheim/Heidelberg University, Mannheim, Germany
[g] Department of Biomedicine, University of Basel, Basel, Switzerland
[h] Institute for Translational Psychiatry, University of Münster, Münster, Germany
[i] Core-Facility Brainimaging, Faculty of Medicine, University of Marburg, Marburg, Germany
[j] Department of Psychiatry and Psychotherapy, University Hospital, LMU Munich, Munich, Germany
[k] Institute of Psychiatric Phenomics and Genomics (IPPG), University Hospital, LMU Munich, Munich, Germany
[l] Department of Psychiatry & Behavioral Sciences, SUNY Upstate Medical University, Syracuse, NY, USA
[m] Department of Psychiatry und Psychotherapy, University Hospital Bonn, Bonn, Germany
[n] Institute of Neuroscience and Medicine (INM-1), Research Center Jülich, Jülich, Germany
[o] Centre for Human Genetics, University of Marburg, Marburg, Germany

ARTICLE INFO

Keywords:
Factor model
Psychosis
Affective disorders
GWAS
Polygenic risk scores

ABSTRACT

Major depressive disorder (MDD), bipolar disorder (BD), and schizophrenia spectrum disorders (SZ) exhibit considerable phenotypic and genetic overlap. However, the contribution of genetic factors to their shared psychopathological symptom dimensions remains unclear. The present exploratory study investigated genetic contributions to the symptom dimensions "Depression", "Negative syndrome", "Positive formal thought disorder", "Paranoid-hallucinatory syndrome", and "Increased appetite" in a transdiagnostic subset of the German FOR2107 cohort (n = 1042 patients with MDD, BD, or SZ). As replication cohort, a subset of the German/Austrian PsyCourse study (n = 816 patients with MDD, BD, or SZ) was employed. First, the relationship between symptom dimensions and common variants associated with MDD, BD, and SZ was investigated via polygenic risk score (PRS) association analyses, with disorder-specific PRS as predictors and symptom dimensions as outcomes. In the FOR2107 study sample, PRS for BD and SZ were positively associated with "Positive formal thought disorder", the PRS for SZ was positively associated with "Paranoid-hallucinatory syndrome", and the PRS for BD was negatively associated with "Depression". The effects of PRS for SZ were replicated in PsyCourse. No significant associations were observed for the MDD PRS. Second, genome-wide association studies (GWAS) were performed for the five symptom dimensions. No genome-wide significant associations and no replicable suggestive associations (p < 1e−6 in the GWAS) were identified. In summary, our results suggest that, similar to

* Corresponding author at: Institute of Human Genetics, University of Bonn, Venusberg-Campus 1, 53127 Bonn, Germany.
E-mail address: forstner@uni-bonn.de (A.J. Forstner).
[1] These authors jointly supervised this work.

https://doi.org/10.1016/j.schres.2023.01.002
Received 23 August 2021; Received in revised form 22 December 2022; Accepted 2 January 2023
Available online 16 January 2023

diagnostic categories, transdiagnostic psychiatric symptom dimensions are attributable to polygenic contributions with small effect sizes. Further studies in larger thoroughly phenotyped psychiatric cohorts are required to elucidate the genetic factors that shape psychopathological symptom dimensions.

1. Introduction

Major depressive disorder (MDD), bipolar disorder (BD), and schizophrenia spectrum disorders (SZ) such as schizophrenia and schizoaffective disorder represent major psychiatric disorders with overlapping clinical phenotypes (Kircher et al., 2019). Although considered distinct diagnostic groups in clinical practice, these disorders have several mood and psychotic features in common. Depressed mood, for example, represents the key symptom in patients diagnosed with MDD, but also occurs in patients diagnosed with BD and SZ (Cotton et al., 2012; Tondo et al., 2017; Upthegrove et al., 2017). Similarly, while psychotic features, such as delusions, typically occur in patients diagnosed with SZ, they are also observed in patients with a diagnosis of BD or MDD (Rosen et al., 2012; Toh et al., 2015; van Bergen et al., 2019; Varghese et al., 2011). Therefore, in recent years, several authors have proposed that the major psychiatric disorders should be conceptualized as a dimensional clinical spectrum, rather than as distinct diagnostic categories (Benazzi, 2005; Keshavan et al., 2011; Lynham et al., 2018). This has prompted transdiagnostic studies to investigate shared phenotypic dimensions and etiological factors.

1.1. Factor model of transdiagnostic symptoms

Factor models represent a valuable approach to the dissection of latent variables in psychiatric disorders (Baek et al., 2019; Emsley et al., 2003; Li et al., 2014). Recently, Stein et al. (2020) conducted exploratory and confirmatory transdiagnostic factor analyses of psychopathological symptoms in individuals with a DSM-IV diagnosis of MDD, BD, or SZ ($n = 1182$). The study sample consisted of a subset of the German FOR2107 cohort, a deeply phenotyped transdiagnostic psychiatric cohort that allows for a plethora of analyses (Kircher et al., 2019). Using a comprehensive set of psychopathological scales for the rating of acute psychopathology (SANS, SAPS, HAMA, HAM-D, YMRS), Stein et al. (2020) identified and validated five transdiagnostic symptom dimensions, i.e., "Depression", "Negative syndrome", "Positive formal thought disorder", "Paranoid-hallucinatory syndrome", and "Increased appetite". While some symptom dimensions exhibited higher average factor scores in specific diagnostic groups, e.g., "Depression" in MDD, distributions of factor scores overlapped between diagnoses. This concordance of the factor model with the clinical observation of heterogeneous phenotypes in the major psychiatric disorders provides a sound basis for the further investigation of shared etiological factors.

1.2. Genetic contributions to major psychiatric disorders

Research has shown that genetic factors play a substantial role in the etiology of the major psychiatric disorders. As demonstrated by twin and family studies, each of the disorders has a strong genetic component, with heritability estimates ranging from 40 % for MDD to 60–80 % for schizophrenia and 60–85 % for BD (Nöthen et al., 2019). For these disorders, genome-wide association studies (GWAS) have identified multiple risk variants that are common in the general population (Mullins et al., 2021; Trubetskoy et al., 2022; Wray et al., 2018). This, in turn, has enabled the calculation of polygenic risk scores (PRS), which summarize disorder-specific genetic risk at the individual level (Andlauer and Nöthen, 2020; Lewis and Vassos, 2020). Furthermore, cross-disorder studies have demonstrated high genetic correlations between the major psychiatric disorders, which supports the hypothesis of shared etiological factors (Cross-Disorder Group of the Psychiatric Genomics Consortium, 2019). In view of the observed genetic and phenotypic overlap between the major psychiatric disorders, the investigation of associations between genetic factors and transdiagnostic measures is becoming an increasingly valued approach (e.g., Andlauer et al., 2021b; Pelin et al., 2021).

1.3. Present study

The aim of the present study was to investigate the association of common genetic variants with the five transdiagnostic symptom dimensions identified by Stein et al. (2020). PRS association analyses were performed to determine the impact of genetic liability for MDD, BD, and SZ. Further, exploratory GWAS were performed for each of the five symptom dimensions. Results were followed up via replication analyses in an independent cohort.

2. Methods

2.1. Sample description and factor scores

All individuals from the German FOR2107 cohort (Kircher et al., 2019) who had been included in the factor analysis study of Stein et al. (2020) were selected for the present analyses. Written informed consent was obtained from all participants during the original FOR2107 recruitment process. Ethical approval was obtained from the local ethics committees of the Universities of Marburg and Münster, Germany. The sample in Stein et al. (2020) comprised 1182 individuals (465 males/717 females) with a DSM-IV diagnosis of MDD, BD, or SZ. For each of these subjects, factor scores for the transdiagnostic symptom dimensions "Depression", "Negative syndrome", "Positive formal thought disorder", "Paranoid-hallucinatory syndrome", and "Increased appetite" were provided by Stein et al. (2020) for the purposes of the present analyses. The individual items contributing to each factor are listed in Supplementary Table 1.

2.2. Genotyping, quality control, and imputation

Genotyping, quality control (QC), and imputation were previously performed as described elsewhere (Meller et al., 2019; Pelin et al., 2021) for the complete FOR2107 cohort, of which the selected sample constitutes a subset. Briefly, genomic DNA was extracted from blood samples and used for genome-wide genotyping with the Infinium PsychArray BeadChip (Illumina, San Diego, CA, USA). After initial QC and population substructure analyses in PLINK v1.90 (Chang et al., 2015), genotype data were imputed to the 1000 Genomes phase 3 reference panel (Auton et al., 2015) using SHAPEIT and IMPUTE2 (Delaneau et al., 2011; Howie et al., 2009). Post-imputation QC included removal of variants with any of the following characteristics: a minor allele frequency < 1 %; a Hardy-Weinberg equilibrium test $p < 1e{-}6$; or an INFO metric < 0.8. From the 1182 subjects for which factor scores were available, 140 individuals were excluded from the present analyses due to: a lack of high-quality genotype data ($n = 124$); a mismatch between information on phenotypic and genotypic sex ($n = 2$); or intra-sample relatedness ($n = 14$) ($\hat{\pi} \geq 0.125$). In the remaining sample ($n = 1042$), henceforth referred to as FOR2107 study sample, imputed genotype probabilities were available for 8565143 common variants.

2.3. Replication sample

For replication analyses, an independent subset of the German/Austrian PsyCourse study (Budde et al., 2019) was selected. Genotyping,

QC, and imputation were previously performed within the complete PsyCourse cohort in the same way as conducted for the FOR2107 cohort. The replication sample comprised $n = 816$ unrelated individuals with a DSM-IV diagnosis of MDD, BD, or SZ. Some of the psychopathology assessment scales applied in the PsyCourse study differed to those used by Stein et al. (2020). However, the respective scales are well correlated (e.g., Rush et al., 2006; van Erp et al., 2014), and mapping symptoms between correlated scales, e.g. SANS/SAPS and PANSS, is common practice in corsortia comprising different study sites, e.g. the ENIGMA consortium. Differing scales included the Positive and Negative Syndrome Scale (PANSS) (Kay et al., 1987) and the Inventory of Depressive Symptoms (IDS-C_{30}) (Rush et al., 1986). To emulate the previously established factor model by Stein et al. (2020), single symptoms included in the original model were mapped to symptoms of the scales used in the replication sample (Supplementary Table 2). For the symptom mapping, equivalent items available in the replication sample were sought for each item of the original factor model. In some cases, a 1:1 mapping of items was not possible, resulting in both 1:n (one-to-many, e.g. HAMA4 to IDS-C 1, 2, 3, and 4) and n:1 (many-to-one) mappings (e.g. both HAMA6 and HAMD1 to IDS-C 5). For few items (HAMA2, HAMD16, SAPS28) that were part of the first four factors in the original model, no corresponding item was identified in the replication sample. Due to lack of data on appetite behavior, the fifth factor "Increased appetite" was not included in the model replication step. Confirmatory factor analysis (CFA) in the replication sample was performed in accordance with Stein et al. (2020) using *MPlus* (version 8.4) (Muthén and Muthén, 1998-2017). As in Stein et al. (2020), the model was estimated using the MLR method. Goodness of fit was assessed using the chi-square significance test, the comparative fit index (CFI) (Bentler, 1990), and the root mean square error of approximation (RMSEA) (Browne and Cudeck, 1993). Latent standardized factor scores for each patient were extracted using *MPlus*.

2.4. Polygenic risk score association analyses

In the complete FOR2107 cohort, disorder-specific PRS for MDD, BD, and SZ were calculated using publicly available GWAS summary statistics (Mullins et al., 2021; Trubetskoy et al., 2022; Wray et al., 2018, meta-analysis excluding 23andMe). For PRS calculation, variant weights were estimated via the PRS-CS approach (Ge et al., 2019), using multiple pre-selected values for the global shrinkage parameter φ (1e−4, 1e−3, 1e−2). PRS were calculated in R (R Core Team, 2019), as described previously (Andlauer et al., 2021a,b). For the FOR2107 study sample, linear regression models were then fitted in R with one of the three z-scaled disorder-specific PRS as predictor and one of the five symptom dimensions as outcome. Due to non-normal distribution of factor scores within the symptom dimensions, rank-based inverse normal transformed values (McCaw et al., 2020) were used. Sex, age, and the first four ancestry components calculated via multidimensional scaling (MDS) were included as covariates (Supplementary Fig. 1). Linear models were first fitted for the complete transdiagnostic FOR2107 study sample, and then separately for each diagnostic subgroup. Adjustment of p values for multiple testing was performed using the Benjamini-Hochberg approach (Benjamini and Hochberg, 1995) for 45 tests (5 symptom dimensions × 3 PRS models × 3 values for φ). Model coefficients were considered statistically significant at $p < 0.05$. The variance explained (R^2) by each PRS was calculated as the difference between R^2 of the full model and R^2 of the null model containing the covariates only. The same analysis was conducted in the PsyCourse study sample for the four approximated symptom dimensions. As the PsyCourse sample was part of the BD GWAS by Mullins et al. (2021), the PRS for BD was calculated in the replication sample using summary statistics of the respective leave-one-out GWAS from Mullins et al. (2021), in which the PsyCourse sample had been excluded (41670 BD cases, 371261 controls).

2.5. Genome-wide association studies and downstream analyses

For each of the five symptom dimensions, a GWAS was performed in the FOR2107 study sample using the linear regression approach in PLINK. Rank-based inverse normal transformed values were used as quantitative phenotypes. Sex, age, and the first four ancestry components calculated via MDS were included as covariates. To facilitate comparisons between symptom dimensions, all variables were z-scaled via the 'standard-beta' modifier. Clumping of genetic markers was performed using a maximum p value of 1e−4 for index variants ('–clump-p1 1e−4'); a linkage disequilibrium threshold of 0.1 ('–clump-r2 0.1'); and a window size of 1000 kb ('–clump-kb 500'). Genetic associations with $p < 5e{-}8$ were considered genome-wide significant, and genetic associations with $p < 1e{-}6$ were considered suggestive of association (e.g., Forstner et al., 2021; Risch and Merikangas, 1996). FUMA (Watanabe et al., 2017) was used for basic annotation of summary statistics, and MAGMA (de Leeuw et al., 2015) as implemented in FUMA was used for gene-based analyses. LocusZoom (Pruim et al., 2010) was used to generate regional plots. Power calculations were performed in accordance with the formulas provided in Visscher et al. (2017).

2.6. Replication analysis

For variants with suggestive evidence of association in the discovery GWAS, association testing was performed in the PsyCourse study sample using PLINK. Rank-based inverse normal transformed factor scores were used as outcomes and sex, age, and the first four ancestry components calculated via MDS as covariates. For the examined variants, a sign test was performed to evaluate concordance of the directions of effect between cohorts, and combined effects and p values were calculated using inverse variance-weighted meta-analysis in METAL (Willer et al., 2010). Since all lead variants of the discovery GWAS with $p < 1e{-}6$ were present in the imputed PsyCourse genotype data, no investigation of other variants at these loci in linkage disequilibrium with the lead variants was performed.

3. Results

The distribution of the FOR2107 study sample ($n = 1042$ individuals post-QC) across diagnostic subgroups is shown in Table 1. The distribution of factor scores and PRS (Fig. 1) demonstrates the transdiagnostic nature of the psychopathological symptom dimensions and the relationships between disorder-specific PRS and diagnostic groups.

Emulation of the factor model of Stein et al. (2020) in the PsyCourse replication sample (Table 2) yielded four factors (i.e., "Depression", "Negative syndrome", "Positive formal thought disorder", "Paranoid-hallucinatory syndrome") comparable to the first four symptom dimensions in Stein et al. (2020). The dimension "Increased appetite" could not be approximated. Results of the CFA of the emulated four-factor model suggested an acceptable fit: $\chi^2 = 939.24$, $df = 187$, $p < 0.0001$, $CFI = 0.868$, $RMSEA = 0.058$.

3.1. PRS association analyses

Associations between disorder-specific PRS for MDD, BD, and SZ and the five transdiagnostic symptom dimensions were investigated in the FOR2107 study sample (Fig. 2A, Supplementary Table 3) to determine whether disorder-specific PRS explain any of the transdiagnostic variance in disorder-related symptom dimensions. A significant positive effect of the PRS for SZ was observed on "Positive formal thought disorder" (maximum $\beta = 0.102$ at $\varphi = 1e{-}3$ with $R^2 = 0.010$ and adjusted $p = 0.0087$), and "Paranoid-hallucinatory syndrome" (maximum $\beta = 0.084$ at $\varphi = 1e{-}3$ with $R^2 = 0.007$ and adjusted $p = 0.0362$). A significant negative effect of the PRS for BD was observed on "Depression" (minimum $\beta = -0.084$ at $\varphi = 1e{-}4$ with $R^2 = 0.007$ and adjusted $p = 0.0362$) as well as a significant positive effect on "Positive formal

Table 1
Characteristics of the FOR2107 study sample.

	Major depressive disorder	Bipolar disorder	Schizophrenia spectrum disorders	Total sample
n	783	134	125	1042
Sex	f = 506, m = 277	f = 70, m = 64	f = 53, m = 72	f = 629, m = 413
Age (SD)	37.1 (13.3)	42.1 (12.5)	38.4 (11.5)	37.9 (13.1)
Age at onset (SD)	26.2 (12.7)	25.4 (12.0)	22.2 (9.1)	25.6 (12.5)
Depressive episodes (SD)	4.1 (6.9)	7.7 (8.1)	6.2 (8.2)	4.6 (7.2)
Manic episodes (SD)	–	5.8 (6.9)	4.7 (9.4)	5.6 (7.4)
Psychotic episodes (SD)	–	–	5.1 (8.2)	5.1 (8.2)
Antidepressants (%)	471 (60.2 %)	52 (38.8 %)	36 (28.8 %)	559 (53.6 %)
Antipsychotics (%)	150 (19.2 %)	64 (47.8 %)	106 (84.8 %)	320 (30.7 %)
Mood stabilizers (%)	36 (4.6 %)	76 (56.7 %)	13 (10.4 %)	125 (12.0 %)
Inpatient treatment (%)	240 (30.7 %)	37 (27.6 %)	58 (46.4 %)	335 (32.2 %)
Remission (%)	308 (39.3 %)	47 (35.1 %)	37 (29.6 %)	392 (37.6 %)

Sample size (n), number of female (f) and male (m) participants, the mean age at recruitment and at onset in years as well as selected clinical characteristics are shown. The latter includes the mean number of depressive, manic, and psychotic episodes where applicable, the number and percentage of participants on different groups of medications, the number and percentage of participants undergoing inpatient treatment, and the number and percentage of participants in remission. SD, standard deviation.

thought disorder" (maximum $\beta = 0.129$ at $\varphi = 1e{-}3$ with $R^2 = 0.017$ and adjusted $p = 0.0009$). For the PRS for MDD, no significant associations were found with any of the symptom dimensions. In PsyCourse, the significant positive effects of the PRS for SZ could be replicated, while the PRS for BD did not have a significant effect on either "Depression" or "Positive formal thought disorder" (Supplementary Fig. 2).

When regression modeling was conducted in the FOR2107 study sample separately for each diagnostic subgroup, heterogeneous effects were observed (Fig. 2B). While the positive effect of PRS for SZ on "Positive formal thought disorder" and "Paranoid-hallucinatory syndrome" reached statistical significance in the total sample, no significant signal was present in any of the diagnostic subgroups. Similarly, with regard to the significant negative effect of the PRS for BD on the "Depression" dimension observed in the total sample, no significant signals were found in the analysis of diagnostic subgroups. Interestingly, PRS for BD showed some evidence for subgroup-specific effects on "Positive formal thought disorder" and "Paranoid-hallucinatory

Table 2
Characteristics of the PsyCourse replication sample.

	Major depressive disorder	Bipolar disorder	Schizophrenia spectrum disorders	Total sample
n	65	353	398	816
Sex	f = 37, m = 28	f = 175, m = 178	f = 166, m = 232	f = 378, m = 438
Age (SD)	41.9 (15.2)	45.9 (13.0)	41.3 (11.9)	43.3 (12.8)
Age at first inpatient treatment (SD)	33.2 (13.3)	33.6 (12.3)	27.5 (9.9)	30.5 (11.6)
Depressive episodes (SD)	4.5 (3.3)	9.0 (10.8)	5.2 (6.2)	7.4 (9.3)
Manic episodes (SD)	–	5.9 (6.8)	3.6 (2.8)	5.5 (6.3)
Antidepressants (%)	52 (80.0 %)	165 (46.7 %)	118 (29.7 %)	335 (41.1 %)
Antipsychotics (%)	30 (46.2 %)	238 (67.4 %)	382 (96.0 %)	650 (79.7 %)
Mood stabilizers (%)	5 (7.7 %)	245 (69.4 %)	50 (12.6 %)	300 (36.8 %)
Inpatient treatment (%)	34 (52.3 %)	102 (28.9 %)	205 (51.5 %)	341 (41.8 %)
Remission (%)	8 (12.3 %)	12 (3.4 %)	7 (1.8 %)	27 (3.3 %)

Sample size (n), number of female (f) and male (m) participants, the mean age at recruitment and at first inpatient treatment in years as well as selected clinical characteristics are shown. The latter includes the mean number of depressive and manic episodes where applicable, the number and percentage of participants on different groups of medications, the number and percentage of participants undergoing inpatient treatment, and the number and percentage of participants in remission. SD, standard deviation.

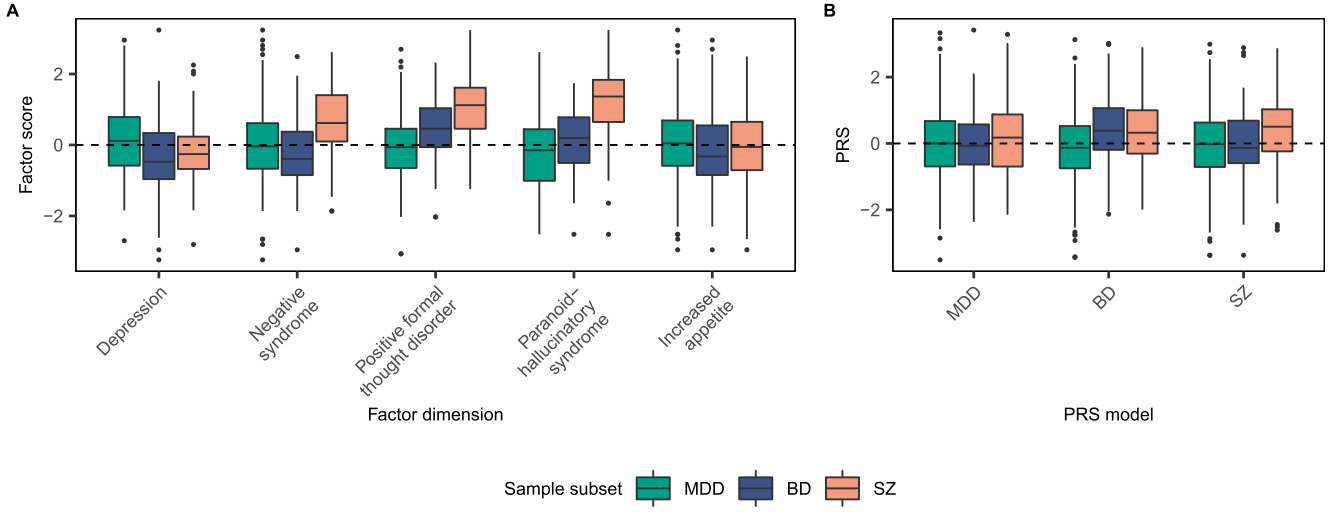


Fig. 1. Distribution of factor scores and PRS across diagnostic groups.
(A) The boxplots of factor scores from Stein et al. (2020) demonstrate overlapping distributions between diagnostic groups. Rank-based inverse normal transformed factor scores used for statistical analyses are shown. (B) Distributions of PRS for MDD, BD, and SZ at $\varphi = 0.001$ after z-scaling (mean = 0, standard deviation = 1) by PRS model. BD, bipolar disorder; MDD, major depressive disorder; PRS, polygenic risk score; SZ, schizophrenia spectrum disorders.

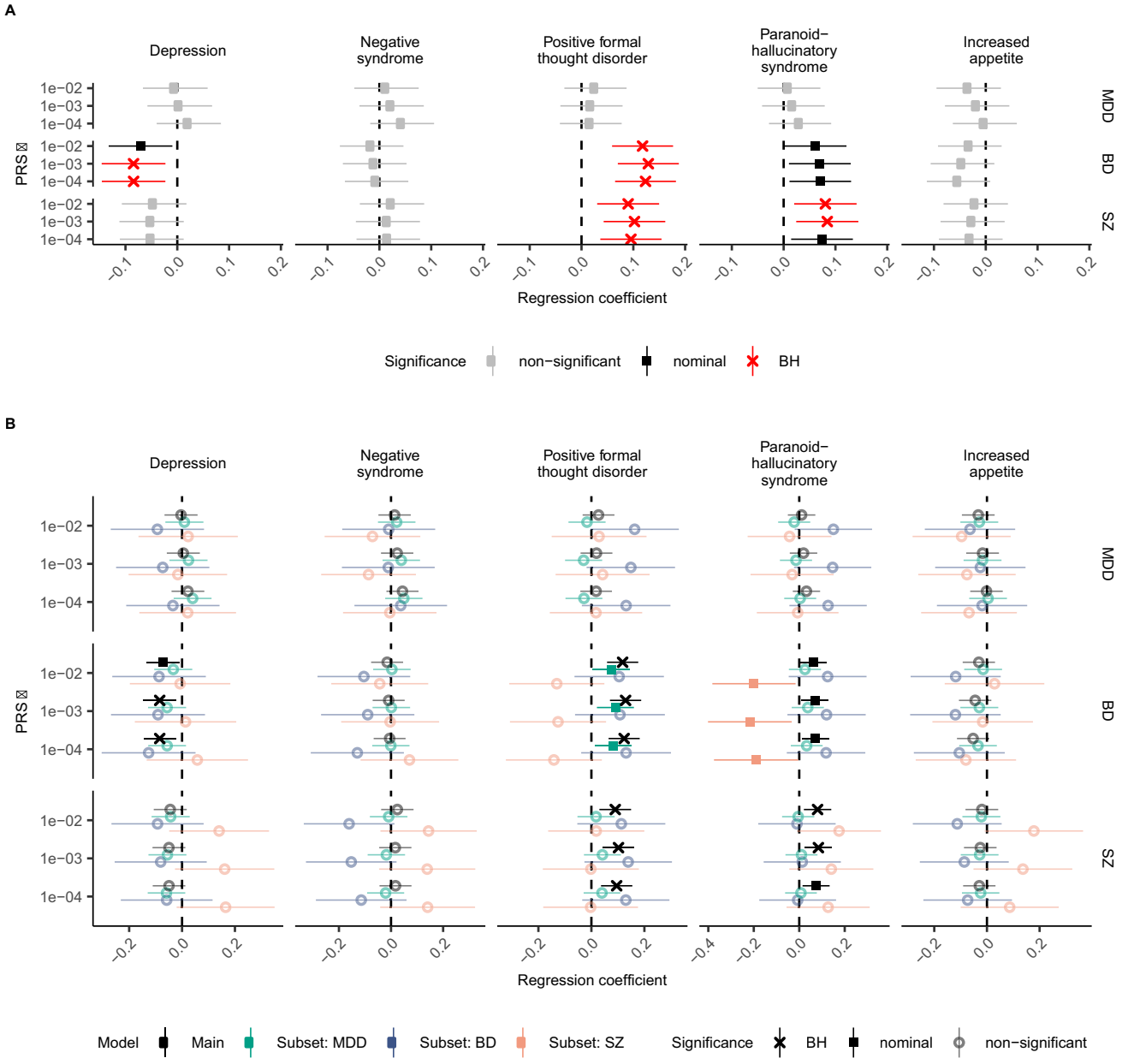


Fig. 2. PRS association analyses.
(A) Regression coefficients of the disorder-specific PRS for MDD, BD, and SZ at different PRS φ values on the transdiagnostic symptom dimensions are shown with 95 % confidence intervals for models based on the complete FOR2107 study sample. Model coefficients with $p < 0.05$ after correction for multiple testing with the Benjamini-Hochberg approach are indicated in red, while nominally significant effects (unadjusted $p < 0.05$) are indicated in black. (B) The regression analysis was repeated within each diagnostic subgroup. Regression coefficients of these subset analyses are shown in comparison to the main model presented in (A). BD, bipolar disorder; BH, Benjamini-Hochberg; MDD, major depressive disorder; PRS, polygenic risk score; SZ, schizophrenia spectrum disorders.

syndrome", albeit non-significant after correction for multiple testing, with a tendency towards negative effects in the subgroup of SZ cases and towards positive effects in the subgroup of BD cases.

3.2. Genome-wide association analyses

In the FOR2107 study sample, GWAS were performed for each of the five symptom dimensions. No evidence of genomic inflation was observed, with λ values between 0.99 and 1.01. No genome-wide significant associations were identified. However, one to three loci per symptom dimension showed suggestive evidence of association ($p <$ 1e−6; Table 3, Supplementary Figs. 3–7). The MAGMA gene analysis and gene-set analysis in FUMA yielded no statistically significant results for any of the symptom dimensions after correction for multiple testing (data not shown). Among the nine loci with association at $p <$ 1e−6, four lead variants were located in introns of protein-coding and noncoding genes and five at intergenic positions. Closer examination of the loci with suggestive evidence of association (Fig. 3, Supplementary Figs. 8–11) revealed that genes at two of the loci have been implicated previously in major psychiatric disorders, i.e., *RELN* and *NEFH*.

Table 3
Lead variants of associated loci ($p < 1e{-}6$) in FOR2107 study sample. Subsequent rows with the same shade represent genetic variants that are associated with the same factor. The respective factor for each group of rows is indicated in the first column.

Factor	CHR	RSID	BP	A1/A2	BETA	SE	P	GENE(S)
Depression	7	rs12536739	103387409	C/T	0.2231	0.0451	9.02E-07	*RELN*
Negative syndrome	1	rs2787876	48948331	A/G	0.5786	0.1091	1.37E-07	*SPATA6, AGBL4*
	4	rs6832060	173985694	A/T	-0.3247	0.0657	8.99E-07	*GALNTL6, GALNT7*
Positive formal thought	1	rs74574738	219286141	T/C	0.9628	0.1828	1.70E-07	*LYPLAL1-AS1*
disorder	11	rs11407840	23543893	G/GA	-0.5118	0.1008	4.55E-07	*MIR8054*
	22	rs420395	29851183	T/C	-0.4079	0.0820	7.72E-07	*RFPL1, NEFH*
Paranoid-hallucinatory	7	rs4726988	148326035	T/C	0.2151	0.0434	8.24E-07	*C7orf33, CUL1*
syndrome	8	rs35831749	69000845	AC/A	-0.2330	0.0469	7.81E-07	*PREX2*
Increased appetite	20	rs193035887	15947060	G/T	-1.0856	0.2142	4.76E-07	*MACROD2*

CHR, chromosome; RSID, reference variant identifier; BP, base pair position (GRCh37); A1, effect allele; A2, other allele; BETA, effect size estimate; SE, standard error; P, association *p* value; GENE(S), nearest gene(s), up- and downstream within 250 kb distance in case of intergenic variants.

3.3. Replication analysis

In the replication analysis of variants with $p < 1e{-}6$ in the discovery GWAS, none of the eight variants showed a nominally significant association ($p < 0.05$) with the respective emulated factor dimension in the PsyCourse replication sample (Supplementary Table 4). For five of the eight evaluated variants, the direction of effect was identical in the FOR2107 study sample and the PsyCourse replication sample. In a binomial sign test, the null hypothesis of random directions of effect could not be rejected ($p = 0.36$).

4. Discussion

Genetic factors play a substantial role in the development of MDD, BD, and SZ (Sullivan and Geschwind, 2019). Although multiple studies have investigated genetic associations within individual diagnostic groups, few studies analyzed genetic associations with transdiagnostic measures. Among others, McCoy et al. (2018) conducted GWAS of five dimensional phenotypes of psychopathology and successfully identified genome-wide significant loci in three dimensions, indicating the merit of this approach. In view of the observed genetic and phenotypic overlap between the major psychiatric disorders, the present study explored genetic contributions to the five transdiagnostic symptom dimensions described by Stein et al. (2020), i.e., "Depression", "Negative syndrome", "Positive formal thought disorder", "Paranoid-hallucinatory syndrome", and "Increased appetite" via PRS analyses and GWAS.

As hypothesized by previous authors (e.g. Bipolar Disorder and Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2018; Guzman-Parra et al., 2021), the manifestation of cross-disorder symptom dimensions may be shaped by certain sets of genetic risk variants. Thus, some associations between genetic factors and symptom dimensions may have already been captured by disorder-specific GWAS conducted by large international consortia (Mullins et al., 2021; Trubetskoy et al., 2022; Wray et al., 2018). For example, the genetic associations identified to date by GWAS of SZ may include variants that contribute to positive formal thought disorder in all major psychiatric disorders, rather than in SZ only. This hypothesis is supported by the positive, although non-significant effect of PRS for SZ on "Positive formal thought disorder" found in the present subgroup analysis of BD cases. However, the opposite direction of effects of the PRS for BD on symptom dimensions "Positive formal thought disorder" and "Paranoid-hallucinatory syndrome" in the subgroups BD and SZ might point to the existence of genetic factors with divergent effects between disorders and therefore distinct biological differences between groups. These results are consistent with previous work by the Bipolar Disorder and Schizophrenia Working Group of the Psychiatric Genomics Consortium (2018), who also described genetic factors with both concordant and divergent effects in BD and SZ. Thus, while transdiagnostic approaches might facilitate the identification of concordant effects due to increase in sample size, their value may be limited if effects differ between groups.

Interestingly, we identified a significant negative effect of PRS for BD on "Depression" in the transdiagnostic FOR2107 sample, which was most pronounced in the subgroup of BD cases. However, as this association was not found in the PsyCourse replication sample, additional studies in independent samples are required before definitive conclusions can be drawn about the effect of the PRS for BD on the "Depression" symptom dimension.

Surprisingly, no significant effect of the PRS for MDD on the symptom dimension "Depression" was observed in either the total transdiagnostic sample or in the MDD subgroup. These findings may reflect the lower heritability of MDD compared to the other major psychiatric disorders and its high degree of etiological heterogeneity (Cai et al., 2020; Nöthen et al., 2019). Therefore, the investigation of endophenotypes in MDD may be valuable to disentangling its genetic etiology (Kendall et al., 2021). The question whether the symptom dimension "Depression" is impacted by the same genetic risk variants in the different diagnostic groups remains to be answered.

In the present GWAS, two of the loci with suggestive evidence of association harbored genes that have been implicated previously in major psychiatric disorders. On chromosome 7, the lead variant rs12536739, suggestively associated with "Depression", mapped to an intron of *RELN*. This gene encodes the Reelin protein, which is involved in cortex development and synaptic function (Ishii et al., 2016; Jossin, 2020). Interestingly, Reelin has been associated with several psychiatric disorders, including MDD, BD, and SZ (Fatemi et al., 2001; Fatemi et al., 2000; Lussier et al., 2011). The second gene, *NEFH*, is located around 25 kb downstream of the lead variant rs420395 on chromosome 22, suggestively associated with "Positive formal thought disorder". *NEFH* encodes the neurofilament heavy polypeptide, for which altered protein levels have been observed in the dorsolateral prefrontal cortex of SZ patients (Pinacho et al., 2016). Notably, none of the present suggestive genetic associations achieved replication in the independent PsyCourse sample. These suggestive associations should thus be viewed with caution until replication is achieved in future studies.

From a clinical perspective, the present analyses generated support for the factor model of Stein et al. (2020) in the independent PsyCourse sample, after symptom mapping to take into account differences in the applied psychopathological assessment scales. These results point to a cross-disorder and cross-scale psychopathological classification into four to five factors. In line with current approaches on dimensional

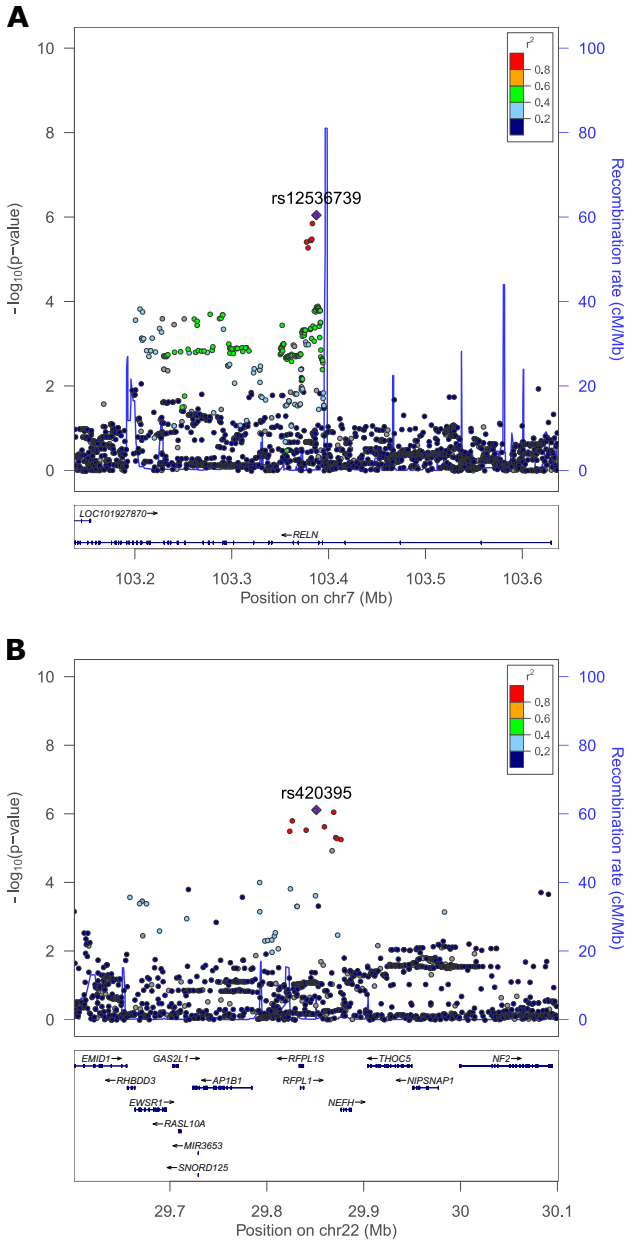


Fig. 3. Regional plots for loci with suggestive evidence of association and a link to major psychiatric disorders.
For the lead variant rs12536739, which displayed suggestive evidence of association with factor dimension "Depression" (A), and lead variant rs420395, which displayed suggestive evidence of association with factor dimension "Positive formal thought disorder" (B), regional plots with a window size of 500 kb are shown. The lead variants represent the linkage disequilibrium reference variants. Below the regional plots, annotation tracks for genes are included. Both *RELN* (A) and *NEFH* (B) have been associated with the major psychiatric disorders in previous research. cM, centimorgan; Mb, megabase.

psychiatry (Anderson et al., 2018; Conway et al., 2019; Kaczkurkin et al., 2018), this type of approach addresses the issue of cross-diagnostic heterogeneity in nosology, and might be a useful framework for studying the shared neurobiology of the major psychiatric disorders.

4.1. Limitations

The present study had four main limitations. First, since the effect sizes of individual risk alleles in complex diseases, such as psychiatric disorders, are generally low (Watanabe et al., 2019), the statistical power to detect genome-wide associations was limited at the given GWAS sample size (Supplementary Fig. 12; Visscher et al., 2017), despite the fact that the FOR2107 cohort represents one of the largest deeply phenotyped transdiagnostic psychiatric cohorts worldwide. In the replication analysis in PsyCourse ($n = 816$), however, the power to detect associations exceeded 90 % at the p value threshold of 0.05 (Supplementary Table 4). Second, due to the unavailability of some of the psychopathology scales used in the original factor model, the symptom dimensions were approximated in the PsyCourse sample via CFA. Although this resulted in a comparable factor model with reasonable fit, an identical phenotype definition would have been preferable for the genetic replication (Kraft et al., 2009). In future studies, efforts should be made to achieve uniform standards in the collection of phenotype data across psychiatric studies in order to overcome this issue. Third, the sample distribution across diagnostic groups differed both within and between the FOR2107 and the PsyCourse sample, with a higher percentage of MDD cases (75 %) in FOR2107 and a lower percentage (8 %) in PsyCourse. Despite the higher heritability of BD and SZ, the power to discover effects driven by BD and SZ cases in the FOR2107 sample may have been reduced compared to MDD. While we assume that the distribution in PsyCourse did not affect the replication of effects driven by BD and SZ cases, since they accounted for the majority of cases in PsyCourse, it is likely that the power to replicate effects driven by MDD cases was reduced. Lastly, the factor model was based on ratings of acute symptoms experienced at the time of data collection. However, factor scores do not remain stable throughout the patient's lifetime. The power of the present analyses may thus have been reduced, depending on the phase of illness that was being experienced by the participants at the time of assessment. Despite this limitation, these data can be considered as an exploratory starting point for the assessment of genetic contributions to transdiagnostic symptom dimensions. To overcome this limitation, future work by the present authors will focus on factor models that are based on lifetime psychopathological symptom dimensions and their genetic underpinnings.

4.2. Conclusion

The present study explored genetic contributions to the five transdiagnostic symptom dimensions reported by Stein et al. (2020). The PRS association analyses generated some degree of evidence for transdiagnostic effects of PRS for SZ on the symptom dimensions "Positive formal thought disorder" and "Paranoid hallucinatory syndrome". In contrast, no effect on the five symptom dimensions was found for PRS for MDD, and the effect of PRS for BD on "Positive formal thought disorder" pointed towards opposite effects in BD and SZ. The GWAS identified no genome-wide significant associations at the given sample size, which suggests that polygenic contributions with small effect sizes are implicated in the individual symptom dimensions. While the small effect sizes in psychiatric genetics studies of common variants limit the immediate clinical utility at this point in time, these studies are nevertheless essential for understanding the biological basis of the major psychiatric disorders. Further studies involving larger deeply phenotyped cohorts of multiple diagnostic subgroups and the evaluation of longitudinally stable measures are required to elucidate the genetic factors that impact the manifestation of different psychopathological symptom dimensions in the major psychiatric disorders.

CRediT authorship contribution statement

FSD, MMN, AK and AJF designed the study. FSD, FStein, TFMA, FStreit, SHW, SH, PH, SH-H, NO, JR, AJ, IN, SP, UH, JLK, SKS, FSenner, ECS, PGF, TGS, UD, TK, MR, MMN, AK, and AJF participated in data acquisition, quality control and preparation. FSD, FStein, and TFMA performed the statistical analyses. FSD, FStein, MMN, AK, and AJF

evaluated and interpreted the results. FSD wrote the first draft of the manuscript; FSD, FStein, TFMA, AK and AJF revised it. All authors contributed to and have approved the final manuscript.

Declaration of competing interest

TK received unrestricted educational grants from Servier, Janssen, Recordati, Aristo, Otsuka, neuraxpharm. ECS received funding through the Munich Clinician Scientist Program (MCSP). All other authors report no conflict of interest.

Acknowledgements

We are deeply indebted to all study participants and staff. A list of acknowledgments can be found here: www.for2107.de/acknowledgements. We thank Christine Schmäl for her critical reading of the manuscript. All PIs take responsibility for the integrity of the respective study data and their components. All authors and coauthors had full access to all study data.

Role of the funding source

This work is part of the German multicenter consortium "Neurobiology of Affective Disorders. A translational perspective on brain structure and function", funded by the German Research Foundation (DFG; Research Unit FOR2107). Principal investigators are: Work Package FOR2107/MACS cohort: TK (speaker FOR2107; DFG grant numbers KI 588/14-1, KI 588/14-2), UD (co-speaker FOR2107; DA 1151/5-1, DA 1151/5-2), AK (KR 3822/5-1, KR 3822/7-2), IN (NE 2254/1-2, NE 2254/3-1, and NE 2254/4-1). Work Package Genetics: MR (RI 908/11-1, RI 908/11-2), MMN (NO 246/10-1, NO 246/10-2), SHW (WI 3439/3-1, WI 3439/3-2). Work Package Data Analytics: Andreas Jansen (JA 1890/7-1, JA 1890/7-2). The study was in part supported by the German Federal Ministry of Education and Research (BMBF) through the Integrated Network IntegraMent (Integrated Understanding of Causes and Mechanisms in Mental Disorders), under the auspices of the e:Med Programme (01ZX1314A/01ZX1614A to MMN, 01ZX1314G/01ZX1614G to MR, 01ZX1614K to TGS) and through ERA-NET NEURON, "SynSchiz—Linking synaptic dysfunction to disease mechanisms in schizophrenia—a multilevel investigation" (01EW1810 to MR). MMN is a member of the DFG-funded Excellence-Cluster ImmunoSensation. TGS received additional support from the BMBF within the framework of the BipoLife network (01EE1404H) and the Dr. Lisa Oehler Foundation, Kassel (Germany). TGS and PGF were supported by the DFG within the framework of the projects www.kfo241.de and www.PsyCourse.de (SCHU 1603/4-1, 5-1, 7-1, FA 241/16-1).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.schres.2023.01.002.

References

Anderson, A.E., Marder, S., Reise, S.P., Savitz, A., Salvadore, G., Fu, D.J., Li, Q., Turkoz, I., Han, C., Bilder, R.M., 2018. Bifactor modeling of the positive and negative syndrome scale: generalized psychosis spans schizoaffective, bipolar, and schizophrenia diagnoses. Schizophr. Bull. 44, 1204–1216. https://doi.org/10.1093/schbul/sbx163.

Andlauer, T.F.M., Guzman-Parra, J., Streit, F., Strohmaier, J., González, M.J., Gil Flores, S., Cabaleiro Fabeiro, F.J., Del Río Noriega, F., Perez, F.P., Haro González, J., Orozco Diaz, G., Diego-Otero, Y.de, Moreno-Küstner, B., Auburger, G., Degenhardt, F., Heilmann-Heimbach, S., Herms, S., Hoffmann, P., Frank, J., Foo, J. C., Treutlein, J., Witt, S.H., Cichon, S., Kogevinas, M., Rivas, F., Mayoral, F., Müller-Myhsok, B., Forstner, A.J., Nöthen, M.M., Rietschel, M., 2021a. Bipolar multiplex families have an increased burden of common risk variants for psychiatric disorders. Mol. Psychiatry 26, 1286–1298. https://doi.org/10.1038/s41380-019-0558-2.

Andlauer, T.F.M., Mühleisen, T.W., Hoffstaedter, F., Teumer, A., Wittfeld, K., Teuber, A., Reinbold, C.S., Grotegerd, D., Bülow, R., Caspers, S., Dannlowski, U., Herms, S., Hoffmann, P., Kircher, T., Minnerup, H., Moebus, S., Nenadić, I., Teismann, H., Völker, U., Etkin, A., Berger, K., Grabe, H.J., Nöthen, M.M., Amunts, K., Eickhoff, S. B., Sämann, P.G., Müller-Myhsok, B., Cichon, S., 2021b. Genetic factors influencing a neurobiological substrate for psychiatric disorders. Transl. Psychiatry 11, 192. https://doi.org/10.1038/s41398-021-01317-7.

Andlauer, T.F.M., Nöthen, M., 2020. Polygenic scores for psychiatric disease: from research tool to clinical application. Med. Genet. 32, 39–45.

Auton, A., Brooks, L.D., Durbin, R.M., Garrison, E.P., Kang, H.M., Korbel, J.O., Marchini, J.L., McCarthy, S., McVean, G.A., Abecasis, G.R., 2015. A global reference for human genetic variation. Nature 526, 68–74. https://doi.org/10.1038/nature15393.

Baek, J.H., Ha, K., Kim, Y., Cho, Y.-A., Yang, S.Y., Choi, Y., Jang, S.-L., Park, T., Ha, T.H., Hong, K.S., 2019. Psychopathologic structure of bipolar disorders: exploring dimensional phenotypes, their relationships, and their associations with bipolar I and II disorders. Psychol. Med. 49, 2177–2185. https://doi.org/10.1017/S003329171800301X.

Benazzi, F., 2005. The relationship of major depressive disorder to bipolar disorder: continuous or discontinuous? Curr. Psychos. Ther. Rep. 3, 131–139. https://doi.org/10.1007/BF02629401.

Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R.Stat. Soc. Series B Stat. Methodol. 57, 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x.

Bentler, P.M., 1990. Comparative fit indexes in structural models. Psychol. Bull. 107, 238–246. https://doi.org/10.1037/0033-2909.107.2.238.

Bipolar Disorder Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2018. Genomic dissection of bipolar disorder and schizophrenia, including 28 subphenotypes. Cell 173, 1705–1715. https://doi.org/10.1016/j.cell.2018.05.046 e16.

Browne, M.W., Cudeck, R., 1993. Alternative ways of assessing model fit. In: Bollen, K.A., Long, J.S. (Eds.), Testing Structural Equation Models. SAGE, Newbury Park, CA, pp. 136–162.

Budde, M., Anderson-Schmidt, H., Gade, K., Reich-Erkelenz, D., Adorjan, K., Kalman, J. L., Senner, F., Papiol, S., Andlauer, T.F.M., Comes, A.L., Schulte, E.C., Klöhn-Saghatolislam, F., Gryaznova, A., Hake, M., Bartholdi, K., Flatau, L., Reitt, M., Quast, S., Stegmaier, S., Meyers, M., Emons, B., Haußleiter, I.S., Juckel, G., Nieratschker, V., Dannlowski, U., Schaupp, S.K., Schmauß, M., Zimmermann, J., Reimer, J., Schulz, S., Wiltfang, J., Reininghaus, E., Anghelescu, I.-G., Arolt, V., Baune, B.T., Konrad, C., Thiel, A., Fallgatter, A.J., Figge, C., von Hagen, M., Koller, M., Lang, F.U., Wigand, M.E., Becker, T., Jäger, M., Dietrich, D.E., Stierl, S., Scherk, H., Spitzer, C., Folkerts, H., Witt, S.H., Degenhardt, F., Forstner, A.J., Rietschel, M., Nöthen, M.M., Falkai, P., Schulze, T.G., Heilbronner, U., 2019. A longitudinal approach to biological psychiatric research: the PsyCourse study. Am. J. Med. Genet. B Neuropsychiatr. Genet. 180, 89–102. https://doi.org/10.1002/ajmg.b.32639.

Cai, N., Choi, K.W., Fried, E.I., 2020. Reviewing the genetics of heterogeneity in depression: operationalizations, manifestations and etiologies. Hum. Mol. Genet. 29, R10–R18. https://doi.org/10.1093/hmg/ddaa115.

Chang, C.C., Chow, C.C., Tellier, L.C., Vattikuti, S., Purcell, S.M., Lee, J.J., 2015. Second-generation PLINK: rising to the challenge of larger and richer datasets. GigaScience 4, 7. https://doi.org/10.1186/s13742-015-0047-8.

Conway, C.C., Forbes, M.K., Forbush, K.T., Fried, E.I., Hallquist, M.N., Kotov, R., Mullins-Sweatt, S.N., Shackman, A.J., Skodol, A.E., South, S.C., Sunderland, M., Waszczuk, M.A., Zald, D.H., Afzali, M.H., Bornovalova, M.A., Carragher, N., Docherty, A.R., Jonas, K.G., Krueger, R.F., Patalay, P., Pincus, A.L., Tackett, J.L., Reininghaus, U., Waldman, I.D., Wright, A.G.C., Zimmermann, J., Bach, B., Bagby, R.M., Chmielewski, M., Cicero, D.C., Clark, L.A., Dalgleish, T., DeYoung, C. G., Hopwood, C.J., Ivanova, M.Y., Latzman, R.D., Patrick, C.J., Ruggero, C.J., Samuel, D.B., Watson, D., Eaton, N.R., 2019. A hierarchical taxonomy of psychopathology can transform mental health research. Perspect. Psychol. Sci. 14, 419–436. https://doi.org/10.1177/1745691618810696.

Cotton, S.M., Lambert, M., Schimmelmann, B.G., Mackinnon, A., Gleeson, J.F.M., Berk, M., Hides, L., Chanen, A., McGorry, P.D., Conus, P., 2012. Depressive symptoms in first episode schizophrenia spectrum disorder. Schizophr. Res. 134, 20–26. https://doi.org/10.1016/j.schres.2011.08.018.

Cross-Disorder Group of the Psychiatric Genomics Consortium, 2019. Genomic relationships, novel loci, and pleiotropic mechanisms across eight psychiatric disorders. Cell 179, 1469–1482.e11. https://doi.org/10.1016/j.cell.2019.11.020.

Delaneau, O., Marchini, J., Zagury, J.-F., 2011. A linear complexity phasing method for thousands of genomes. Nat. Methods 9, 179–181. https://doi.org/10.1038/nmeth.1785.

Emsley, R., Rabinowitz, J., Torreman, M., 2003. The factor structure for the positive and negative syndrome scale (PANSS) in recent-onset psychosis. Schizophr. Res. 61, 47–57. https://doi.org/10.1016/S0920-9964(02)00302-X.

Fatemi, S.H., Earle, J.A., McMenomy, T., 2000. Reduction in reelin immunoreactivity in hippocampus of subjects with schizophrenia, bipolar disorder and major depression. Mol. Psychiatry 5 (654–63), 571. https://doi.org/10.1038/sj.mp.4000783.

Fatemi, S.H., Kroll, J.L., Stary, J.M., 2001. Altered levels of Reelin and its isoforms in schizophrenia and mood disorders. Neuroreport 12, 3209–3215. https://doi.org/10.1097/00001756-200110290-00014.

Forstner, A.J., Awasthi, S., Wolf, C., Maron, E., Erhardt, A., Czamara, D., Eriksson, E., Lavebratt, C., Allgulander, C., Friedrich, N., Becker, J., Hecker, J., Rambau, S., Conrad, R., Geiser, F., McMahon, F.J., Moebus, S., Hess, T., Buerfent, B.C., Hoffmann, P., Herms, S., Heilmann-Heimbach, S., Kockum, I., Olsson, T., Alfredsson, L., Weber, H., Alpers, G.W., Arolt, V., Fehm, L., Fydrich, T., Gerlach, A.L., Hamm, A., Kircher, T., Pané-Farré, C.A., Pauli, P., Rief, W., Ströhle, A., Plag, J., Lang, T., Wittchen, H.-U., Mattheisen, M., Meier, S., Metspalu, A., Domschke, K.,

Reif, A., Hovatta, I., Lindefors, N., Andersson, E., Schalling, M., Mbarek, H., Milaneschi, Y., de Geus, E.J.C., Boomsma, D.I., Penninx, B.W.J.H., Thorgeirsson, T. E., Steinberg, S., Stefansson, K., Stefansson, H., Müller-Myhsok, B., Hansen, T.F., Børglum, A.D., Werge, T., Mortensen, P.B., Nordentoft, M., Hougaard, D.M., Hultman, C.M., Sullivan, P.F., Nöthen, M.M., Woldbye, D.P.D., Mors, O., Binder, E. B., Rück, C., Ripke, S., Deckert, J., Schumacher, J., 2021. Genome-wide association study of panic disorder reveals genetic overlap with neuroticism and depression. Mol. Psychiatry. https://doi.org/10.1038/s41380-019-0590-2.

Ge, T., Chen, C.Y., Ni, Y., Feng, Y.A., Smoller, J.W., 2019. Polygenic prediction via Bayesian regression and continuous shrinkage priors. Nat. Commun. 10 (1), 1776. https://doi.org/10.1038/s41467-019-09718-5.

Guzman-Parra, J., Streit, F., Forstner, A.J., Strohmaier, J., González, M.J., Gil Flores, S., Cabaleiro Fabeiro, F.J., Del Río Noriega, F., Perez Perez, F., Haro González, J., Orozco Diaz, G., de Diego-Otero, Y., Moreno-Kustner, B., Auburger, G., Degenhardt, F., Heilmann-Heimbach, S., Herms, S., Hoffmann, P., Frank, J., Foo, J. C., Sirignano, L., Witt, S.H., Cichon, S., Rivas, F., Mayoral, F., Nöthen, M.M., Andlauer, T.F.M., Rietschel, M., 2021. Clinical and genetic differences between bipolar disorder type 1 and 2 in multiplex families. Transl. Psychiatry 11, 31. https://doi.org/10.1038/s41398-020-01146-0.

Howie, B.N., Donnelly, P., Marchini, J., 2009. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. PLoS Genet. 5, e1000529 https://doi.org/10.1371/journal.pgen.1000529.

Ishii, K., Kubo, K.-I., Nakajima, K., 2016. Reelin and neuropsychiatric disorders. Front. Cell. Neurosci. 10, 229. https://doi.org/10.3389/fncel.2016.00229.

Jossin, Y., 2020. Reelin functions, mechanisms of action and signaling pathways during brain development and maturation. Biomolecules 10, 964. https://doi.org/10.3390/biom10060964.

Kaczkurkin, A.N., Moore, T.M., Calkins, M.E., Ciric, R., Detre, J.A., Elliott, M.A., Foa, E. B., Garcia, La., de Garza, A., Roalf, D.R., Rosen, A., Ruparel, K., Shinohara, R.T., Xia, C.H., Wolf, D.H., Gur, R.E., Gur, R.C., Satterthwaite, T.D., 2018. Common and dissociable regional cerebral blood flow differences associate with dimensions of psychopathology across categorical diagnoses. Mol. Psychiatry 23, 1981–1989. https://doi.org/10.1038/mp.2017.174.

Kay, S.R., Fiszbein, A., Opler, L.A., 1987. The positive and negative syndrome scale (PANSS) for schizophrenia. Schizophr. Bull. 13, 261–276. https://doi.org/10.1093/schbul/13.2.261.

Kendall, K.M., van Assche, E., Andlauer, T.F.M., Choi, K.W., Luykx, J.J., Schulte, E.C., Lu, Y., 2021. The genetic basis of major depression. Psychol. Med. 1–14. https://doi.org/10.1017/S0033291721000441.

Keshavan, M.S., Morris, D.W., Sweeney, J.A., Pearlson, G., Thaker, G., Seidman, L.J., Eack, S.M., Tamminga, C., 2011. A dimensional approach to the psychosis spectrum between bipolar disorder and schizophrenia: the schizo-bipolar scale. Schizophr. Res. 133, 250–254. https://doi.org/10.1016/j.schres.2011.09.005.

Kircher, T., Wöhr, M., Nenadic, I., Schwarting, R., Schratt, G., Alferink, J., Culmsee, C., Garn, H., Hahn, T., Müller-Myhsok, B., Dempfle, A., Hahmann, M., Jansen, A., Pfefferle, P., Renz, H., Rietschel, M., Witt, S.H., Nöthen, M., Krug, A., Dannlowski, U., 2019. Neurobiology of the major psychoses: a translational perspective on brain structure and function-the FOR2107 consortium. Eur. Arch. Psychiatry Clin. Neurosci. 269, 949–962. https://doi.org/10.1007/s00406-018-0943-x.

Kraft, P., Zeggini, E., Ioannidis, J.P.A., 2009. Replication in genome-wide association studies. Stat. Sci. 24, 561–573. https://doi.org/10.1214/09-STS290.

de Leeuw, C.A., Mooij, J.M., Heskes, T., Posthuma, D., 2015. MAGMA: generalized gene-set analysis of GWAS data. PLoS Comput. Biol. 11, e1004219 https://doi.org/10.1371/journal.pcbi.1004219.

Lewis, C.M., Vassos, E., 2020. Polygenic risk scores: from research tools to clinical instruments. Genome Med. 12, 44. https://doi.org/10.1186/s13073-020-00742-5.

Li, Y., Aggen, S., Shi, S., Gao, J., Tao, M., Zhang, K., Wang, X., Gao, C., Yang, L., Liu, Y., Li, K., Shi, J., Wang, G., Liu, L., Zhang, J., Du, B., Jiang, G., Shen, J., Zhang, Z., Liang, W., Sun, J., Hu, J., Liu, T., Miao, G., Meng, H., Hu, C., Huang, G., Li, G., Ha, B., Deng, H., Mei, Q., Zhong, H., Gao, S., Sang, H., Zhang, Y., Fang, X., Yu, F., Yang, D., Chen, Y., Hong, X., Wu, W., Chen, G., Cai, M., Song, Y., Pan, J., Dong, J., Pan, R., Zhang, W., Shen, Z., Liu, Z., Gu, D., Liu, X., Zhang, Q., Flint, J., Kendler, K.S., 2014. The structure of the symptoms of major depression: exploratory and confirmatory factor analysis in depressed Han Chinese women. Psychol. Med. 44, 1391–1401. https://doi.org/10.1017/S003329171300192X.

Lussier, A.L., Romay-Tallón, R., Kalynchuk, L.E., Caruncho, H.J., 2011. Reelin as a putative vulnerability factor for depression: examining the depressogenic effects of repeated corticosterone in heterozygous reeler mice. Neuropharmacology 60, 1064–1074. https://doi.org/10.1016/j.neuropharm.2010.09.007.

Lynham, A.J., Hubbard, L., Tansey, K.E., Hamshere, M.L., Legge, S.E., Owen, M.J., Jones, I.R., Walters, J.T.R., Walters, J.T., 2018. Examining cognition across the bipolar/schizophrenia diagnostic spectrum. J. Psychiatry Neurosci. 43, 245–253. https://doi.org/10.1503/jpn.170076.

McCaw, Z.R., Lane, J.M., Saxena, R., Redline, S., Lin, X., 2020. Operating characteristics of the rank-based inverse normal transformation for quantitative trait analysis in genome-wide association studies. Biometrics 76, 1262–1272. https://doi.org/10.1111/biom.13214.

McCoy, T.H., Castro, V.M., Hart, K.L., Pellegrini, A.M., Yu, S., Cai, T., Perlis, R.H., 2018. Genome-wide association study of dimensional psychopathology using electronic health records. Biol. Psychiatry 83, 1005–1011. https://doi.org/10.1016/j.biopsych.2017.12.004.

Meller, T., Schmitt, S., Stein, F., Brosch, K., Mosebach, J., Yüksel, D., Zaremba, D., Grotegerd, D., Dohm, K., Meinert, S., Förster, K., Redlich, R., Opel, N., Repple, J., Hahn, T., Jansen, A., Andlauer, T.F., Forstner, A.J., Heilmann-Heimbach, S., Streit, F., Witt, S.H., Rietschel, M., Müller-Myhsok, B., Nöthen, M.M., Dannlowski, U., Krug, A., Kircher, T., Nenadić, I., Andlauer, T.F.M., 2019. Associations of schizophrenia risk genes ZNF804A and CACNA1C with schizotypy and modulation of attention in healthy subjects. Schizophr. Res. 208, 67–75. https://doi.org/10.1016/j.schres.2019.04.018.

Mullins, N., Forstner, A.J., O'Connell, K.S., Coombes, B., Coleman, J.R.I., Qiao, Z., Als, T. D., Bigdeli, T.B., Børte, S., Bryois, J., Charney, A.W., Drange, O.K., Gandal, M.J., Hagenaars, S.P., Ikeda, M., Kamitaki, N., Kim, M., Krebs, K., Panagiotaropoulou, G., Schilder, B.M., Sloofman, L.G., Steinberg, S., Trubetskoy, V., Winsvold, B.S., Won, H.-H., Abramova, L., Adorjan, K., Agerbo, E., Al Eissa, M., Albani, D., Alliey-Rodriguez, N., Anjorin, A., Antilla, V., Antoniou, A., Awasthi, S., Baek, J.H., Bækvad-Hansen, M., Bass, N., Bauer, M., Beins, E.C., Bergen, S.E., Birner, A., Bøcker Pedersen, C., Bøen, E., Boks, M.P., Bosch, R., Brum, M., Brumpton, B.M., Brunkhorst-Kanaan, N., Budde, M., Bybjerg-Grauholm, J., Byerley, W., Cairns, M., Casas, M., Cervantes, P., Clarke, T.-K., Cruceanu, C., Cuellar-Barboza, A., Cunningham, J., Curtis, D., Czerski, P.M., Dale, A.M., Dalkner, N., David, F.S., Degenhardt, F., Djurovic, S., Dobbyn, A.L., Douzenis, A., Elvsåshagen, T., Escott-Price, V., Ferrier, I. N., Fiorentino, A., Foroud, T.M., Forty, L., Frank, J., Frei, O., Freimer, N.B., Frisén, L., Gade, K., Garnham, J., Gelernter, J., Giørtz Pedersen, M., Gizer, I.R., Gordon, S.D., Gordon-Smith, K., Greenwood, T.A., Grove, J., Guzman-Parra, J., Ha, K., Haraldsson, M., Hautzinger, M., Heilbronner, U., Hellgren, D., Herms, S., Hoffmann, P., Holmans, P.A., Huckins, L., Jamain, S., Johnson, J.S., Kalman, J.L., Kamatani, Y., Kennedy, J.L., Kittel-Schneider, S., Knowles, J.A., Kogevinas, M., Koromina, M., Kranz, T.M., Kranzler, H.R., Kubo, M., Kupka, R., Kushner, S.A., Lavebratt, C., Lawrence, J., Leber, M., Lee, H.-J., Lee, P.H., Levy, S.E., Lewis, C., Liao, C., Lucae, S., Lundberg, M., MacIntyre, D.J., Magnusson, S.H., Maier, W., Maihofer, A., Malaspina, D., Maratou, E., Martinsson, L., Mattheisen, M., McCarroll, S.A., McGregor, N.W., McGuffin, P., McKay, J.D., Medeiros, H., Medland, S.E., Millischer, V., Montgomery, G.W., Moran, J.L., Morris, D.W., Mühleisen, T.W., O'Brien, N., O'Donovan, C., Olde Loohuis, L.M., Oruc, L., Papiol, S., Pardiñas, A.F., Perry, A., Pfennig, A., Porichi, E., Potash, J.B., Quested, D., Raj, T., Rapaport, M.H., DePaulo, J.R., Regeer, E.J., Rice, J.P., Rivas, F., Rivera, M., Roth, J., Roussos, P., Ruderfer, D.M., Sánchez-Mora, C., Schulte, E.C., Senner, F., Sharp, S., Shilling, P.D., Sigurdsson, E., Sirignano, L., Slaney, C., Smeland, O.B., Smith, D.J., Sobell, J.L., Søholm Hansen, C., Soler Artigas, M., Spijker, A.T., Stein, D. J., Strauss, J.S., Świątkowska, B., Terao, C., Thorgeirsson, T.E., Toma, C., Tooney, P., Tsermpini, E.-E., Vawter, M.P., Vedder, H., Walters, J.T.R., Witt, S.H., Xi, S., Xu, W., Yang, J.M.K., Young, A.H., Young, H., Zandi, P.P., Zhou, H., Zillich, L., Adolfsson, R., Agartz, I., Alda, M., Alfredsson, L., Babadjanova, G., Backlund, L., Baune, B.T., Bellivier, F., Bengesser, S., Berrettini, W.H., Blackwood, D.H.R., Boehnke, M., Børglum, A.D., Breen, G., Carr, V.J., Catts, S., Corvin, A., Craddock, N., Dannlowski, U., Dikeos, D., Esko, T., Etain, B., Ferentinos, P., Frye, M., Fullerton, J. M., Gawlik, M., Gershon, E.S., Goes, F.S., Green, M.J., Grigoroiu-Serbanescu, M., Hauser, J., Henskens, F., Hillert, J., Hong, K.S., Hougaard, D.M., Hultman, C.M., Hveem, K., Iwata, N., Jablensky, A.V., Jones, I., Jones, L.A., Kahn, R.S., Kelsoe, J.R., Kirov, G., Landén, M., Leboyer, M., Lewis, C.M., Li, Q.S., Lissowska, J., Lochner, C., Loughland, C., Martin, N.G., Mathews, C.A., Mayoral, F., McElroy, S.L., McIntosh, A. M., McMahon, F.J., Melle, I., Michie, P., Milani, L., Mitchell, P.B., Morken, G., Mors, O., Mortensen, P.B., Mowry, B., Müller-Myhsok, B., Myers, R.M., Neale, B.M., Nievergelt, C.M., Nordentoft, M., Nöthen, M.M., O'Donovan, M.C., Oedegaard, K.J., Olsson, T., Owen, M.J., Paciga, S.A., Pantelis, C., Pato, C., Pato, M.T., Patrinos, G.P., Perlis, R.H., Posthuma, D., Ramos-Quiroga, J.A., Reif, A., Reininghaus, E.Z., Ribasés, M., Rietschel, M., Ripke, S., Rouleau, G.A., Saito, T., Schall, U., Schalling, M., Schofield, P.R., Schulze, T.G., Scott, L.J., Scott, R.J., Serretti, A., Shannon Weickert, C., Smoller, J.W., Stefansson, H., Stefansson, K., Stordal, E., Streit, F., Sullivan, P.F., Turecki, G., Vaaler, A.E., Vieta, E., Vincent, J.B., Waldman, I.D., Weickert, T.W., Werge, T., Wray, N.R., Zwart, J.-A., Biernacka, J.M., Nurnberger, J.I., Cichon, S., Edenberg, H.J., Stahl, E.A., McQuillin, A., Di Florio, A., Ophoff, R.A., Andreassen, O.A., 2021. Genome-wide association study of more than 40,000 bipolar disorder cases provides new insights into the underlying biology. Nat. Genet. 53, 817–829. https://doi.org/10.1038/s41588-021-00857-4.

Muthén, L.K., Muthén, B.O., 1998-2017. Mplus User's Guide: Eighth Edition. Los Angeles, CA.

Nöthen, M.M., Degenhardt, F., Forstner, A.J., 2019. Durchbrüche im Verständnis der molekularen Ursachen psychiatrischer Störungen [Breakthrough in understanding the molecular causes of psychiatric disorders]. Nervenarzt 90, 99–106. https://doi.org/10.1007/s00115-018-0670-6.

Pelin, H., Ising, M., Stein, F., Meinert, S., Meller, T., Brosch, K., Winter, N.R., Krug, A., Leenings, R., Lemke, H., Nenadić, I., Heilmann-Heimbach, S., Forstner, A.J., Nöthen, M.M., Opel, N., Repple, J., Pfarr, J., Ringwald, K., Schmitt, S., Thiel, K., Waltemate, L., Winter, A., Streit, F., Witt, S., Rietschel, M., Dannlowski, U., Kircher, T., Hahn, T., Müller-Myhsok, B., Andlauer, T.F.M., 2021. Identification of transdiagnostic psychiatric disorder subtypes using unsupervised learning. Neuropsychopharmacology. https://doi.org/10.1038/s41386-021-01051-0.

Pinacho, R., Villalmanzo, N., Meana, J.J., Ferrer, I., Berengueras, A., Haro, J.M., Villén, J., Ramos, B., 2016. Altered CSNK1E, FABP4 and NEFH protein levels in the dorsolateral prefrontal cortex in schizophrenia. Schizophr. Res. 177, 88–97. https://doi.org/10.1016/j.schres.2016.04.050.

Pruim, R.J., Welch, R.P., Sanna, S., Teslovich, T.M., Chines, P.S., Gliedt, T.P., Boehnke, M., Abecasis, G.R., Willer, C.J., 2010. LocusZoom: regional visualization of genome-wide association scan results. Bioinformatics 26, 2336–2337. https://doi.org/10.1093/bioinformatics/btq419.

R Core Team, 2019. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing.

Risch, N., Merikangas, K., 1996. The future of genetic studies of complex human diseases. Science 273, 1516–1517. https://doi.org/10.1126/science.273.5281.1516.

Rosen, C., Marvin, R., Reilly, J.L., Deleon, O., Harris, M.S.H., Keedy, S.K., Solari, H., Weiden, P., Sweeney, J.A., 2012. Phenomenology of first-episode psychosis in schizophrenia, bipolar disorder, and unipolar depression: a comparative analysis. Clin. Schizophr. Relat. Psychoses 6, 145–151. https://doi.org/10.3371/CSRP.6.3.6.

Rush, A.J., Bernstein, I.H., Trivedi, M.H., Carmody, T.J., Wisniewski, S., Mundt, J.C., Shores-Wilson, K., Biggs, M.M., Woo, A., Nierenberg, A.A., Fava, M., 2006. An evaluation of the quick inventory of depressive symptomatology and the Hamilton rating scale for depression: a sequenced treatment alternatives to relieve depression trial report. Biol. Psychiatry 59, 493–501. https://doi.org/10.1016/j.biopsych.2005.08.022.

Rush, J.A., Giles, D.E., Schlesser, M.A., Fulton, C.L., Weissenburger, J., Burns, C., John Rush, A., Giles, D.E., Schlesser, M.A., Fulton, C.L., Weissenburger, J., Burns, C., 1986. The inventory for depressive symptomatology (IDS): preliminary findings. Psychiatry Res. 18, 65–87. https://doi.org/10.1016/0165-1781(86)90060-0.

Stein, F., Lemmer, G., Schmitt, S., Brosch, K., Meller, T., Fischer, E., Kraus, C., Lenhard, L., Köhnlein, B., Murata, H., Bäcker, A., Müller, M., Franz, M., Förster, K., Meinert, S., Enneking, V., Koch, K., Grotegerd, D., Nagels, A., Nenadić, I., Dannlowski, U., Kircher, T., Krug, A., 2020. Factor analyses of multidimensional symptoms in a large group of patients with major depressive disorder, bipolar disorder, schizoaffective disorder and schizophrenia. Schizophr. Res. 218, 38–47. https://doi.org/10.1016/j.schres.2020.03.011.

Sullivan, P.F., Geschwind, D.H., 2019. Defining the genetic, genomic, cellular, and diagnostic architectures of psychiatric disorders. Cell 177, 162–183. https://doi.org/10.1016/j.cell.2019.01.015.

Toh, W.L., Thomas, N., Rossell, S.L., 2015. Auditory verbal hallucinations in bipolar disorder (BD) and major depressive disorder (MDD): a systematic review. J. Affect. Disord. 184, 18–28. https://doi.org/10.1016/j.jad.2015.05.040.

Tondo, L., Vázquez, G.H., Baldessarini, R.J., 2017. Depression and mania in bipolar disorder. Curr. Neuropharmacol. 15, 353–358. https://doi.org/10.2174/1570159X14666160606210811.

Trubetskoy, V., Pardiñas, A.F., Qi, T., Panagiotaropoulou, G., Awasthi, S., Bigdeli, T.B., Bryois, J., Chen, C.-Y., Dennison, C.A., Hall, L.S., Lam, M., Watanabe, K., Frei, O., Ge, T., Harwood, J.C., Koopmans, F., Magnusson, S., Richards, A.L., Sidorenko, J., Wu, Y., Zeng, J., Grove, J., Kim, M., Li, Z., Voloudakis, G., Zhang, W., Adams, M., Agartz, I., Atkinson, E.G., Agerbo, E., Al Eissa, M., Albus, M., Alexander, M., Alizadeh, B.Z., Alptekin, K., Als, T.D., Amin, F., Arolt, V., Arrojo, M., Athanasiu, L., Azevedo, M.H., Bacanu, S.A., Bass, N.J., Begemann, M., Belliveau, R.A., Bene, J., Benyamin, B., Bergen, S.E., Blasi, G., Bobes, J., Bonassi, S., Braun, A., Bressan, R.A., Bromet, E.J., Bruggeman, R., Buckley, P.F., Buckner, R.L., Bybjerg-Grauholm, J., Cahn, W., Cairns, M.J., Calkins, M.E., Carr, V.J., Castle, D., Catts, S.V., Chambert, K. D., Chan, R.C.K., Chaumette, B., Cheng, W., Cheung, E.F.C., Chong, S.A., Cohen, D., Consoli, A., Cordeiro, Q., Costas, J., Curtis, C., Davidson, M., Davis, K.L., Haan, L.de, Degenhardt, F., DeLisi, L.E., Demontis, D., Dickerson, F., Dikeos, D., Dinan, T., Djurovic, S., Duan, J., Ducci, G., Dudbridge, F., Eriksson, J.G., Fañanás, L., Faraone, S.V., Fiorentino, A., Forstner, A., Frank, J., Freimer, N.B., Fromer, M., Frustaci, A., Gadelha, A., Genovese, G., Gershon, E.S., Giannitelli, M., Giegling, I., Giusti-Rodríguez, P., Godard, S., Goldstein, J.I., González Peñas, J., González-Pinto, A., Gopal, S., Gratten, J., Green, M.F., Greenwood, T.A., Guillin, O., Gülöksüz, S., Gur, R.E., Gur, R.C., Gutiérrez, B., Hahn, E., Hakonarson, H., Haroutunian, V., Hartmann, A.M., Harvey, C., Hayward, C., Henskens, F.A., Herms, S., Hoffmann, P., Howrigan, D.P., Ikeda, M., Iyegbe, C., Joa, I., Julià, A., Kähler, A.K., Kam-Thong, T., Kamatani, Y., Karachanak-Yankova, S., Kebir, O., Keller, M.C., Kelly, B.J., Khrunin, A., Kim, S.-W., Klovins, J., Kondratiev, N., Konte, B., Kraft, J., Kubo, M., Kučinskas, V., Kučinskiene, Z.A., Kusumawardhani, A., Kuzelova-Ptackova, H., Landi, S., Lazzeroni, L.C., Lee, P.H., Legge, S.E., Lehrer, D.S., Lencer, R., Lerer, B., Li, M., Lieberman, J., Light, G.A., Limborska, S., Liu, C.-M., Lönnqvist, J., Loughland, C.M., Lubinski, J., Luykx, J.J., Lynham, A., Macek, M., Mackinnon, A., Magnusson, P.K.E., Maher, B.S., Maier, W., Malaspina, D., Mallet, J., Marder, S.R., Marsal, S., Martin, A.R., Martorell, L., Mattheisen, M., McCarley, R.W., McDonald, C., McGrath, J.J., Medeiros, H., Meier, S., Melegh, B., Melle, I., Mesholam-Gately, R.I., Metspalu, A., Michie, P.T., Milani, L., Milanova, V., Mitjans, M., Molden, E., Molina, E., Molto, M.D., Mondelli, V., Moreno, C., Morley, C.P., Muntané, G., Murphy, K.C., Myin-Germeys, I., Nenadić, I., Nestadt, G., Nikitina-Zake, L., Noto, C., Nuechterlein, K.H., O'Brien, N.L., O'Neill, F.A., Oh, S.-Y., Olincy, A., Ota, V.K., Pantelis, C., Papadimitriou, G.N., Parellada, M., Paunio, T., Pellegrino, R., Periyasamy, S., Perkins, D.O., Pfuhlmann, B., Pietiläinen, O., Pimm, J., Porteous, D., Powell, J., Quattrone, D., Quested, D., Radant, A.D., Rampino, A., Rapaport, M.H., Rautanen, A., Reichenberg, A., Roe, C., Roffman, J.L., Roth, J., Rothermundt, M., Rutten, B.P.F., Saker-Delye, S., Salomaa, V., Sanjuan, J., Santoro, M.L., Savitz, A., Schall, U., Scott, R.J., Seidman, L.J., Sharp, S.I., Shi, J., Siever, L.J., Sigurdsson, E., Sim, K., Skarabis, N., Slominsky, P., So, H.-C., Sobell, J. L., Söderman, E., Stain, H.J., Steen, N.E., Steixner-Kumar, A.A., Stögmann, E., Stone, W.S., Straub, R.E., Streit, F., Strengman, E., Stroup, T.S., Subramaniam, M., Sugar, C.A., Suvisaari, J., Svrakic, D.M., Swerdlow, N.R., Szatkiewicz, J.P., Ta, T.M. T., Takahashi, A., Terao, C., Thibaut, F., Toncheva, D., Tooney, P.A., Torretta, S., Tosato, S., Tura, G.B., Turetsky, B.I., Üçok, A., Vaaler, A., van Amelsvoort, T., van Winkel, R., Veijola, J., Waddington, J., Walter, H., Waterreus, A., Webb, B.T., Weiser, M., Williams, N.M., Witt, S.H., Wormley, B.K., Wu, J.Q., Xu, Z., Yolken, R., Zai, C.C., Zhou, W., Zhu, F., Zimprich, F., Atbaşoğlu, E.C., Ayub, M., Benner, C., Bertolino, A., Black, D.W., Bray, N.J., Breen, G., Buccola, N.G., Byerley, W.F., Chen, W.J., Cloninger, C.R., Crespo-Facorro, B., Donohoe, G., Freedman, R., Galletly, C., Gandal, M.J., Gennarelli, M., Hougaard, D.M., Hwu, H.-G., Jablensky, A. V., McCarroll, S.A., Moran, J.L., Mors, O., Mortensen, P.B., Müller-Myhsok, B., Neil, A.L., Nordentoft, M., Pato, M.T., Petryshen, T.L., Pirinen, M., Pulver, A.E., Schulze, T.G., Silverman, J.M., Smoller, J.W., Stahl, E.A., Tsuang, D.W., Vilella, E., Wang, S.-H., Xu, S., Adolfsson, R., Arango, C., Baune, B.T., Belangero, S.I., Børglum, A.D., Braff, D., Bramon, E., Buxbaum, J.D., Campion, D., Cervilla, J.A., Cichon, S., Collier, D.A., Corvin, A., Curtis, D., Di Forti, M., Domenici, E., Ehrenreich, H., Escott-Price, V., Esko, T., Fanous, A.H., Gareeva, A., Gawlik, M., Gejman, P.V., Gill, M., Glatt, S.J., Golimbet, V., Hong, K.S., Hultman, C.M., Hyman, S.E., Iwata, N., Jönsson, E.G., Kahn, R.S., Kennedy, J.L., Khusnutdinova, E., Kirov, G., Knowles, J.A., Krebs, M.-O., Laurent-Levinson, C., Lee, J., Lencz, T., Levinson, D.F., Li, Q.S., Liu, J., Malhotra, A.K., Malhotra, D., McIntosh, A., McQuillin, A., Menezes, P.R., Morgan, V.A., Morris, D.W., Mowry, B.J., Murray, R. M., Nimgaonkar, V., Nöthen, M.M., Ophoff, R.A., Paciga, S.A., Palotie, A., Pato, C.N., Qin, S., Rietschel, M., Riley, B.P., Rivera, M., Rujescu, D., Saka, M.C., Sanders, A.R., Schwab, S.G., Serretti, A., Sham, P.C., Shi, Y., St Clair, D., Stefánsson, H., Stefansson, K., Tsuang, M.T., van Os, J., Vawter, M.P., Weinberger, D.R., Werge, T., Wildenauer, D.B., Yu, X., Yue, W., Holmans, P.A., Pocklington, A.J., Roussos, P., Vassos, E., Verhage, M., Visscher, P.M., Yang, J., Posthuma, D., Andreassen, O.A., Kendler, K.S., Owen, M.J., Wray, N.R., Daly, M.J., Huang, H., Neale, B.M., Sullivan, P.F., Ripke, S., Walters, J.T.R., O'Donovan, M.C., 2022. Mapping genomic loci implicates genes and synaptic biology in schizophrenia. Nature 604, 502–508. https://doi.org/10.1038/s41586-022-04434-5.

Upthegrove, R., Marwaha, S., Birchwood, M., 2017. Depression and schizophrenia: cause, consequence, or trans-diagnostic issue? Schizophr. Bull. 43, 240–244. https://doi.org/10.1093/schbul/sbw097.

van Bergen, A.H., Verkooijen, S., Vreeker, A., Abramovic, L., Hillegers, M.H., Spijker, A. T., Hoencamp, E., Regeer, E.J., Knapen, S.E., Riemersma-van der Lek, R.F., Schoevers, R., Stevens, A.W., Schulte, P.F.J., Vonk, R., Hoekstra, R., van Beveren, N. J., Kupka, R.W., Sommer, I.E.C., Ophoff, R.A., Kahn, R.S., Boks, M.P.M., 2019. The characteristics of psychotic features in bipolar disorder. Psychol. Med. 49, 2036–2048. https://doi.org/10.1017/S0033291718002854.

van Erp, T.G.M., Preda, A., Nguyen, D., Faziola, L., Turner, J., Bustillo, J., Belger, A., Lim, K.O., McEwen, S., Voyvodic, J., Mathalon, D.H., Ford, J., Potkin, S.G., Fbirn, 2014. Converting positive and negative symptom scores between PANSS and SAPS/SANS. Schizophr. Res. 152, 289–294. https://doi.org/10.1016/j.schres.2013.11.013.

Varghese, D., Scott, J., Welham, J., Bor, W., Najman, J., O'Callaghan, M., Williams, G., McGrath, J., 2011. Psychotic-like experiences in major depression and anxiety disorders: a population-based survey in young adults. Schizophr. Bull. 37, 389–393. https://doi.org/10.1093/schbul/sbp083.

Visscher, P.M., Wray, N.R., Zhang, Q., Sklar, P., McCarthy, M.I., Brown, M.A., Yang, J., 2017. 10 years of GWAS discovery: biology, function, and translation. Am. J. Hum. Genet. 101, 5–22. https://doi.org/10.1016/j.ajhg.2017.06.005.

Watanabe, K., Stringer, S., Frei, O., Umićević Mirkov, M., de Leeuw, C., Polderman, T.J. C., van der Sluis, S., Andreassen, O.A., Neale, B.M., Posthuma, D., 2019. A global overview of pleiotropy and genetic architecture in complex traits. Nat. Genet. 51, 1339–1348. https://doi.org/10.1038/s41588-019-0481-0.

Watanabe, K., Taskesen, E., van Bochoven, A., Posthuma, D., 2017. Functional mapping and annotation of genetic associations with FUMA. Nat. Commun. 8, 1826. https://doi.org/10.1038/s41467-017-01261-5.

Willer, C.J., Li, Y., Abecasis, G.R., 2010. METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics 26, 2190–2191. https://doi.org/10.1093/bioinformatics/btq340.

Wray, N.R., Ripke, S., Mattheisen, M., Trzaskowski, M., Byrne, E.M., Abdellaoui, A., Adams, M.J., Agerbo, E., Air, T.M., Andlauer, T.M.F., Bacanu, S.-A., Bækvad-Hansen, M., Beekman, A.F.T., Bigdeli, T.B., Binder, E.B., Blackwood, D.R.H., Bryois, J., Buttenschøn, H.N., Bybjerg-Grauholm, J., Cai, N., Castelao, E., Christensen, J.H., Clarke, T.-K., Coleman, J.I.R., Colodro-Conde, L., Couvy-Duchesne, B., Craddock, N., Crawford, G.E., Crowley, C.A., Dashti, H.S., Davies, G., Deary, I.J., Degenhardt, F., Derks, E.M., Direk, N., Dolan, C.V., Dunn, E.C., Eley, T.C., Eriksson, N., Escott-Price, V., Kiadeh, F.H.F., Finucane, H.K., Forstner, A.J., Frank, J., Gaspar, H.A., Gill, M., Giusti-Rodríguez, P., Goes, F.S., Gordon, S.D., Grove, J., Hall, L.S., Hannon, E., Hansen, C.S., Hansen, T.F., Herms, S., Hickie, I.B., Hoffmann, P., Homuth, G., Horn, C., Hottenga, J.-J., Hougaard, D.M., Hu, M., Hyde, C.L., Ising, M., Jansen, R., Jin, F., Jorgenson, E., Knowles, J.A., Kohane, I.S., Kraft, J., Kretzschmar, W.W., Krogh, J., Kutalik, Z., Lane, J.M., Li, Y., Li, Y., Lind, P. A., Liu, X., Lu, L., MacIntyre, D.J., MacKinnon, D.F., Maier, R.M., Maier, W., Marchini, J., Mbarek, H., McGrath, P., McGuffin, P., Medland, S.E., Mehta, D., Middeldorp, C.M., Mihailov, E., Milaneschi, Y., Milani, L., Mill, J., Mondimore, F.M., Montgomery, G.W., Mostafavi, S., Mullins, N., Nauck, M., Ng, B., Nivard, M.G., Nyholt, D.R., O'Reilly, P.F., Oskarsson, H., Owen, M.J., Painter, J.N., Pedersen, C.B., Pedersen, M.G., Peterson, R.E., Pettersson, E., Peyrot, W.J., Pistis, G., Posthuma, D., Purcell, S.M., Quiroz, J.A., Qvist, P., Rice, J.P., Riley, B.P., Rivera, M., Saeed Mirza, S., Saxena, R., Schoevers, R., Schulte, E.C., Shen, L., Shi, J., Shyn, S.I., Sigurdsson, E., Sinnamon, G.B.C., Smit, J.H., Smith, D.J., Stefansson, H., Steinberg, S., Stockmeier, C.A., Streit, F., Strohmaier, J., Tansey, K.E., Teismann, H., Teumer, A., Thompson, W., Thomson, P.A., Thorgeirsson, T.E., Tian, C., Traylor, M., Treutlein, J., Trubetskoy, V., Uitterlinden, A.G., Umbricht, D., van der Auwera, S., van Hemert, A.M., Viktorin, A., Visscher, P.M., Wang, Y., Webb, B.T., Weinsheimer, S.M., Wellmann, J., Willemsen, G., Witt, S.H., Wu, Y., Xi, H.S., Yang, J., Zhang, F., Arolt, V., Baune, B.T., Berger, K., Boomsma, D.I., Cichon, S., Dannlowski, U., de Geus, E.C.J., DePaulo, J.R., Domenici, E., Domschke, K., Esko, T., Grabe, H.J., Hamilton, S.P., Hayward, C., Heath, A.C., Hinds, D.A., Kendler, K.S., Kloiber, S., Lewis, G., Li, Q.S., Lucae, S., Madden, P.F.A., Magnusson, P.K., Martin, N. G., McIntosh, A.M., Metspalu, A., Mors, O., Mortensen, P.B., Müller-Myhsok, B., Nordentoft, M., Nöthen, M.M., O'Donovan, M.C., Paciga, S.A., Pedersen, N.L., Penninx, B.W.J.H., Perlis, R.H., Porteous, D.J., Potash, J.B., Preisig, M., Rietschel, M., Schaefer, C., Schulze, T.G., Smoller, J.W., Stefansson, K., Tiemeier, H., Uher, R., Völzke, H., Weissman, M.M., Werge, T., Winslow, A.R., Lewis, C.M., Levinson, D.F., Breen, G., Børglum, A.D., Sullivan, P.F., 2018. Genome-wide

association analyses identify 44 risk variants and refine the genetic architecture of major depression. Nat. Genet. 50, 668–681. https://doi.org/10.1038/s41588-018-0090-3.

3.2 Factor analysis of lifetime psychopathology and its brain morphometric and genetic correlates in a transdiagnostic sample

This study was authored by Krug A*, Stein F*, David FS, Schmitt S, Brosch K, Pfarr J-K, Ringwald KG, Meller T, Thomas-Odenthal F, Meinert S, Thiel K, Winter A, Waltemate L, Lemke H, Grotegerd D, Opel N, Repple J, Hahn T, Streit F, Witt SH, Rietschel M, Andlauer TFM, Nöthen MM, Philipsen A, Nenadić I, Dannlowski U, Kircher T, and Forstner AJ.

*shared first authors

In this study, we investigated symptom dimensions of lifetime psychopathology present in affective and psychotic disorders. Frederike Stein and Axel Krug developed a new factor model comprising three symptom dimensions. In addition, they investigated the relation of the identified psychopathological factors to gray matter volume and cortical thickness, and I investigated the relation to common genetic factors. As the second author, I played a significant role in the planning of the scientific work, data collection, evaluation and interpretation, with a particular focus on the parts of the study related to genetics.

The study was published in *Translational Psychiatry* on June 3, 2024:

Krug A, Stein F, David FS, Schmitt S, Brosch K, … Forstner AJ. Factor analysis of lifetime psychopathology and its brain morphometric and genetic correlates in a transdiagnostic sample. *Transl Psychiatry* 2024; **14**: 235

DOI: https://doi.org/10.1038/s41398-024-02936-6

ARTICLE OPEN

Check for updates

Factor analysis of lifetime psychopathology and its brain morphometric and genetic correlates in a transdiagnostic sample

Axel Krug[1,2,19], Frederike Stein[2,3,19]✉, Friederike S. David[4], Simon Schmitt[2,3,5], Katharina Brosch[2,3,6], Julia-Katharina Pfarr[2,3], Kai G. Ringwald[2,3], Tina Meller[2,3], Florian Thomas-Odenthal[2,3], Susanne Meinert[7,8], Katharina Thiel[7], Alexandra Winter[7], Lena Waltemate[7], Hannah Lemke[7], Dominik Grotegerd[7], Nils Opel[7,9,10], Jonathan Repple[7,11], Tim Hahn[7], Fabian Streit[12,13,14], Stephanie H. Witt[12,15], Marcella Rietschel[12,15], Till F. M. Andlauer[16], Markus M. Nöthen[4], Alexandra Philipsen[1], Igor Nenadić[2,3], Udo Dannlowski[7], Tilo Kircher[2,3] and Andreas J. Forstner[4,17,18]

There is a lack of knowledge regarding the relationship between proneness to dimensional psychopathological syndromes and the underlying pathogenesis across major psychiatric disorders, i.e., Major Depressive Disorder (MDD), Bipolar Disorder (BD), Schizoaffective Disorder (SZA), and Schizophrenia (SZ). Lifetime psychopathology was assessed using the OPerational CRITeria (OPCRIT) system in 1,038 patients meeting DSM-IV-TR criteria for MDD, BD, SZ, or SZA. The cohort was split into two samples for exploratory and confirmatory factor analyses. All patients were scanned with 3-T MRI, and data was analyzed with the CAT-12 toolbox in SPM12. Psychopathological factor scores were correlated with gray matter volume (GMV) and cortical thickness (CT). Finally, factor scores were used for exploratory genetic analyses including genome-wide association studies (GWAS) and polygenic risk score (PRS) association analyses. Three factors (paranoid-hallucinatory syndrome, PHS; mania, MA; depression, DEP) were identified and cross-validated. PHS was negatively correlated with four GMV clusters comprising parts of the hippocampus, amygdala, angular, middle occipital, and middle frontal gyri. PHS was also negatively associated with the bilateral superior temporal, left parietal operculum, and right angular gyrus CT. No significant brain correlates were observed for the two other psychopathological factors. We identified genome-wide significant associations for MA and DEP. PRS for MDD and SZ showed a positive effect on PHS, while PRS for BD showed a positive effect on all three factors. This study investigated the relationship of lifetime psychopathological factors and brain morphometric and genetic markers. Results highlight the need for dimensional approaches, overcoming the limitations of the current psychiatric nosology.

Translational Psychiatry (2024)14:235 ; https://doi.org/10.1038/s41398-024-02936-6

INTRODUCTION

There is a long tradition of investigating the relationship between psychopathological syndromes and brain structure and function in patients suffering from schizophrenia (SZ) and schizoaffective disorder – henceforth referred to as schizophrenia spectrum disorders (SSD), as well as bipolar disorder (BD), and major depressive disorder (MDD). Several studies have linked specific symptoms such as verbal hallucinations to local brain structures, particularly the bilateral superior temporal gyri [1–3]. However, these have been either low in statistical power or variance [4], or limited to a specific diagnosis, such as SZ [5, 6]. This raises the question of generalizability across diagnostic categories: Since almost all symptoms can be present in different diagnoses (e.g., formal thought disorders are found in SZ, as well as in BD, and in MDD) [7–9], it is of major interest to study these syndromes transdiagnostically using dimensional approaches. Moreover, the

[1]Department of Psychiatry and Psychotherapy, University Hospital Bonn, Bonn, Germany. [2]Department of Psychiatry and Psychotherapy, University of Marburg, Marburg, Germany. [3]Center for Mind, Brain and Behavior, University of Marburg, Marburg, Germany. [4]Institute of Human Genetics, University of Bonn, School of Medicine & University Hospital Bonn, Bonn, Germany. [5]Department of Psychiatry, Social Psychiatry and Psychotherapy, Hannover Medical School, Hannover, Germany. [6]Institute of Behavioral Science, Feinstein Institutes for Medical Research, Manhasset, NY, USA. [7]Institute for Translational Psychiatry, University of Münster, Münster, Germany. [8]Institute for Translational Neuroscience, University of Münster, Münster, Germany. [9]German Centre for Mental Health (DZPG), Site Jena-Magdeburg-Halle, Jena, Germany. [10]Department of Psychiatry and Psychotherapy, University Hospital Jena, Jena, Germany. [11]Goethe University Frankfurt, University Hospital, Department of Psychiatry, Psychosomatic Medicine and Psychotherapy, Frankfurt, Germany. [12]Department of Genetic Epidemiology in Psychiatry, Central Institute of Mental Health, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany. [13]Department of Psychiatry and Psychotherapy, Central Institute of Mental Health, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany. [14]Hector Institute for Artificial Intelligence in Psychiatry, Central Institute of Mental Health, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany. [15]Central Institute of Mental Health, Medical Faculty Mannheim, Heidelberg University, J5, 68159 Mannheim, Germany. [16]Department of Neurology, Department of Neurology, Klinikum rechts der Isar, Technical University of Munich, Munich, Germany. [17]Institute of Neuroscience and Medicine (INM-1), Research Center Jülich, Jülich, Germany. [18]Centre for Human Genetics, University of Marburg, Marburg, Germany. [19]These authors contributed equally: Axel Krug, Frederike Stein. ✉email: frederike.stein@staff.uni-marburg.de

Received: 11 November 2022 Revised: 16 May 2024 Accepted: 16 May 2024
Published online: 03 June 2024

phenotypic overlap between psychiatric disorders is also reflected at a brain structural [10–13] as well as genetic level [14].

Factor analyses of lifetime psychopathology have mostly been performed within one categorical disorder. Only a few studies are available, investigating transdiagnostic symptom dimensions of lifetime psychopathology across diagnoses: Investigating patients with DSM-IV diagnosed SZ, BD and delusional disorder, Serretti and Olgiati found that a five-factor model best described lifetime symptom dimensions [15]. In a sample consisting of patients with SZ, BD, MDD, delusional, and psychotic disorder not otherwise specified, a four-factor solution was obtained, consisting of excitement, psychotic features (hallucinations and delusions), depression and disorganization [16]. Studying the factor structure of the OPerational CRITeria (OPCRIT) system in the SZ spectrum and BD, Reininghaus and colleagues obtained a bifactor model with one transdiagnostic psychosis dimension and five specific factors comprising positive, negative, manic, disorganized and depressive symptoms [17].

Previously, most structural and functional magnetic resonance neuroimaging studies focused on categorical comparisons of one patient group (MDD, BD, or SSD) compared to a healthy control (HC) group. However, these studies failed to identify structural and functional brain correlates that separate disorders [18]. In contrast, studies and meta-analyses indicated common alterations across diagnoses [11–13, 19, 20]. Transdiagnostic studies of dimensional psychopathology might thus be more promising regarding identification of common risk factors and might especially lead to a more precise treatment of these syndromes on a transdiagnostic rather than diagnosis-based level. In addition, they should be able to take into account the heterogeneity of psychiatric disorders as well as potential comorbidities. This should also help to identify specific neurobiological markers which in turn can inform personalized treatment interventions.

Twin and family studies demonstrate that genetic factors contribute substantially to the development of MDD, BD and SZ, with heritability estimates of around 60% to 85% for SZ and BD [21–23] and around 40% for MDD [24]. Recent genome-wide association studies (GWAS) have identified numerous genome-wide significant loci for all three psychiatric disorders (e.g., refs. [25, 26]). Furthermore, transdiagnostic GWAS meta-analyses have demonstrated an extensive genetic overlap between MDD, BD and SZ [14]. Byrne et al. provided evidence that only a small subset of the genome-wide significant variants for SZ and MDD have disorder-specific effects [27]. One plausible hypothesis, therefore, is that pleiotropic genetic variants mediate their disease risk via effects on transdiagnostic symptom dimensions. In addition, an analysis of polygenic risk scores (PRS), which summarize the effects of multiple common genetic variants into an individual genetic risk profile [28], by Ruderfer et al. showed that the PRS for SZ was significantly increased in BD patients with psychotic features and SZ patients with prominent negative symptoms [29]. These results suggest that there are genetic factors underlying specific symptom dimensions within both disorders [29].

As symptom presentations can fluctuate within an individual patient over the course of life and even within a single episode, the aim of the present study was to i) assess lifetime symptoms in a transdiagnostic sample to identify underlying symptom factors; and ii) investigate the relationship of detected factors with local GMV and CT. Considering that brain structure is less variable within a short period of time, we hypothesize that this approach would yield more conclusive results than correlating GMV with psychopathology present at any given point in time. In addition, applying both GMV as well as CT measures should render a fuller picture of underlying mechanisms as we would not assume that all potential associations would be based on one measure alone. Finally, iii) it was explored if the detected factor structure can be linked to common genetic variation. Based on previous brain-morphometric and genetic studies, we hypothesized findings from specific DSM-IV diagnostic categories to be present across diagnoses, too.

MATERIAL AND METHODS

Participants

Patients were recruited as part of the FOR2107 cohort [30] (www.for2107.de). Patient recruitment took place via the in-patient facilities of the University hospitals in Marburg and Münster, Germany, through participating hospitals, and via postings in local newspapers. Written informed consent was obtained from all patients before participation. According to the Declaration of Helsinki, all procedures were approved by the local ethics committees. After study participation, all patients received financial compensation. After excluding patients with incomplete data, serious medical illnesses, neurological illnesses, and an IQ < 80, we analyzed a total of $N = 1038$ patients (see Table 1a, b, required sample size is based on [31]) suffering from MDD, BD, and SSD (aged 18–65).

Psychopathological assessment and factor score calculation

The German version of the structured interview SCID-I (DSM-IV-TR [32]) and the OPCRIT (version 4 [33]) were administered in all patients. Lifetime psychopathology was assessed as any occurrence of symptoms during the life span until data acquisition. Trained personnel assessed lifetime symptoms based on patients' reports and additional hospital records, when available. Numerous interview trainings assured data quality. Interrater-reliability was assessed with the interclass coefficient, achieving good reliability of $r > 0.86$. For the present study, only symptomatic items were included (items 17–77). Following the procedures described in Stein et al. [34], we separated the total cohort ($N = 1038$) into two samples using the "mindiff" [35] package in R [36] accounting for age, sex, and diagnostic category (i.e., same distribution of categorical diagnoses across both samples). In the first sample of $n = 520$, we performed varimax rotated principal axis factor analyses with bootstrapping (5000 permutations) using the *psych* [37] and *EFAutilities* [38] packages in R (v4.0.5.) for models with 2–5 factors. Hereof, z-transformed values were used since the data was differently scaled. For interpretation purposes, items with factor loadings <0.5 were not considered in the analysis [34]. Cronbach's alpha coefficients [39] were used to test the internal consistency of the explorative factors. Using the second sample of $n = 518$, we cross-validated the explorative models using confirmatory factor analysis in Mplus (version 8.4 [40]). Confirmatory model estimation was performed using the maximum-likelihood-method (MLM) since this estimator is robust to standard errors and is one of the most common estimators [41]. The following fit indices were used: chi-square significance test, comparative fit index (CFI [42]) and Root Mean Square Error of Approximation (RMSEA [43]). Based on these fit indices, we evaluated the different models and selected the one with best fit. After cross-validating the explorative model in the second sample, we tested the model for the whole sample ($N = 1038$).

As the DSM-IV-TR diagnostic groups were unequally distributed, we wanted to rule out potential confounding effects of formal diagnosis. Therefore, we tested the selected factorial model in an age- and sex-matched sample with an equal diagnosis distribution (each $n = 108$ of MDD, BD, SSD, total $n = 324$) (see supplement eTable1). Matching was performed using the "MatchIt" package [44] in R [36]. Furthermore, the factorial model was also tested within each of the three diagnostic categories and factor loadings were compared between DSM diagnosis using non-parametric Kruskal–Wallis tests (see supplement).

MRI assessment and preprocessing

Subjects were scanned with a 3-T MRI at the Department of Psychiatry and Psychotherapy in Marburg (Tim Trio, Siemens, Erlangen, Germany; 12-channel head coil) and the Institute for Translational Psychiatry in Münster (Prisma, Siemens, Erlangen, Germany; 20-channel head coil). MRI data were acquired according to an extensive quality assurance protocol [45]. A fast gradient echo MP-RAGE sequence with a slice thickness of 1.0 mm consisting of 176 sagittal orientated slices in Marburg and 192 in Münster and a FOV of 256 mm was used to acquire T1 weighted images. Parameters differed across sites: Marburg: TR = 1.9 s, TE = 2.26 ms, TI = 900 ms, flip angle = 9°; Münster: TR = 2.13 s, TE = 2.28 ms, TI = 900 ms, flip angle = 8°.

For a detailed description of the preprocessing of MRI data please see refs. [31, 46]. In short, both voxel-based-morphometry GMV and cortical thickness

Table 1. a: Characteristics of the explorative sample $n = 520$. b: Characteristics of the confirmatory sample $n = 518$.

	Major depressive disorder ($n = 402$)	Bipolar disorder ($n = 64$)	Schizophrenia spectrum disorders ($n = 54$)	Group comparison (F/Chi-values in brackets)
a				
Age	36.53 (13.45)	41.05 (11.96)	37.46 (11.75)	$p = 0.005$[a] (5.34)
Sex	$m = 140$ $f = 262$	$m = 30$ $f = 34$	$m = 28$ $f = 26$	$p = 0.016$ (8.15)
Years of education	13.25 (2.68)	13.5 (2.79)	12.54 (2.45)	$p = 0.213$ (1.55)
Age of onset	26.14 (12.64)	21.97 (10.66)	22.54 (10.26)	$p = 0.034$[b] (3.4)
TIV	1563.7 (160.75)	1578.16 (149.26)	1580.85 (152.16)	$p = 0.685$ (0.38)
	Major depressive disorder ($n = 401$)	**Bipolar disorder ($n = 63$)**	**Schizophrenia spectrums disorders ($n = 54$)**	**Group comparison (F/Chi-values in brackets)**
b				
Age	36.53 (12.84)	41.03 (12.31)	37.48 (11.27)	$p = 0.023$[a] (3.8)
Sex	$m = 140$ $f = 261$	$m = 30$ $f = 33$	$m = 29$ $f = 25$	$p = 0.008$ (9.67)
Years of education	13.25 (2.79)	14.47 (2.8)	12.51 (2.84)	$p = 0.001$[b] (1.55)
Age of onset	25.52 (12.28)	26.0 (11.18)	22.89 (8.27)	$p = 0.23$ (1.48)
TIV	1558.8 (148.18)	1596.85 (134.22)	1584.76 (206.19)	$p = 0.193$ (1.65)

TIV total intracranial volume.
[a]MDD < BD.
[b]MDD < BD; SSD < BD.

(CT) data were preprocessed using the default parameters as implemented in the CAT12-Toolbox (Computation Anatomy Toolbox for SPM, build 1184, Structural Brain Mapping group, Jena University Hospital, Germany) building on SPM12 (Statistical Parametric Mapping, Institute of Neurology, London, UK). We opted for GMV and CT over other MRI-derived metrics for two primary reasons. Firstly, volume and thickness measures, commonly employed in large-scale analyses such as those by the ENIGMA consortium, were selected to facilitate result comparisons. Second, recent neuroimaging research has underscored the complementary nature of GMV and CT measurements. GMV provides insight into overall gray matter volume, which can reflect global brain atrophy or neurodevelopmental factors. In contrast, CT offers information about the thickness of the cortical mantle, allowing for the detection of localized changes. By analyzing both GMV and CT, we aimed to capture both global and local structural alterations in the context of these psychiatric disorders [47, 48]. Images were spatially registered, segmented [49] and normalized [50]. CT preprocessing included fully-automated methods projecting local maxima to other GM voxels using a neighbor relationship described by the white matter distance [51]. Quality control of processed data was performed as implemented in CAT12. For GMV data, a Gaussian kernel of 8 mm FWHM was used for smoothing. For CT data, a Gaussian kernel of 20 mm FWHM was used.

Statistical analyses: gray matter volume and cortical thickness

For both GMV and CT analyses, we used separate linear regression models for each factor using CAT12 and SPM12. The following nuisance variables were included in brain structural analyses: age, sex, and two dummy-coded variables accounting for the different MRI scanners and a body coil exchange in Marburg (Marburg pre body coil: yes/no, Marburg post body coil: yes/no, Münster as reference category [30, 45]). To control for potential medication effects, we used three dummy coded (yes/no) covariates accounting for the current medication with antidepressants, mood stabilizers and antipsychotics. For GMV analyses total intracranial volume was additionally accounted as covariate of no interest and absolute threshold masking with a threshold value of 0.1 was used.

To further test confounding effects of unequally distributed diagnostic categories, we performed multiple regression analyses in the age and sex matched sample ($n = 324$) with same n per diagnosis, again. Besides this whole brain analysis, we additionally performed ROI analyses of the detected clusters from the total sample in the matched sub-sample (see supplement).

In addition to multiple regression analyses, we performed full factorial ANCOVA whole brain interaction analyses (factor x diagnosis) for each of the three factors to test whether transdiagnostic brain correlates were driven by single DSM-IV-TR diagnostic categories for both the total and the matched sample with same n per diagnosis. Moreover, post hoc interaction analyses (factor x diagnosis) were performed specifically within each detected cluster of the total sample using the "lm-function" in R.

Cluster labeling was applied using the dartel space Neuromorphometrics atlas (http://www.neuromorphometrics.com/) for GMV analyses and for CT analyses using the Desikan–Killiany atlas [52]. Results were suggested significant at $p < 0.05$ peak-level, family wise error (FWE) corrected, cluster extend $k = 35$ voxels in the total and matched sample.

GWAS and PRS association analysis

DNA extraction, genome-wide genotyping, quality control and imputation were carried out as previously described [53] in the full FOR2107 cohort. Briefly, genotyping was performed using genomic DNA from blood samples and the Infinium PsychArray BeadChip (Illumina, San Diego, CA, USA). Pre-imputation quality control (QC) was performed in GenomeStudio, PLINK v1.9 [54], and R [36], with removal of genetic variants and individuals according to standard filter criteria. Genotype data were imputed to the 1000 Genomes phase 3 reference panel [55] using SHAPEIT [56] and IMPUTE2 [57]. In post-imputation QC, variants with a minor allele frequency <1%, Hardy-Weinberg equilibrium $p < 1e{-}6$, and an INFO metric <0.8 were removed. From the total sample of the present study ($N = 1038$), high-quality genotype data were available for 951 individuals. From these, 13 individuals were excluded due to intra-sample relatedness ($\hat{\pi} > 0.125$), resulting in a sample of $n = 938$ individuals used for genetic analyses.

For each of the three factors, GWAS, which should be considered exploratory at the given sample size, were conducted via linear regression in PLINK with rank-based inverse normal transformed values [58] as quantitative phenotypes due to the non-normal distribution of factor scores. Sex, age and the first four multidimensional scaling (MDS) components were included as covariates. All variables were z-scaled via the 'standard-beta' modifier for better comparability between factor dimensions. We performed clumping of genetic markers in the GWAS results using a maximum p value of 1e−4 for index variants ('--clump-p1 1e−4'), an LD threshold of 0.1 ('--clump-r2 0.1'), and a window size of 1000 kb ('--clump-kb 500'). We considered genetic associations with $p < 5e{-}8$ to be genome-wide significant and with $p < 1e{-}6$ to be suggestive. We performed gene-based and gene-set analyses with MAGMA [59] as implemented in FUMA [60]. The resulting p values were corrected for multiple testing using the Bonferroni method taking into account the number of tested genes ($n = 18{,}846$) or gene sets ($n = 10{,}678$). We used LocusZoom [61] to generate regional plots.

Table 2. Explorative factor model of sample 1, $n = 520$.

Factor	Item	Symptom	Loading	Cronbach´s Alpha
1 (PHS)	Opcrit61	Delusions of passivity	0.711	0.910
	Opcrit64	Delusions and hallucinations last for one week	0.700	
	Opcrit68	Thought broadcast	0.697	
	Opcrit66	Thought insertion	0.681	
	Opcrit58	Delusions of influence	0.681	
	Opcrit62	Primary delusional perception	0.657	
	Opcrit55	Well organized delusions	0.650	
	Opcrit60	Widespread delusions	0.621	
	Opcrit54	Persecutory delusions	0.620	
	Opcrit74	Running commentary voices	0.602	
	Opcrit73	Third person auditory hallucinations	0.571	
	Opcrit59	Bizarre delusions	0.552	
	Opcrit67	Thought withdrawal	0.549	
	Opcrit63	Other primary delusions	0.543	
	Opcrit77	Non-affective hallucination in any modality	0.522	
2 (MA)	Opcrit35	Elevated mood	0.890	0.921
	Opcrit19	Excessive activity	0.829	
	Opcrit30	Pressured speech	0.804	
	Opcrit56	Increased self esteem	0.795	
	Opcrit20	Reckless activity	0.765	
	Opcrit22	Reduced need for sleep	0.760	
	Opcrit31	Thoughts racing	0.732	
	Opcrit21	Distractibility	0.525	
3 (DEP)	Opcrit39	Loss of pleasure	0.678	0.736
	Opcrit25	Loss of energy/tiredness	0.635	
	Opcrit37	Dysphoria	0.603	
	Opcrit41	Lack of concentration	0.507	

PRS for MDD, BD and SZ were calculated based on publicly available summary statistics from three studies [25, 26, 62]. Variant weights for PRS calculation were estimated with PRS-CS [63] using default parameters and a set of pre-defined values for the global shrinkage parameter φ (1e−4, 1e−3, 1e−2). PRS were subsequently calculated in R [36] as described previously [64]. Linear additive models with rank-based inverse normal transformed factor scores as outcome, one of the z-scaled disorder-specific PRS as predictor and sex, age and the first four MDS components as covariates were fitted in R. The PRS association analysis was conducted for both the complete set of $n = 938$ individuals as well as for each diagnostic subgroup separately. Adjustment of p values for multiple testing was performed with the Benjamini–Hochberg approach [65] within each set of 27 tests (3 factor dimensions * 3 PRS models * 3 values for φ). Model coefficients were considered to be statistically significant at $p < 0.05$. We calculated the variance explained (R^2) by each PRS as the difference between R^2 of the full model and R^2 of the null model containing only the covariates.

RESULTS

Exploratory and confirmatory factor analyses

We tested explorative models ranging from 2-5 factors. Results of these models can be found in Supplementary eTables 2a–d. In a next step, we evaluated the four explorative models using confirmatory analyses in the second sample. Model fits were as follows: a) 2 factors: $\chi^2 = 393.645$, $df = 224$, $p < 0.001$, $CFI = 0.903$, $RMSEA = 0.038$; b) 3 factors: $\chi^2 = 543.005$, $df = 316$, $p < 0.001$, $CFI = 0.904$, $RMSEA = 0.037$; c) 4 factors: $\chi^2 = 588.773$, $df = 314$, $p < 0.001$, $CFI = 0.875$, $RMSEA = 0.042$; d) 5 factors: $\chi^2 = 748.705$, $df = 391$, $p < 0.001$, $CFI = 0.884$, $RMSEA = 0.041$. Based on the fit indices, we decided to use model b) with 3 factors (Table 2) for further analyses as this model outperformed the other ones. Moreover, a 3-factor model is also in line with the Scree Plot (Supplementary eFigure 1). The model included the factors paranoid-hallucinatory syndrome (PHS) (explaining 14% of total variance), mania (MA) (explaining 11% of total variance), and depression (DEP) (explaining 5% of total variance). Furthermore, we performed a confirmatory factor analysis in the whole sample ($N = 1038$) showing a good fit $\chi^2 = 605.667$, $df = 316$, $p < 0.0001$, $CFI = 0.932$, $RMSEA = 0.03$. Results of the confirmatory analyses of the matched sample and within each diagnostic category are presented in the supplement (Supplementary eResults1 and 2). We investigated differences of the factor loadings between diagnostic categories using a non-parametric ANOVA (Kruskal–Wallis). Diagnostic groups differed significantly in all three factors identified (Supplementary eResults3 and Supplementary eFigure 2).

Brain morphometric correlates of life-time psychopathology

Results of the multiple regression analyses of the total sample are displayed in Table 3 (GMV) and 4 (CT). For the paranoid-hallucinatory syndrome (PHS), negative GMV correlations were observed in the bilateral hippocampus, amygdala, and right angular gyrus (see Fig. 1). CT was negatively correlated with the paranoid-hallucinatory syndrome (PHS) comprising left supramarginal, bilateral superior temporal, and right lateral occipital clusters (see Fig. 2). Whole-brain interaction analyses revealed no significant interaction of psychopathological factor and DSM-IV-TR diagnosis for both GMV and CT. Post hoc interaction analyses on the significant clusters in Tables 3 and 4 revealed no significant

Table 3. Results of the lifetime paranoid-hallucinatory syndrome (PHS) and its local gray matter (GMV) correlates.

		MNI coordinates				
Anatomical region	**H**	**X**	**Y**	**Z**	**t-value**	**Cluster size**
Factor I: Paranoid-hallucinatory syndrome: gray matter volume						
Entorhinal area, hippocampus, amygdala, parahippocampal gyrus, fusiform gyrus, temporal pole	L	−27	−8	−27	5.53	560
Angular gyrus, middle occipital gyrus	R	57	−64.5	22.5	5.13	150
Amygdala, hippocampus, entorhinal area	R	25.5	−4.5	−27	5.00	83
Medial frontal cerebrum	R	3	57	−10.5	4.81	64

Only negative correlations are reported as no positive correlations were detected.
H hemisphere, *L* left, *R* right.

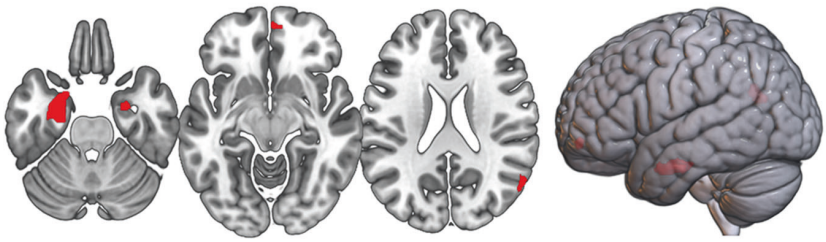


Fig. 1 Local GMV correlates of the lifetime paranoid-hallucinatory syndrome (PHS). Negative association of factor 1 paranoid-hallucinatory syndrome (PHS) and gray matter volume (GMV) comprising parts of the bilateral hippocampus, amygdala, and right angular gyrus across patients with major depressive disorder, bipolar disorder, and schizophrenia spectrum disorders. Clusters are shown at $p < 0.05$ peak-level, family-wise error-corrected.

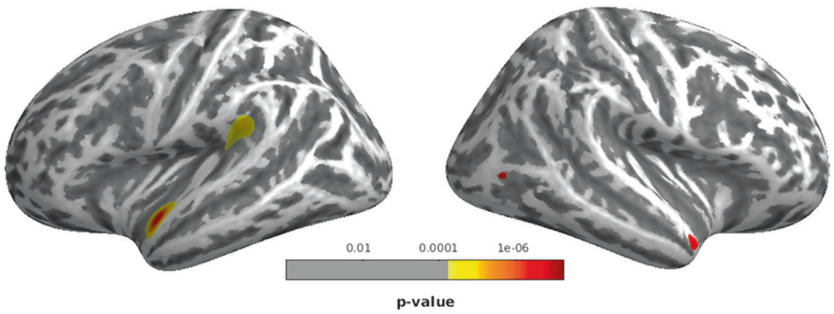


Fig. 2 CT correlates of the paranoid-hallucinatory syndrome (PHS). Negative association of factor 1 paranoid-hallucinatory syndrome (PHS) and cortical thickness (CT) comprising parts of left supramariginal, bilateral superior temporal, and right lateral occipital clusters across patients with major depressive disorder, bipolar disorder, and schizophrenia spectrum disorders. Clusters are shown at $p < 0.05$ peak-level, family-wise error-corrected.

interactions of factor x diagnosis (all *ps* > 0.05, see Supplementary eResults4 and Supplementary eFig. 3–10 for details). Results of the GMV and CT analyses in the matched sample are presented in the supplement (Supplementary eResults 5, Supplementary eTables 3 and 4). Here, results from the total sample could be replicated. We did not find any associations with the DEP and MA factors for both GMV and CT.

Genetic correlates of life-time psychopathology

Our exploratory GWAS revealed genome-wide significant associations for MA and DEP (Fig. 3, Supplementary eFigs. 11–13, Supplementary eTable 5), with intronic lead variants rs10062519 ($p = 1.10e{-}8$) located in *ADAMTS19* for MA and rs11131155 ($p = 4.12e{-}8$) located in *RAD18* for DEP. In the MAGMA gene analysis, a genome-wide significant association was identified for *SYTL1* (DEP, $p = 1.79e{-}6$). The MAGMA gene-set analysis yielded no statistically significant results for any of the three factors after correction for multiple testing (data not shown).

In the PRS association analysis of the complete sample (Fig. 4), we detected a positive effect of PRS for BD on all three factors (PHS: maximum $\beta = 0.13$ at $\varphi = 1e{-}3$ with $R^2 = 0.021$ and adjusted $p = 5.48e{-}5$; MA: maximum $\beta = 0.18$ at $\varphi = 1e{-}3$ with $R^2 = 0.031$ and adjusted $p = 1.17e{-}6$; DEP: maximum $\beta = 0.08$ at $\varphi = 1e{-}4$ with $R^2 = 0.006$ and adjusted $p = 0.038$). Further, a positive effect on PHS was observed for the PRS for MDD (maximum $\beta = 0.07$ at $\varphi = 1e{-}2$ with $R^2 = 0.006$ and adjusted $p = 0.038$) and SZ (maximum $\beta = 0.13$ at $\varphi = 1e{-}3$ with $R^2 = 0.020$ and adjusted $p = 7.06e{-}5$). In the subset analysis of each diagnostic group, none of the effects observed in the complete transdiagnostic sample reached statistical significance (Supplementary eFig. 14).

DISCUSSION

In the present study, exploratory and confirmatory factor analyses of lifetime psychopathology revealed a three-factor model with

Table 4. Results of the lifetime paranoid-hallucinatory syndrome (PHS) and its cortical thickness correlates.

		MNI Coordinates				
Anatomical region	**H**	**X**	**Y**	**Z**	**t-value**	**Cluster size**
Factor I: Paranoid-hallucinatory syndrome: Cortical thickness						
Superior temporal cortex	L	−48.3	−3.62	−12.1	5.07	657
Supramarginal cortex	L	−48.3	−38.5	25.9	4.34	777
Superior temporal cortex	R	48.6	15.6	−22.6	4.11	236
Lateral occipital cortex	R	44.6	−71.7	3.1	4.00	47

Only negative correlations are reported as no positive correlations were detected.
H hemisphere, *L* left, *R* right.

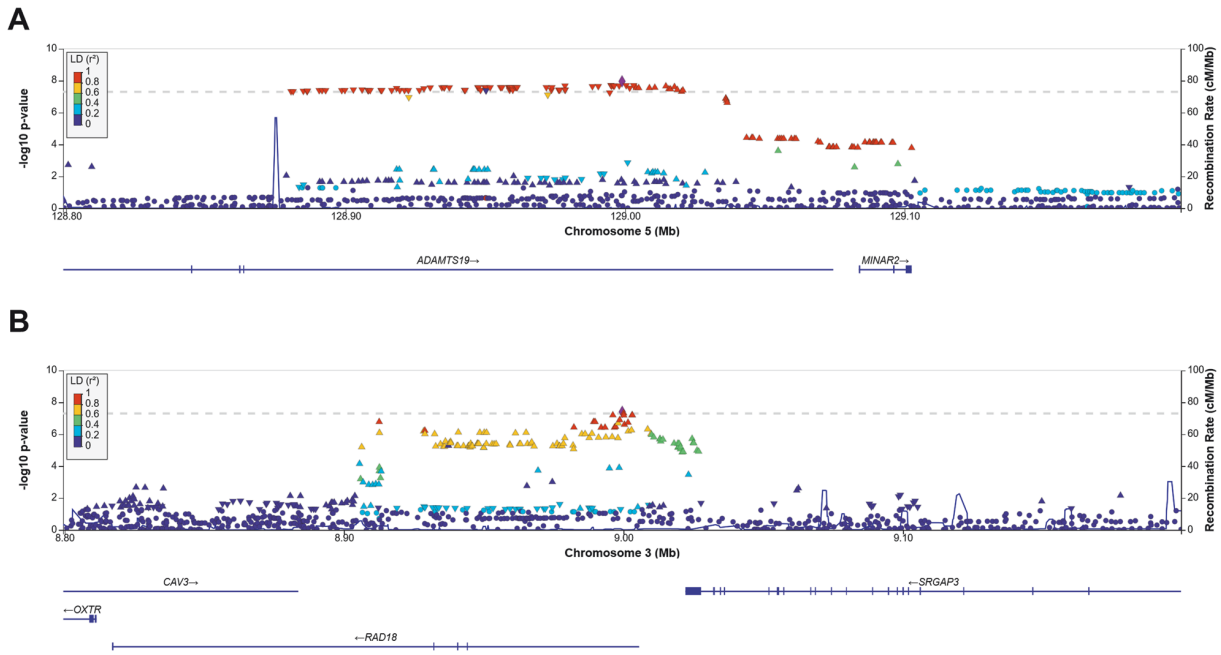


Fig. 3 Genetic loci with genome-wide significant association. Regional plots with a window size of 500 kb are shown for the genome-wide significant associations with MA "mania" (**A**), and DEP "depression" (**B**). The respective lead variants rs10062519 and rs11131155 are depicted as linkage disequilibrium reference variants (purple diamonds). cM centimorgan, LD Ref Var, linkage disequilibrium reference variant, Mb megabase.

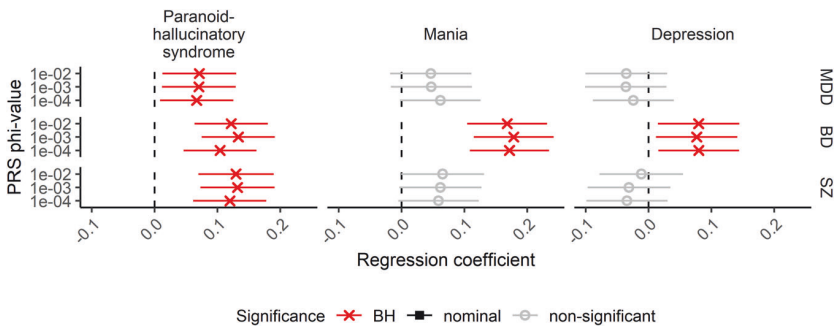


Fig. 4 PRS association analysis. Regression of the three factors on the PRS for MDD, BD, and SZ shows significant effects of PRS for MDD, BD, and SZ on PHS "paranoid-hallucinatory syndrome" and of PRS for BD on MA "mania" and DEP "depression" in the full transdiagnostic sample. BD bipolar disorder, BH Benjamini–Hochberg, MDD major depressive disorder, SZ schizophrenia.

superior fit properties compared to models with less or more factors. Factors were the paranoid-hallucinatory syndrome (PHS), mania (MA) and depression (DEP). In addition, several associations with both brain morphometry and genetics were reported. This study represents a successful advancement of previous research by Stein et al. [34] and David et al. [66] all of them part of the FOR2107 cohort, wherein five factors of acute psychopathology were described and genetically investigated.

Compared to previous factor analytical research in the three diagnoses included in our study, utilizing the OPCRIT, the present factor solution features three factors, while other studies showed an additional negative factor, which was not present in our study. Nevertheless, an overlap exists as previous models also comprised a depression factor and a mania factor (e.g., refs. [15, 67]). The often-reported factors positive and negative symptoms are split into all three factors in the present results while disorganization best fits the present second factor MA.

The derived lifetime psychopathological factors were used to investigate underlying GMV and CT correlates. We were able to detect numerous associations between the PHS and both GMV and CT in both temporal and frontal regions. We did not detect any interactions for both factor x diagnosis on a whole-brain level, nor in post hoc analyses of the significant clusters. These findings do not exclude that the severity of both brain structural alterations and psychopathological syndromes may vary by diagnosis. Our study aligns with previous studies proposing overlaps in acute psychopathology, brain structure as well as genetics across MDD, BD and SSD [11–13, 68]. Combining a data-driven approach to psychopathology with studying neuroanatomical and genetic correlates may help elucidate the biological underpinnings of complex syndromes in psychiatric disorders. Approaches such as those applied in the present study can reveal intra- and inter-disorder heterogeneity and thus could support the establishment of treatments specific to symptom or syndrome in the next step.

When comparing our results to previous dimensional studies, a recent study also identified subcortical volume reductions associated with hallucinations as well as delusions [69], but reductions of superior temporal areas have also been well established in SSD [1, 3, 70, 71]. The present findings are also in line with a recent investigation where psychotic symptoms were negatively correlated with CT in a large sample of SSD patients, relatives and healthy controls [72]. Consistent with previous studies in SSD, we found cortical thinning in the bilateral STG to be correlated with the PHS factor [73], indicating this brain structure to be a core feature of positive symptomatology.

Exploratory GWAS and PRS analyses suggest a contribution of common genetic variants to all three factor dimensions, supporting the hypothesis that symptoms observed in different diagnostic groups may be influenced by the same genetic variants across diagnostic boundaries [14, 29]. Interestingly, the genome-wide significant loci of our GWAS implicated protein-coding genes that both might be linked to psychiatric disorders. *ADAMTS19* is a member of the ADAMTS (a disintegrin and metalloproteinase with thrombospondin motif) family [74], which might be involved in neuroplasticity [75]. The *RAD18* gene encodes for a DNA damage repair protein [74, 76]. Notably, a study by Alsulami and colleagues provided evidence that RAD18 interacts with SETD1A [76], which has previously been associated with SSD at the rare variant level [77, 78]. As it is known that genome-wide significant lead variants do not necessarily exert their effects through the nearest genes (e.g., ref. [79]), the above discussed functional interpretations should be viewed with caution, as further bioinformatic and functional analyses are needed to identify the gene(s) relevant at the identified loci.

Finally, despite the associations at the genetic level, we did not detect an association between the MA or DEP factor and brain morphology. This suggests that even though aspects of lifetime psychopathology might at least be partially influenced by genetic factors, this might not necessarily be detectable on a neural level. It could thus be argued that a dimensional approach is even more important than a narrow nosology as these associations might be subtle and implications for translation into treatment options are not as clear, yet.

Limitations

There are several limitations to be considered: First, as lifetime psychopathology was assessed only at one point in this study, a bias may arise in favor of symptoms that have occurred recently or are currently present, as they could be more salient in the individual's memory. This bias could lead to an overemphasis on these symptoms during the assessment process [80]. As a result, symptoms that occurred in the past may be underreported or forgotten entirely. We tried to circumvent these biases by carefully examining every hospital record available for each patient, but these were not available for all patients included here. In addition, the used psychopathological scale did not include the full symptomatic spectrum, which restricted the identification of psychopathological factors.

Second, sample sizes of each diagnostic category were unequal. The aim of the present study was to investigate syndrome-brain structure and syndrome-genetic associations dimensionally rather than within categorical diagnoses. The presence of psychotic and manic symptoms in the MDD group might be limited, which may result in restricted variance found for the PHS factor. While results can be considered as diagnosis-shared, severity may be differing across diagnoses.

Third, MRI techniques in general might not be able to detect subtle differences in locations of effects if these occurs in close proximity. In addition, true effects between groups might be mapped onto the same neural circuit while in fact there are differences on the underlying cellular level [81].

Fourth, pharmacological treatment was considered as three dummy coded variables to account for the current intake of antidepressants, antipsychotics, and mood stabilizers. This approach does not take into account both the dosage and lifetime cumulative intake of psychotropic medication, which might have influenced our results.

Finally, the available sample size represents a limitation for the genetic analyses, as the robust detection of genetic associations with small effect sizes usually requires meta-analytical efforts involving multiple cohorts [82]. Thus, the exploratory nature of the presented GWAS should be considered in the interpretation of our findings.

CONCLUSION

This study comprehensively investigated the association of lifetime psychopathological dimensions and brain morphometric markers as well as underlying genetic factors. At the level of brain imaging, GMV and CT reductions in temporal, occipital, and limbic structures were found to be correlated with paranoid-hallucinatory symptoms in a transdiagnostic sample. On the genetic level, we identified genome-wide significant loci for MA and DEP factors, as well as positive effects of specific PRS on different factors. These findings suggest that genetic factors contribute to the identified factor dimensions. The results presented in this study highlight the importance of i) dimensional modeling and ii) transdiagnostic research gaining a better understanding of pathophysiological mechanisms underlying MDD, BD and SSD.

DATA AVAILABILITY

The data and code supporting the findings of this study can be accessed by contacting the corresponding author (FS).

REFERENCES

1. Modinos G, Costafreda SG, van Tol M-J, McGuire PK, Aleman A, Allen P. Neuroanatomy of auditory verbal hallucinations in schizophrenia: A quantitative meta-analysis of voxel-based morphometry studies. Cortex. 2013;49:1046–55.
2. Wong TY, Radua J, Pomarol-Clotet E, Salvador R, Albajes-Eizagirre A, Solanes A et al. An overlapping pattern of cerebral cortical thinning is associated with both

positive symptoms and aggression in schizophrenia via the ENIGMA consortium. Psychol Med. 2020;50:2034–45.
3. Palaniyappan L, Balain V, Radua J, Liddle PF. Structural correlates of auditory hallucinations in schizophrenia: A meta-analysis. Schizophr Res. 2012;137:169–73.
4. Kim GW, Kim YH, Jeong GW. Whole brain volume changes and its correlation with clinical symptom severity in patients with schizophrenia: A DARTEL-based VBM study. PLoS One. 2017;12. https://doi.org/10.1371/journal.pone.0177251.
5. Kircher T, Markov V, Krug A, Eggermann T, Zerres K, Nöthen MM, et al. Association of the DTNBP1 genotype with cognition and personality traits in healthy subjects. Psychol Med. 2009;39:1657–65.
6. Nenadic I, Sauer H, Gaser C. Distinct pattern of brain structural deficits in subsyndromes of schizophrenia delineated by psychopathology. Neuroimage. 2010;49:1153–60.
7. Kircher T, Krug A, Stratmann M, Ghazi S, Schales C, Frauenheim M, et al. A rating scale for the assessment of objective and subjective formal thought and language disorder (TALD). Schizophr Res. 2014;160:216–21.
8. Andreasen NC, Grove WM. Thought, language, and communication in schizophrenia: diagnosis and prognosis. Schizophr Bull. 1986;12:348–59.
9. Stein F, Buckenmayer E, Brosch K, Meller T, Schmitt S, Ringwald KG et al. Dimensions of Formal Thought Disorder and Their Relation to Gray- and White Matter Brain Structure in Affective and Psychotic Disorders. Schizophr Bull. 2022. https://doi.org/10.1093/SCHBUL/SBAC002.
10. Patel Y, Parker N, Shin J, Howard D, French L, Thomopoulos SI et al. Virtual Histology of Cortical Thickness and Shared Neurobiology in 6 Psychiatric Disorders. JAMA Psychiatry. 2020. https://doi.org/10.1001/jamapsychiatry.2020.2694.
11. Goodkind M, Eickhoff SB, Oathes DJ, Jiang Y, Chang A, Jones-Hagata LB, et al. Identification of a Common Neurobiological Substrate for Mental Illness. JAMA Psychiatry. 2015;72:305.
12. Brosch K, Stein F, Schmitt S, Pfarr JK, Ringwald KG, Thomas-Odenthal F et al. Reduced hippocampal gray matter volume is a common feature of patients with major depression, bipolar disorder, and schizophrenia spectrum disorders. Mol Psychiatry. 2022;27. https://doi.org/10.1038/S41380-022-01687-4.
13. Repple J, Gruber M, Mauritz M, de Lange SC, Winter NR, Opel N et al. Shared and specific patterns of structural brain connectivity across affective and psychotic disorders. Biol Psychiatry. 2022. https://doi.org/10.1016/J.BIOPSYCH.2022.05.031.
14. Lee PH, Anttila V, Won H, Feng YCA, Rosenthal J, Zhu Z, et al. Genomic Relationships, Novel Loci, and Pleiotropic Mechanisms across Eight Psychiatric Disorders. Cell. 2019;179:1469–82.e11.
15. Serretti A, Olgiati P. Dimensions of major psychoses: A confirmatory factor analysis of six competing models. Psychiatry Res. 2004;127:101–9.
16. Serretti A, Rietschel M, Lattuada E, Krauss H, Schulze TG, Müller DJ, et al. Major psychoses symptomatology: Factor analysis of 2241 psychotic subjects. Eur Arch Psychiatry Clin Neurosci. 2001;251:193–8.
17. Reininghaus U, Böhnke JR, Hosang G, Farmer A, Burns T, McGuffin P, et al. Evaluation of the validity and utility of a transdiagnostic psychosis dimension encompassing schizophrenia and bipolar disorder. Br J Psychiatry. 2016;209:107–13.
18. Lalousis PA, Wood SJ, Schmaal L, Chisholm K, Griffiths SL, Reniers RLEP et al. Heterogeneity and Classification of Recent Onset Psychosis and Depression: A Multimodal Machine Learning Approach. Schizophr Bull. 2021. https://doi.org/10.1093/schbul/sbaa185.
19. Chang M, Womer FY, Edmiston EK, Bai C, Zhou Q, Jiang X, et al. Neurobiological Commonalities and Distinctions Among Three Major Psychiatric Diagnostic Categories: A Structural MRI Study. Schizophr Bull. 2018;44:65–74.
20. Koutsouleris N, Meisenzahl EM, Borgwardt S, Riecher-Rössler A, Frodl T, Kambeitz J, et al. Individualized differential diagnosis of schizophrenia and mood disorders using neuroanatomical biomarkers. Brain. 2015;138:2059–73.
21. Sullivan PF, Kendler KS, Neale MC. Schizophrenia as a Complex Trait: Evidence from a Meta-analysis of Twin Studies. Arch Gen Psychiatry. 2003;60:1187–92.
22. Lichtenstein P, Yip BH, Björk C, Pawitan Y. Common genetic influences for schizophrenia and bipolar disorder: A population-based study of 2 million nuclear families. Lancet; 2009;373. https://doi.org/10.1016/S0140-6736(09)60072-6.
23. Bienvenu OJ, Davydow DS, Kendler KS. Psychiatric diseases versus behavioral disorders and degree of genetic influence. Psychol Med. 2011;41:33–40.
24. Sullivan PF, Neale MC, Kendler KS. Genetic epidemiology of major depression: Review and meta-analysis. Am J Psychiatry. 2000;157:1552–62.
25. Stahl EA, Breen G, Forstner AJ, McQuillin A, Ripke S, Trubetskoy V, et al. Genome-wide association study identifies 30 loci associated with bipolar disorder. Nat Genet. 2019;51:793–803.
26. Wray NR, Ripke S, Mattheisen M, Trzaskowski M, Byrne EM, Abdellaoui A, et al. Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. Nat Genet. 2018;50:668–81.
27. Byrne EM, Zhu Z, Qi T, Skene NG, Bryois J, Pardinas AF et al. Conditional GWAS analysis to identify disorder-specific SNPs for psychiatric disorders. Mol Psychiatry. 2020. https://doi.org/10.1038/s41380-020-0705-9.
28. Lewis CM, Vassos E. Polygenic risk scores: From research tools to clinical instruments. Genome Med. 2020;12. https://doi.org/10.1186/s13073-020-00742-5.
29. Ruderfer DM, Ripke S, McQuillin A, Boocock J, Stahl EA, Pavlides JMW, et al. Genomic Dissection of Bipolar Disorder and Schizophrenia, Including 28 Subphenotypes. Cell. 2018;173:1705–1715.e16.
30. Kircher T, Wöhr M, Nenadic I, Schwarting R, Schratt G, Alferink J, et al. Neurobiology of the major psychoses: a translational perspective on brain structure and function-the FOR2107 consortium. Eur Arch Psychiatry Clin Neurosci. 2018;1:3.
31. Stein F, Meller T, Brosch K, Schmitt S, Ringwald K, Pfarr JK et al. Psychopathological Syndromes Across Affective and Psychotic Disorders Correlate With Gray Matter Volumes. Schizophr Bull. 2021;47:1740–0.
32. Wittchen HU, Wunderlich U, Gruschwitz S, Zaudig M. SKID I. Strukturiertes Klinisches Interview für DSM-IV. Achse I: Psychische Störungen. Interviewheft und Beurteilungsheft. Eine deutschsprachige, erweiterte Bearb. d. amerikanischen Originalversion des SKID I. 1997. Göttingen: Hogrefe.
33. McGuffin P, Farmer A, Harvey I. A polydiagnostic application of operational criteria in studies of psychotic illness: Development and reliability of the OPCRIT system. Arch Gen Psychiatry. 1991;48:764–70.
34. Stein F, Lemmer G, Schmitt S, Brosch K, Meller T, Fischer E, et al. Factor analyses of multidimensional symptoms in a large group of patients with major depressive disorder, bipolar disorder, schizoaffective disorder and schizophrenia. Schizophr Res. 2020;218:38–47.
35. Papenberg M. minDiff: Minimize Differences Between Groups (R package version 0.01-3) 2019. [Computer software]. Available online at: https://github.com/m-Py/minDiff.
36. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing. 2021. Vienna, Austria. https://www.R-project.org/.
37. Revelle W. psych: Procedures for Psychological, Psychometric, and Personality Research. 2022. https://cran.r-project.org/web/packages/psych/citation.html. Accessed 16 Mar 2023.
38. Luo L, Arizmendi C, Gates KM. Exploratory Factor Analysis (EFA) Programs in R. Multidiscipl J. 2019;26:819–26.
39. Cronbach LJ. Coefficient alpha and the internal structure of tests. Psychometrika. 1951;16:234–97.
40. Muthén LK, Muthén BO. Mplus User's Guide. Eighth Edition. (1998-2017) Los Angeles, CA: Muthén & Muthén.
41. Maydeu-Olivares A. Maximum Likelihood Estimation of Structural Equation Models for Continuous Data: Standard Errors and Goodness of Fit. Struct Equ Model. 2017;24:383–94.
42. Bentler PM. Comparative Fit Indexes in Structural Models. Psychol Bull. 1990;107:238–46.
43. Steiger JH. Structural Model Evaluation and Modification: An Interval Estimation Approach. Multivar Behav Res. 1990;25:173–80.
44. Ho D, Kosuke I, King G, Stuart E. Matchit: Nonparametric Preprocessing for Parametric Causal Inference. J Stat Softw. 2011;42:1–28.
45. Vogelbacher C, Möbius TWD, Sommer J, Schuster V, Dannlowski U, Kircher T, et al. The Marburg-Münster Affective Disorders Cohort Study (MACS): A quality assurance protocol for MR neuroimaging data. Neuroimage. 2018;172:450–60.
46. Brosch K, Stein F, Meller T, Schmitt S, Yuksel D, Ringwald KG et al. DLPFC volume is a neural correlate of resilience in healthy high-risk individuals with both childhood maltreatment and familial risk for depression. Psychol Med. 2022;52:4139–45.
47. Rimol LM, Hartberg CB, Nesvåg R, Fennema-Notestine C, Hagler DJ, Pung CJ, et al. Cortical Thickness and Subcortical Volumes in Schizophrenia and Bipolar Disorder. Biol Psychiatry. 2010;68:41–50.
48. Hibar DP, Westlye LT, Van Erp TGM, Rasmussen J, Leonardo CD, Faskowitz J, et al. Subcortical volumetric abnormalities in bipolar disorder. Mol Psychiatry. 2016;21:1710–6.
49. Ashburner J, Friston KJ. Unified segmentation. Neuroimage. 2005;26:839–51.
50. Ashburner J. A fast diffeomorphic image registration algorithm. Neuroimage. 2007;38:95–113.
51. Dahnke R, Yotter RA, Gaser C. Cortical thickness and central surface estimation. Neuroimage. 2013;65:336–48.
52. Desikan RS, Ségonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, et al. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. Neuroimage. 2006;31:968–80.
53. Meller T, Schmitt S, Stein F, Brosch K, Mosebach J, Yüksel D et al. Associations of schizophrenia risk genes ZNF804A and CACNA1C with schizotypy and modulation of attention in healthy subjects. Schizophr Res. 2019;208. https://doi.org/10.1016/j.schres.2019.04.018.
54. Chang CC, Chow CC, Tellier LCAM, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: Rising to the challenge of larger and richer datasets. Gigascience. 2015;4. https://doi.org/10.1186/s13742-015-0047-8.

55. Auton A, Abecasis GR, Altshuler DM, Durbin RM, Bentley DR, Chakravarti A, et al. A global reference for human genetic variation. Nature. 2015;526:68–74.
56. Delaneau O, Marchini J, Zagury JF. A linear complexity phasing method for thousands of genomes. Nat Methods. 2012;9:179–81.
57. Howie BN, Donnelly P, Marchini J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. PLoS Genet. 2009;5. https://doi.org/10.1371/journal.pgen.1000529.
58. McCaw ZR, Lane JM, Saxena R, Redline S, Lin X. Operating characteristics of the rank-based inverse normal transformation for quantitative trait analysis in genome-wide association studies. Biometrics. 2020;76:1262–72.
59. de Leeuw CA, Mooij JM, Heskes T, Posthuma D. MAGMA: Generalized Gene-Set Analysis of GWAS Data. PLoS Comput Biol. 2015;11:1–19.
60. Watanabe K, Taskesen E, Van Bochoven A, Posthuma D. Functional mapping and annotation of genetic associations with FUMA. Nat Commun. 2017;8:1826 https://doi.org/10.1038/s41467-017-01261-5.
61. Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Gliedt TP, et al. LocusZoom: regional visualization of genome-wide association scan results. Bioinformatics Appl Note. 2010;26:2336–7.
62. Ripke S, Neale BM, Corvin A, Walters JTR, Farh KH, Holmans PA, et al. Biological insights from 108 schizophrenia-associated genetic loci. Nature. 2014;511:421–7.
63. Ge T, Chen CY, Ni Y, Feng YCA, Smoller JW. Polygenic prediction via Bayesian regression and continuous shrinkage priors. Nat Commun. 2019;10:1–10.
64. Andlauer TFM, Guzman-Parra J, Streit F, Strohmaier J, González MJ, Gil Flores S, et al. Bipolar multiplex families have an increased burden of common risk variants for psychiatric disorders. Mol Psychiatry. 2021;26:1286–98.
65. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society: Series B (Methodological). 1995;57:289-300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x1995.
66. David FS, Stein F, Andlauer TFM, Streit F, Witt SH, Herms S, et al. Genetic contributions to transdiagnostic symptom dimensions in patients with major depressive disorder, bipolar disorder, and schizophrenia spectrum disorders. Schizophr Res. 2023;252:161–71.
67. Reininghaus U, Böhnke JR, Chavez-Baldini U, Gibbons R, Ivleva E, Clementz BA, et al. Transdiagnostic dimensions of psychosis in the Bipolar-Schizophrenia Network on Intermediate Phenotypes (B-SNIP). World Psychiatry. 2019;18:67–76.
68. Anttila V, Bulik-Sullivan B, Finucane HK, Walters RK, Bras J, Duncan L, et al. Analysis of shared heritability in common disorders of the brain. Science. 2018;360:eaap8757.
69. Schoorl J, Barbu MC, Shen X, Harris MR, Adams MJ, Whalley HC et al. Grey and white matter associations of psychotic-like experiences in a general population sample (UK Biobank). Transl Psychiatry. 2021;11. https://doi.org/10.1038/s41398-020-01131-7.
70. Van Tol MJ, Van Der Meer L, Bruggeman R, Modinos G, Knegtering H, Aleman A. Voxel-based gray and white matter morphometry correlates of hallucinations in schizophrenia: The superior temporal gyrus does not stand alone. Neuroimage Clin. 2014;4:249–57.
71. Nickl-Jockschat T, Schneider F, Pagel AD, Laird AR, Fox PT, Eickhoff SB. Progressive pathology is functionally linked to the domains of language and emotion: Meta-analysis of brain structure changes in schizophrenia patients. Eur Arch Psychiatry Clin Neurosci. 2011;261. https://doi.org/10.1007/s00406-011-0249-8.
72. Stan AD, Tamminga CA, Han K, Kim JB, Padmanabhan J, Tandon N, et al. Associating Psychotic Symptoms with Altered Brain Anatomy in Psychotic Disorders Using Multidimensional Item Response Theory Models. Cereb Cortex. 2020;30:2939–47.
73. van Erp TGM, Walton E, Hibar DP, Schmaal L, Jiang W, Glahn DC, et al. Cortical Brain Abnormalities in 4474 Individuals With Schizophrenia and 5098 Control Subjects via the Enhancing Neuro Imaging Genetics Through Meta Analysis (ENIGMA) Consortium. Biol Psychiatry. 2018;84:644–54.
74. O'Leary NA, Wright MW, Brister JR, Ciufo S, Haddad D, McVeigh R, et al. Reference sequence (RefSeq) database at NCBI: Current status, taxonomic expansion, and functional annotation. Nucleic Acids Res. 2016;44:D733–D745.
75. Gottschall PE, Howell MD. ADAMTS expression and function in central nervous system injury and disorders. Matrix Biol. 2015;44–46:70–76.
76. Alsulami M, Munawar N, DIllon E, Oliviero G, Wynne K, Alsolami M, et al. SETD1A methyltransferase is physically and functionally linked to the DNA damage repair protein RAD18. Mol Cell Proteom. 2019;18:1428–36.
77. Singh T, Kurki MI, Curtis D, Purcell SM, Crooks L, McRae J, et al. Rare loss-of-function variants in SETD1A are associated with schizophrenia and developmental disorders. Nat Neurosci. 2016;19:571–7.
78. Singh T, Neale BM, Daly MJ. Exome sequencing identifies rare coding variants in 10 genes which confer substantial risk for schizophrenia on behalf of the Schizophrenia Exome Meta-Analysis (SCHEMA) Consortium. medRxiv. 2020. https://www.medrxiv.org/content/10.1101/2020.09.18.20192815v1.
79. Zhu Z, Zhang F, Hu H, Bakshi A, Robinson MR, Powell JE, et al. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. Nat Genet. 2016;48:481–7.
80. Wenze SJ, Gunthert KC, German RE. Biases in Affective Forecasting and Recall in Individuals With Depression and Anxiety Symptoms. Pers Soc Psychol Bull. 2012;38:895–906.
81. Zeighami Y, Bakken TE, Nickl-Jockschat T, Peterson Z, Jegga AG, Miller JA et al. A comparison of anatomic and cellular transcriptome structures across 40 human brain diseases. PLoS Biol. 2023;21. https://doi.org/10.1371/JOURNAL.PBIO.3002058.
82. Sullivan PF, Agrawal A, Bulik CM, Andreassen OA, Børglum AD, Breen G, et al. Psychiatric Genomics: An Update and an Agenda. Am J Psychiatry. 2018;175:15–27.

ACKNOWLEDGEMENTS

We thank the patients participating in this study and the staff for their help with recruitment and data collection. A detailed list of acknowledgements can be found here: www.for2107.de/acknowledgements.

AUTHOR CONTRIBUTIONS

AK* & FS*: data collection, study design, analysis, data interpretation, literature search, writing, figures, *contributed equally; FSD: data collection, study design, genetic analysis, data interpretation, figure, literature search, writing; TFMA, SS, KB, JKP, KGR, TM, FTO, SM, KT, AW, LW, HL, DG, NO, JR, FSTR, SW: data collection and curation, review; AP: data interpretation, review; TH, MR, MMN, IN, UD, TK: funding acquisition, data collection, data curation, review; AJF: data collection, study design, genetic analysis, data interpretation, writing.

FUNDING

This work is part of the German multicentre consortium "Neurobiology of Affective Disorders. A translational perspective on brain structure and function", funded by the German Research Foundation (Deutsche Forschungsgemeinschaft DFG; Forschungsgruppe/Research Unit FOR2107). Principal investigators (PIs) with respective areas of responsibility in the FOR2107 consortium are: Work Package WP1, FOR2107/MACS cohort and brainimaging: Tilo Kircher (speaker FOR2107; DFG grant numbers KI588/14-1, and KI588/14-2, and KI588/20-1, KI588/22-1), Udo Dannlowski (co-speaker FOR2107; DA 1151/5-1, DA 1151/5-2), Axel Krug (KR 3822/5-1, KR 3822/7-2), Igor Nenadić (NE2254/1-2, NE2254/2-1, NE2254/3-1, NE2254/4-1), Carsten Konrad (KO 4291/3-1). WP5, genetics: Marcella Rietschel (RI 908/11-1, RI 908/11-2), Markus Nöthen (NO 246/10-1, NO 246/10-2), Stephanie Witt (WI 3439/3-1, WI 3439/3-2). The study was in part supported by grants from UKGM and Forschungscampus Mittelhessen to Igor Nenadić. Biomedical financial interests or potential conflicts of interest: Tilo Kircher received unrestricted educational grants from Servier, Janssen, Recordati, Aristo, Otsuka, neuraxpharm. Open Access funding enabled and organized by Projekt DEAL.

COMPETING INTERESTS

The authors declare no competing interests.

ETHICS APPROVAL

Patients gave written informed consent to the study protocol. This study was approved by the local Ethics Committee Marburg (AZ:07/14) and Münster (AZ:2014-422-b-S) according to the Declaration of Helsinki.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41398-024-02936-6.

Correspondence and requests for materials should be addressed to Frederike Stein.

Reprints and permission information is available at http://www.nature.com/reprints

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

3.3 Genomics yields biological and phenotypic insights into bipolar disorder

More than 300 authors contributed to this large study of the Bipolar Disorder Working Group of the Psychiatric Genomics Consortium. The first, shared second and shared last authors are O'Connell KS, Koromina M*, van der Veen T*, Boltz T*, David FS*, Yang JMK*, Lin K*, Wang X*, Coleman JRI*, Mitchell BL*, McGrouther CC*, Rangan AV*, Lind PA*, Koch E*, Harder A*, Parker N*, Bendl J*, McQuillin A+, Forstner AJ+, Mullins N+, Di Florio A+, Ophoff RA+, and Andreassen OA+.

*shared second authors
+shared last authors

In this study, we conducted the largest GWAS meta-analysis of BD to date and performed a plethora of downstream analyses to obtain novel insights into the biology and genetic architecture of BD. As part of the analytical team, I played a significant role in the planning of the scientific work, data collection, evaluation and interpretation. Above all, I was responsible for the cell type-specific analysis and, together with Maria Koromina and Toni Boltz, for the quantitative trait loci integration analyses.

The study was published in *Nature* on January 22, 2025:

O'Connell KS, Koromina M, van der Veen T, Boltz T, David FS, … Andreassen OA. Genomics yields biological and phenotypic insights into bipolar disorder. *Nature* 2025; **639**: 968–975

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

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 ($X^2 = 730$, $P < 2.2 \times 10^{-16}$; 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

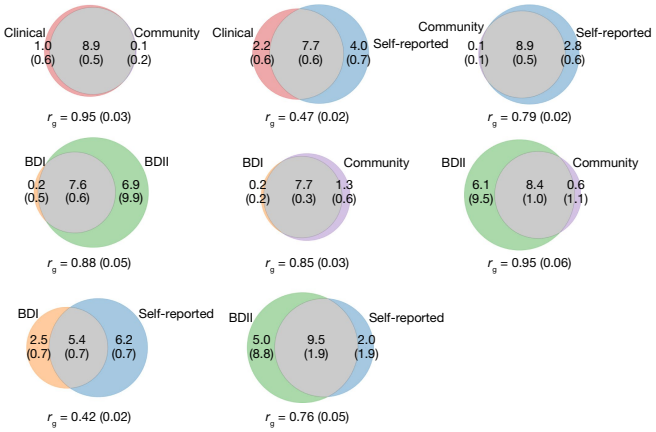


Fig. 1 | Genetic correlation and bivariate MiXeR estimates for the genetic overlap of BD ascertainment and subtypes. Trait-influencing genetic variants shared between each pair (grey) and unique to each trait (colours) are shown. The numbers within the Venn diagrams indicate the estimated number of trait-influencing 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 due to uncontrolled population stratification 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

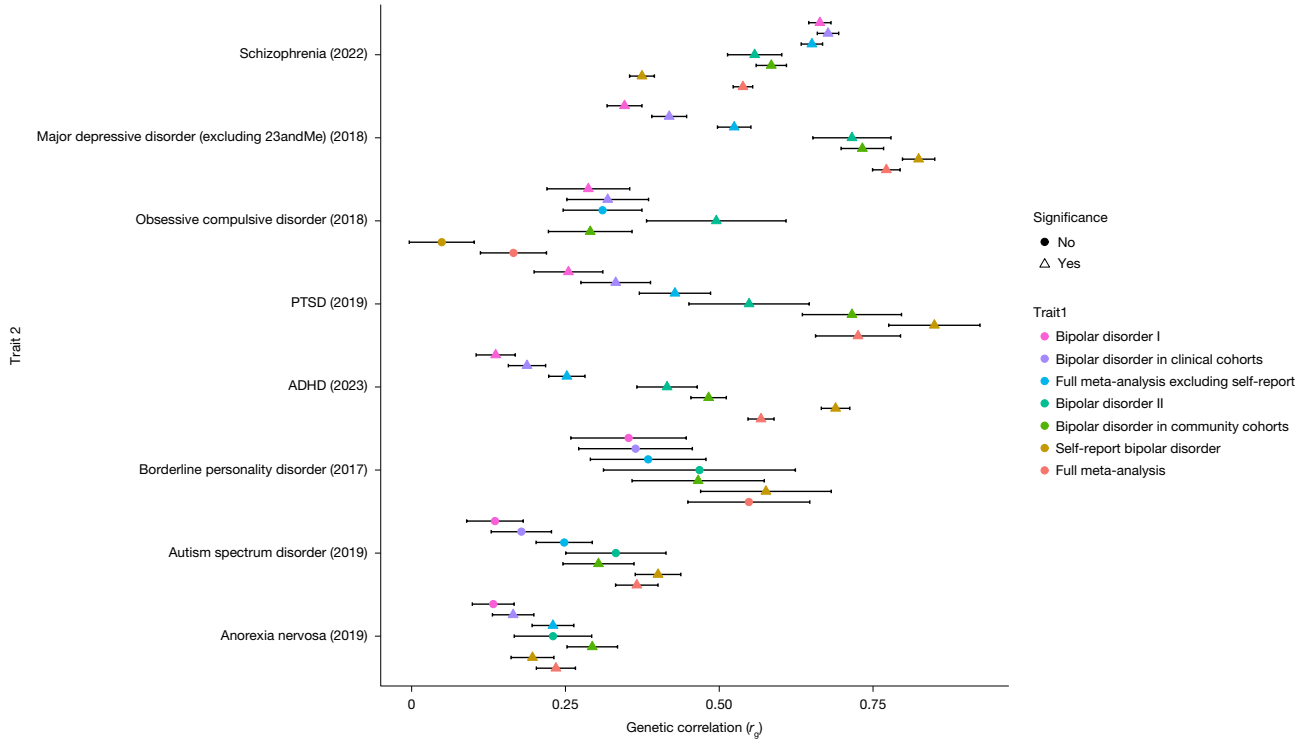


Fig. 2 | Genetic correlations (with standard errors) between BD and other psychiatric disorders. The *y* axis (trait 2) is ordered based on the significance and magnitude of genetic correlation of each trait with BDI. *P* values were calculated from the two-sided *z*-statistics computed by dividing the 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

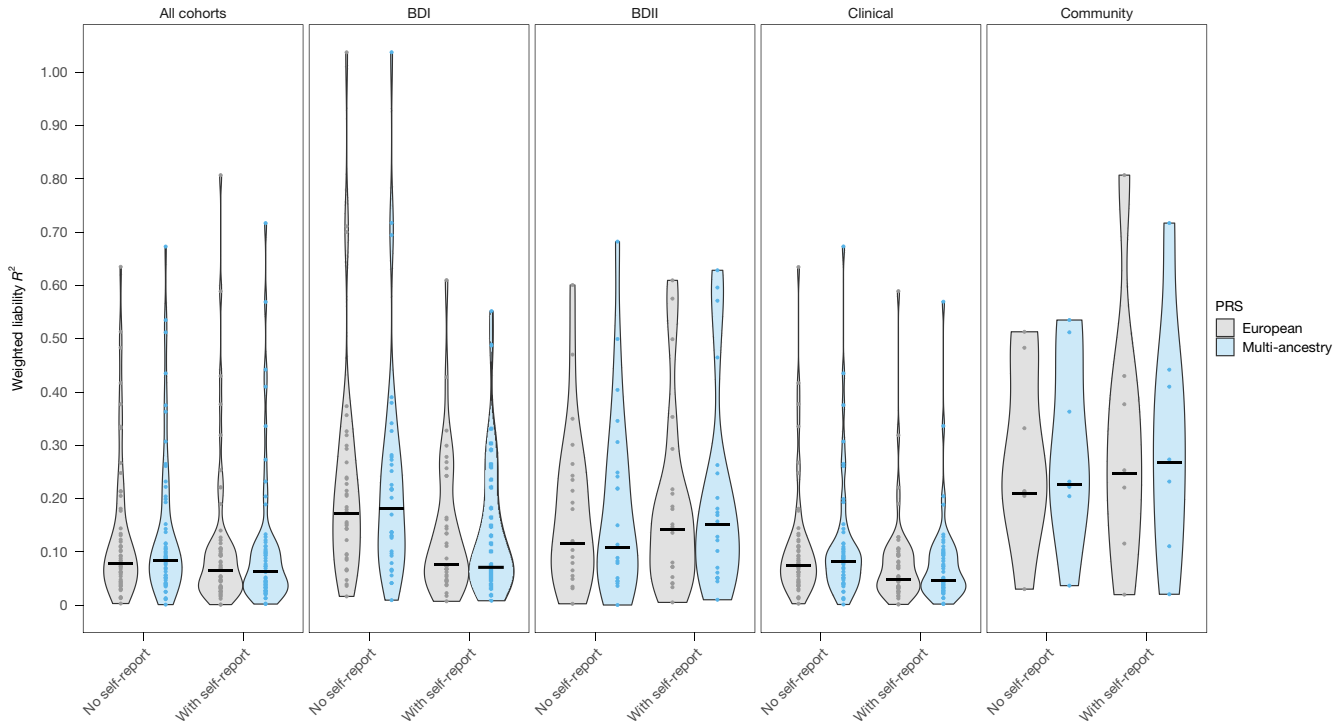


Fig. 3 | Phenotypic variance in BD in EUR cohorts explained by PRSs derived from the multi-ancestry and EUR meta-analyses (with and without self-reported data). Variance explained is presented on the liability scale, assuming a 2% population prevalence of BD. The results (all cohorts) are the median 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.

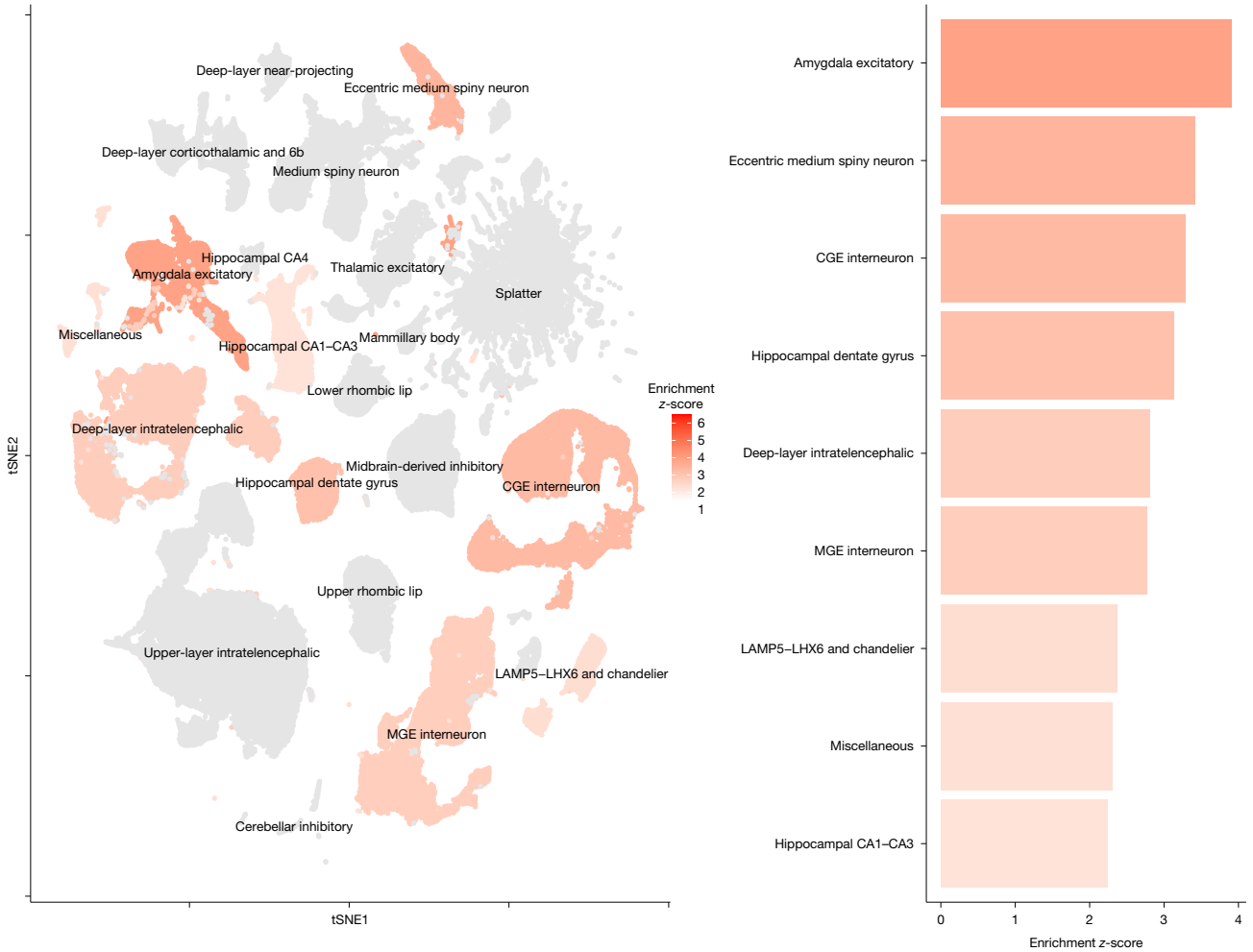


Fig. 4 | Supercluster-level SNP-h^2 enrichment for BD. The *t*-distributed stochastic neighbour embedding (tSNE) plot (naming convention and source of the single-cell data from Siletti et al.[23]; left) is coloured by the enrichment *z*-score. Grey indicates non-significantly enriched superclusters (false 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).

Article

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 Lambtion 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]

Article

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).

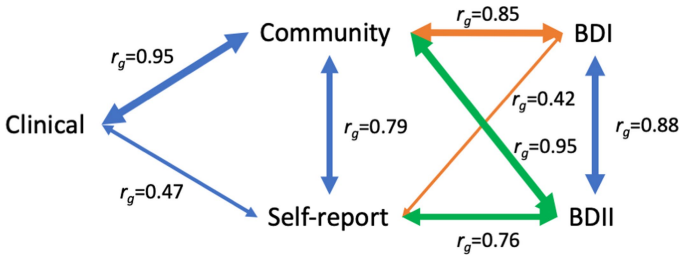
Acknowledgements We thank the participants who donated their time, life experiences and DNA to this research, the clinical and scientific teams that worked with them, and the investigators who comprise the PGC. The PGC has received major funding from the US National Institute of Mental Health (PGC4: 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

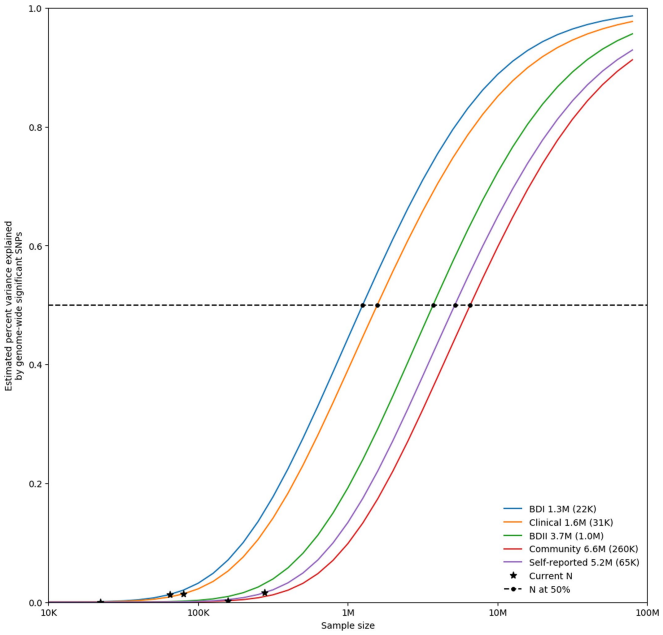
correlation analyses were performed by K.S.O. MiXeR was done by K.S.O. and A. Shadrin. Polygenic association was conducted by T.v.d.V., T.B., B.L.M. and P.A.L. Gene and gene set analyses were done by K.S.O., C.C.M. and A.V.R. Cell-type-specific analyses were performed by F.S.D. Single-nucleus RNA sequencing 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 $30M (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.

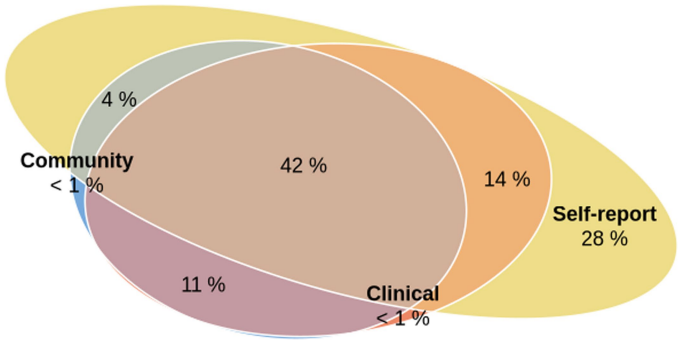
Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41586-024-08468-9.
Correspondence and requests for materials should be addressed to Kevin S. O'Connell or Ole A. Andreassen.
Peer review information *Nature* thanks 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.

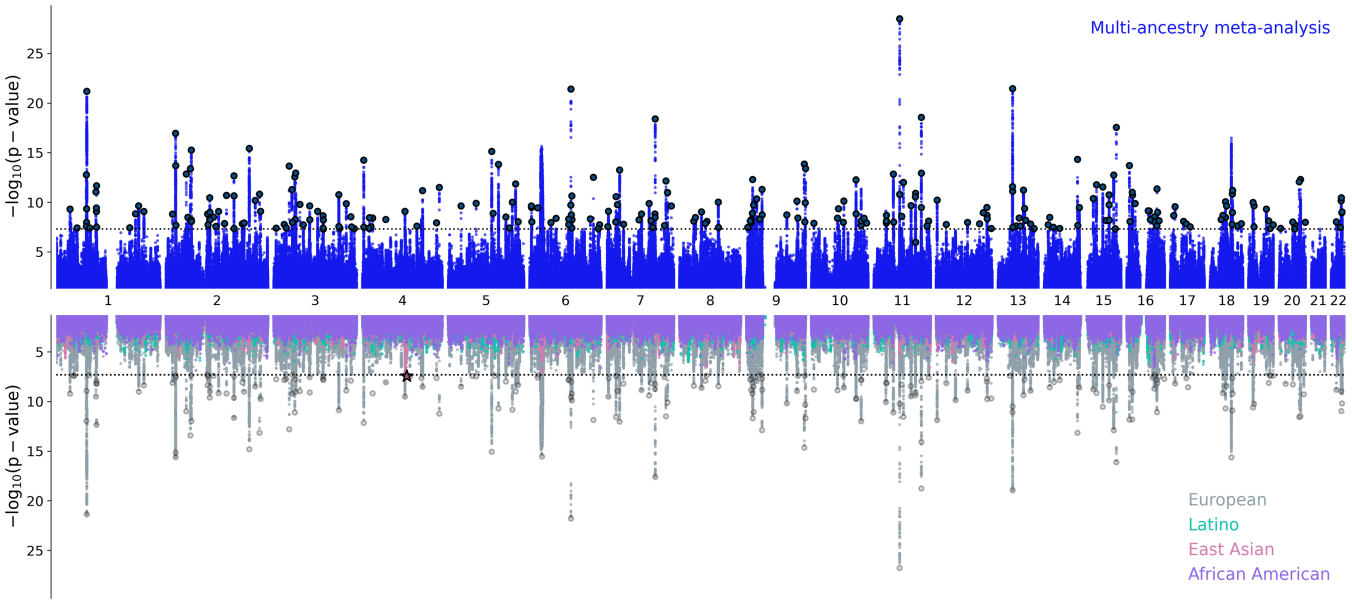


Extended Data Fig. 2 | Univariate MiXeR estimates of the required effective sample size needed to capture 50% of the genetic variance (horizontal dashed line) associated with each BD ascertainment and subtype. N and Sample size refer to the effective sample size. The estimated effective sample size (and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



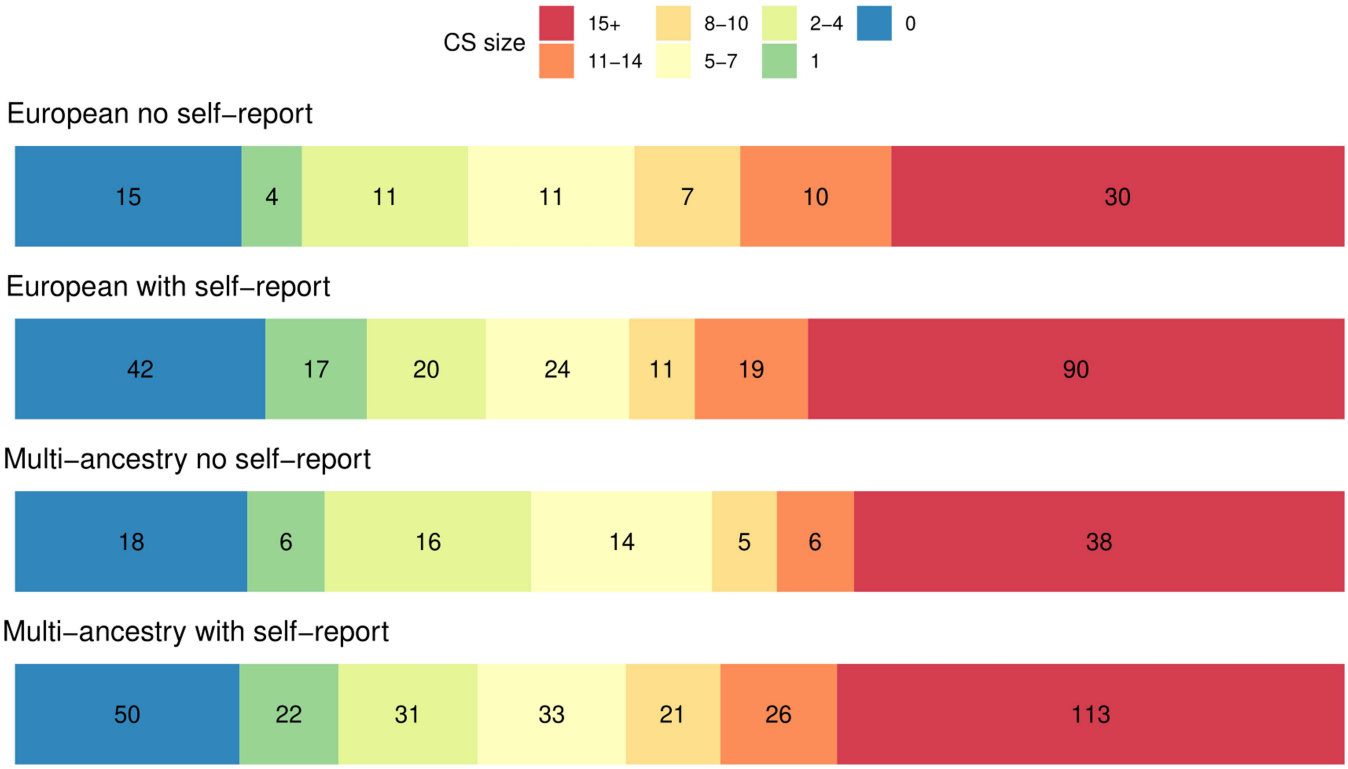
Extended Data Fig. 3 | Trivariate MiXeR estimates for the genetic overlap of BD from Clinical, Community and Self-report samples. The percentages show the proportion of trait-influencing variants within each section of the Venn diagram relative to the sum of all trait-influencing variants across all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

Article

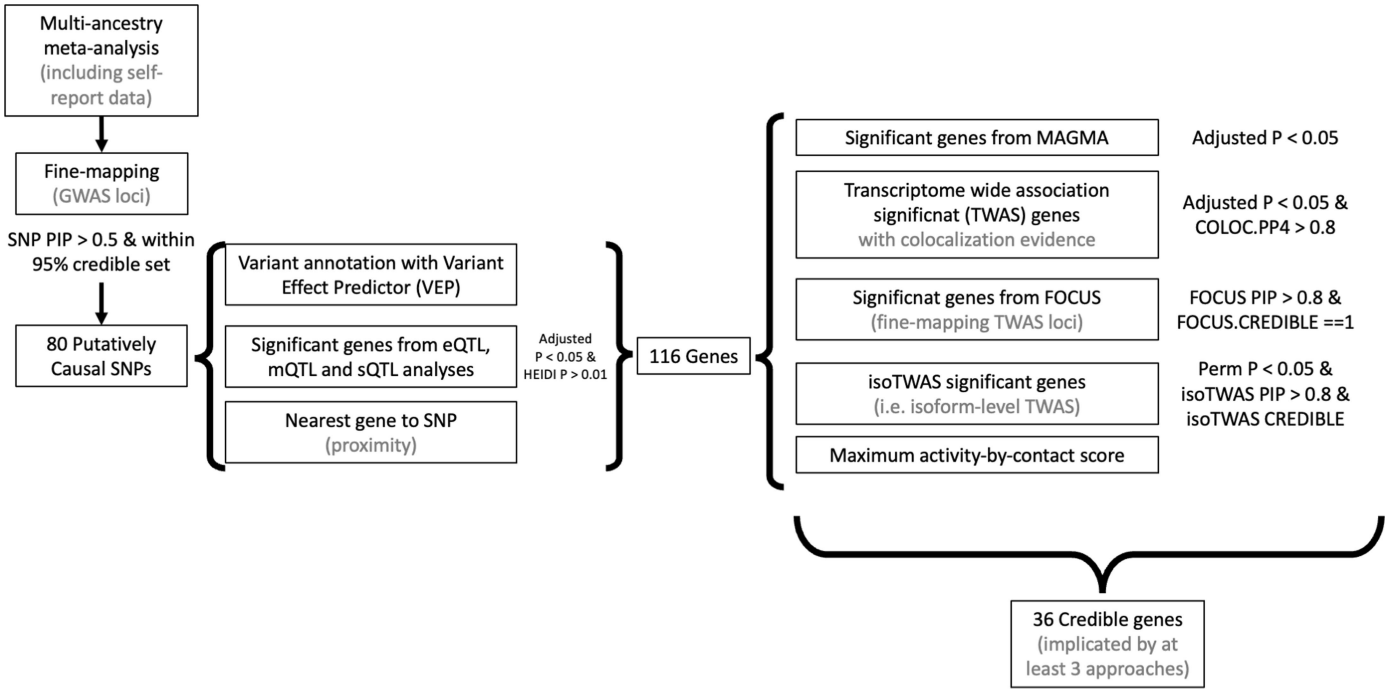


Extended Data Fig. 4 | Miami plot for BD genome-wide meta-analyses, including all cohorts. Upper panel: the multi-ancestry meta-analysis identified 298 genome-wide significant (GWS) loci. Lower panel: porcupine plot showing the results of the Latino (0 GWS loci), African American (0 GWS loci), East Asian (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].

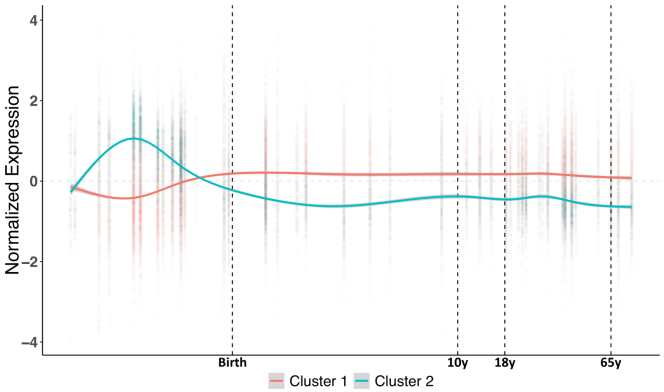


Extended Data Fig. 6 | Number of SNPs within the smallest 95% credible sets (CS) from meta-analysis of European and multi-ancestry meta-analyses when excluding and including self-report data. Colours represent CS of varying size, with blue CS containing 0 SNPs and red CS containing 15+ SNPs. All fine-mapped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

Article



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.

4 Discussion

The following chapter highlights key findings from the three studies included in this doctoral thesis, briefly discusses their significance in relation to existing literature, and outlines directions for future research regarding the role of common genetic variation in affective and psychotic disorders.

In the first and second study we investigated how common genetic variants are associated with symptom dimensions of acute (David *et al.*, 2023) and lifetime (Krug *et al.*, 2024) psychopathology shared across MDD, BD, and SSD. As suggested by initiatives such as the Research Domain Criteria (RDoC) project (Cuthbert, 2014), we adopted a transdiagnostic perspective that focuses on dimensions of behavior observed across diagnostic boundaries rather than on existing diagnostic groupings. In the PRS analyses of the transdiagnostic sample we found associations of individual symptom dimensions with the polygenic liability for BD and SCZ, respectively, e.g., between the PRS for SCZ and the dimension of paranoid-hallucinatory syndrome (David *et al.*, 2023) and between the PRS for BD and the dimension of mania (Krug *et al.*, 2024). This is in line with the substantial genetic overlaps between the disorders (Cross-Disorder Group of the Psychiatric Genomics Consortium, 2019) and indicates that genetic variants identified by disorder-specific case-control GWAS of affective and psychotic disorders might comprise variants related to individual symptom dimensions shared across diagnoses. Similar findings were obtained by studies of polygenic profiles within BD (Allardyce *et al.*, 2023; Richards *et al.*, 2022), supporting this view.

While the exploratory GWAS in David *et al.* (2023) did not reveal any genome-wide significant associations, the exploratory GWAS in Krug *et al.* (2024) yielded two genome-wide significant findings for the lifetime symptom dimensions of mania and depression. This could be interpreted as lifetime measures being more suitable than acute measures of psychopathology for the investigation of genetic correlates. However, neither rs10062519, which was associated with lifetime mania, nor rs11131155, associated with lifetime depression, showed a genome-wide significant disease association (i.e., $p < 5 \times 10^{-8}$) in the latest large-scale GWAS meta-analyses of BD (O'Connell *et al.*, 2025) and major depression (Adams *et al.*, 2025), respectively. Therefore, given the

sample size of our exploratory GWAS, the findings should be interpreted with caution until replication is achieved by other studies.

In contrast to the first two studies adopting a transdiagnostic approach, the third study included in this thesis consists of a disorder-specific investigation of common genetic variants associated with BD disease status (O'Connell *et al.*, 2025). The study represents the largest GWAS meta-analysis of BD to date, encompassing over 158,000 cases and 2.7 million controls from four different ancestral groups. 298 genome-wide significant loci associated with BD were identified. In addition, the results of the meta-analysis fueled a wide variety of downstream analyses, such as quantitative trait loci integration and cell type enrichment analyses. The massive increase in sample size compared to the previous GWAS meta-analysis of BD by the PGC (Mullins *et al.*, 2021; more than 41,000 cases and 371,000 controls) was largely driven by the inclusion of samples from 23andMe, Inc. (more than 90,000 cases and 1.9 million controls), in which case status was ascertained via a single-item self-report question. In addition, over 22,000 cases and 441,000 controls were newly added from community samples, in which case ascertainment was based on medical records, health registries, and questionnaire data. Notably, the SNP-based heritability (h^2_{SNP}) of BD in the community samples (h^2_{SNP} = 0.05) and self-report samples (h^2_{SNP} = 0.08) was substantially lower than in the samples with clinical case ascertainment (h^2_{SNP} = 0.22). This observed difference in the genetic architecture of BD depending on the type of case ascertainment may reflect differences in the phenotypic composition, e.g., in regard to BD subtypes and disease severity (O'Connell *et al.*, 2025).

The three studies included in this thesis illustrate the trade-off between sample size and the depth and quality of phenotypic data in genetic studies of affective and psychotic disorders. The deep phenotyping data of the FOR2107 cohort made it possible to construct transdiagnostic dimensional phenotypes of psychopathology and, in combination with the genotyping and brain imaging data, enabled the investigation of their genetic and brain morphometric correlates in David *et al.* (2023) and Krug *et al.* (2024). However, this came at the cost of a limited sample size, impeding the detection of associated common variants that individually have small effects on highly polygenic complex conditions. The GWAS were underpowered and thus only exploratory, even though well-constructed quantitative phenotypes are known to increase statistical power

for the investigation of common variants in common diseases compared to binary phenotypes (Waszczuk *et al.*, 2023). Conversely, the minimal phenotyping strategy employed in O'Connell *et al.* (2025) resulted in a large sample size and successful discovery of novel BD-associated loci, but the depth of available phenotype data was not sufficient to resolve the observed heterogeneity. Minimal phenotyping can also come at the cost of low disease specificity of GWAS signals, as demonstrated by Cai *et al.* (2020) in the context of MDD. However, the lower heritability and higher phenotypic heterogeneity compared to other affective and psychotic disorders likely exacerbates the consequences of minimal phenotyping in MDD. For SSD, Woolway *et al.* (2024) have recently demonstrated the validity of self-reported diagnoses in genomic research. Together this indicates that the best phenotyping strategy for genetic discovery may differ between different affective and psychotic disorders and phenotypes.

One apparent limitation of both David *et al.* (2023) and Krug *et al.* (2024) is the restriction of the study sample to individuals of European ancestry. Even in O'Connell *et al.* (2025), where we investigated individuals from four different ancestral groups, the majority of individuals were of European ancestry. While ancestry-specific research designs are helpful to reduce the effect of population stratification on common variant associations, they limit the transferability and generalizability of findings, with important implications on equity in health care (Martin *et al.*, 2019). This issue has received increasing attention in the research community throughout the past years, and it should be emphasized that the global efforts towards the recruitment of more ancestrally diverse samples in psychiatric genetics research and the development of suitable methods for multi-ancestry analyses and diversity-aware modelling must continue (Peterson *et al.*, 2019).

Overall, the studies included in this dissertation enhance our understanding of the complex relationships of common genetic variants with symptom dimensions and disease status in affective and psychotic disorders. For future research, a number of directions may prove fruitful. First, standardized digital phenotyping, including ecological momentary assessments, may be used for a more cost efficient collection of detailed phenotype data (Montag, Quintana, 2023), potentially mitigating the trade-off between sample size and phenotyping depth. This may resolve some of the observed genetic heterogeneity due to the ability to define more homogeneous patient subgroups and to construct more specific

quantitative phenotypes. Second, in light of the ever increasing number of GWAS findings, the multimodal integration with rare variant data and other molecular features such as transcriptomics, proteomics, and metabolomics is expected to be of increasing relevance. The multimodal integration will strengthen the identification of genes, pathways, and cell types that are involved in the pathophysiological mechanisms of affective and psychotic disorders (Bruner, Grant, 2024). Lastly, given the growing abundance of various types of data, the application of artificial intelligence, in particular multimodal deep learning models, may lead to a breakthrough in the identification of etiologically defined groups of patients exhibiting distinct genetic and molecular profiles (Chen *et al.*, 2022). Together, these opportunities could be a significant step towards the goal of psychiatric research, which ultimately consists of a more effective prevention and treatment of mental disorders.

4.1 References

Adams MJ, Streit F, Meng X, Awasthi S, Adey BN, … McIntosh AM. Trans-ancestry genome-wide study of depression identifies 697 associations implicating cell types and pharmacotherapies. *Cell* 2025; **188**: 640-652.e9

Allardyce J, Cardno AG, Gordon-Smith K, Jones L, Di Florio A, … O'Donovan MC. Specificity of polygenic signatures across symptom dimensions in bipolar disorder: an analysis of UK Bipolar Disorder Research Network data. *Lancet Psychiatry* 2023; **10**: 623–631

Bruner WS, Grant SFA. Translation of genome-wide association study: from genomic signals to biological insights. *Front Genet* 2024; **15**: 1375481

Cai N, Revez JA, Adams MJ, Andlauer TFM, Breen G, … Flint J. Minimal phenotyping yields genome-wide association signals of low specificity for major depression. *Nat Genet* 2020; **52**: 437–447

Chen ZS, Kulkarni PP, Galatzer-Levy IR, Bigio B, Nasca C, Zhang Y. Modern views of machine learning for precision psychiatry. *Patterns (N Y)* 2022; **3**: 100602

Cross-Disorder Group of the Psychiatric Genomics Consortium. Genomic Relationships, Novel Loci, and Pleiotropic Mechanisms across Eight Psychiatric Disorders. *Cell* 2019; **179**: 1469-1482.e11

Cuthbert BN. The RDoC framework: facilitating transition from ICD/DSM to dimensional approaches that integrate neuroscience and psychopathology. *World Psychiatry Official Journal of the World Psychiatric Association (WPA)* 2014; **13**: 28–35

David FS, Stein F, Andlauer TF, Streit F, Witt SH, … Forstner AJ. Genetic contributions to transdiagnostic symptom dimensions in patients with major depressive disorder, bipolar disorder, and schizophrenia spectrum disorders. *Schizophr Res* 2023; **252**: 161–171

Krug A, Stein F, David FS, Schmitt S, Brosch K, … Forstner AJ. Factor analysis of lifetime psychopathology and its brain morphometric and genetic correlates in a transdiagnostic sample. *Transl Psychiatry* 2024; **14**: 235

Martin AR, Kanai M, Kamatani Y, Okada Y, Neale BM, Daly MJ. Clinical use of current polygenic risk scores may exacerbate health disparities. *Nat Genet* 2019; **51**: 584–591

Montag C, Quintana DS. Digital phenotyping in molecular psychiatry-a missed opportunity? *Mol Psychiatry* 2023; **28**: 6–9

Mullins N, Forstner AJ, O'Connell KS, Coombes B, Coleman JRI, … Andreassen OA. Genome-wide association study of more than 40,000 bipolar disorder cases provides new insights into the underlying biology. *Nat Genet* 2021; **53**: 817–829

O'Connell KS, Koromina M, van der Veen T, Boltz T, David FS, … Andreassen OA. Genomics yields biological and phenotypic insights into bipolar disorder. *Nature* 2025; **639**: 968–975

Peterson RE, Kuchenbaecker K, Walters RK, Chen C-Y, Popejoy AB, … Duncan LE. Genome-wide Association Studies in Ancestrally Diverse Populations: Opportunities, Methods, Pitfalls, and Recommendations. *Cell* 2019; **179**: 589–603

Richards AL, Cardno A, Harold G, Craddock NJ, Di Florio A, … O'Donovan MC. Genetic Liabilities Differentiating Bipolar Disorder, Schizophrenia, and Major Depressive Disorder, and Phenotypic Heterogeneity in Bipolar Disorder. *JAMA Psychiatry* 2022; **79**: 1032–1039

Waszczuk MA, Jonas KG, Bornovalova M, Breen G, Bulik CM, … Waldman ID. Dimensional and transdiagnostic phenotypes in psychiatric genome-wide association studies. *Mol Psychiatry* 2023; **28**: 4943–4953

Woolway GE, Legge SE, Lynham AJ, Smart SE, Hubbard L, … Walters JTR. Assessing the validity of a self-reported clinical diagnosis of schizophrenia. *Schizophrenia (Heidelb)* 2024; **10**: 99

5 Acknowledgement

I would like to express my sincere gratitude to a number of people who were involved in this work or supported me throughout my time as a PhD student.

I am immensely grateful to my doctoral supervisor Prof. Dr. Andreas Forstner for his guidance and mentorship since the very beginning of my doctorate. He was involved in all of my projects, and his thoughts and insights have contributed significantly to the success of this thesis. Thank you for all the scientific and personal conversations, for always valuing my opinion, and for your encouragement and support to pursue my own ideas. I extend my appreciation to Prof. Dr. Peter Krawitz, who took on the role of the second reviewer, and to Prof. Dr. Markus Nöthen and Prof. Dr. Johannes Schumacher as members of my thesis committee. Thank you for your time and consideration.

I would like to thank my colleagues from the German FOR2107 consortium in Marburg, Münster, and Mannheim. Big thank you to Dr. Frederike Stein for her continuous support with everything related to the FOR2107 database and for the excellent collaboration in several projects. I would like to thank the members of the PGC BD Working Group and in particular Dr. Maria Koromina for the great discussions and fruitful collaboration. I am grateful to all former and current colleagues at the Institute of Human Genetics in Bonn, in particular to the members of my working group, for their feedback and support. A special thanks goes to Anna Siewert, Charlotte Pahnke, and Dr. Carina Mathey for making the daily adventures of PhD life a worthwhile experience.

My warmest gratitude goes to my family, especially my mom, and my friends for their many different ways of supporting me, for their continuous encouragement, and for the welcome distractions when I needed a break from work. To Milo, for always listening to my rants and for being my first-level emotional support person.

Finally, this work would not have been possible without the many people who participated in the studies, who donated their time and data. Without direct contact to the participants it is easy to forget that behind every sample ID is a real person with a real life. I am deeply indebted to all study participants, and in particular to those who struggle with the burden of the mental health conditions that are the focus of my thesis. This work is for you.