

Genome-wide analyses identify 30 loci associated with obsessive–compulsive disorder

Received: 13 March 2024

Accepted: 7 April 2025

Published online: 13 May 2025

 Check for updates

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

Obsessive–compulsive disorder (OCD) affects ~1% of children and adults and is partly caused by genetic factors. We conducted a genome-wide association study (GWAS) meta-analysis combining 53,660 OCD cases and 2,044,417 controls and identified 30 independent genome-wide significant loci. Gene-based approaches identified 249 potential effector genes for OCD, with 25 of these classified as the most likely causal candidates, including *WDR6*, *DALRD3* and *CTNND1* and multiple genes in the major histocompatibility complex (MHC) region. We estimated that ~11,500 genetic variants explained 90% of OCD genetic heritability. OCD genetic risk was associated with excitatory neurons in the hippocampus and the cortex, along with D₁ and D₂ type dopamine receptor-containing medium spiny neurons. OCD genetic risk was shared with 65 of 112 additional phenotypes, including all the psychiatric disorders we examined. In particular, OCD shared genetic risk with anxiety, depression, anorexia nervosa and Tourette syndrome and was negatively associated with inflammatory bowel diseases, educational attainment and body mass index.

OCD is a chronic psychiatric disorder that affects 1–3% of the population¹ and is characterized by obsessions and compulsions that vary in type and severity and over time. OCD is responsible for profound personal and societal costs², including increased risk of suicide³ and overall mortality⁴. OCD is moderately heritable; twin-based heritability estimates range between 27% and 47% in adults and between 45% and 65% in children^{5–8}, with SNP-based heritability estimates between 28% and 37%^{9–11}.

Two earlier OCD GWAS meta-analyses, both containing a subset of the data included in this analysis^{12,13}, showed SNP-based heritabilities of 8.5% (assuming a 3% population prevalence) and 16% (assuming a 2% population prevalence). The first GWAS ($n_{\text{cases}} = 14,140$, $n_{\text{controls}} = 562,117$)¹² found one genome-wide significant locus associated with OCD, while the second ($n_{\text{cases}} = 37,015$, $n_{\text{controls}} = 948,616$)¹³ identified 15 independent genome-wide significant loci. As with other complex traits, increased sample sizes are needed for a more comprehensive understanding of the underlying genetic etiology of OCD and its genetic relationships with related disorders.

The current study combines data from the two unpublished OCD GWASs described above and includes additional cohorts (~9,000 cases). This results in one of the largest and most well-powered GWAS of OCD so far, with a ~20-fold increase of OCD cases compared to the previously published OCD GWASs¹⁰. Based on the results from the meta-analysis, we conducted secondary analyses, including positional and functional fine-mapping of SNPs and genes, structural equation modeling to examine possible genetic differences in sample ascertainment across cohorts, protein and transcriptome-wide association analyses, single-cell enrichment and genetic correlations with other traits (Supplementary Fig. 1). Our results provide more detailed insight into the genetic underpinnings and biology of OCD.

Results

GWAS meta-analysis identifies 30 genome-wide significant loci

We conducted a GWAS meta-analysis of 28 OCD case–control cohorts of European ancestry, comprising 53,660 cases and 2,044,417 controls (effective sample size, ~210,000 individuals). Ascertainment of cases

✉ e-mail: nora.strom@hu-berlin.de; manuel.mattheisen@dal.ca

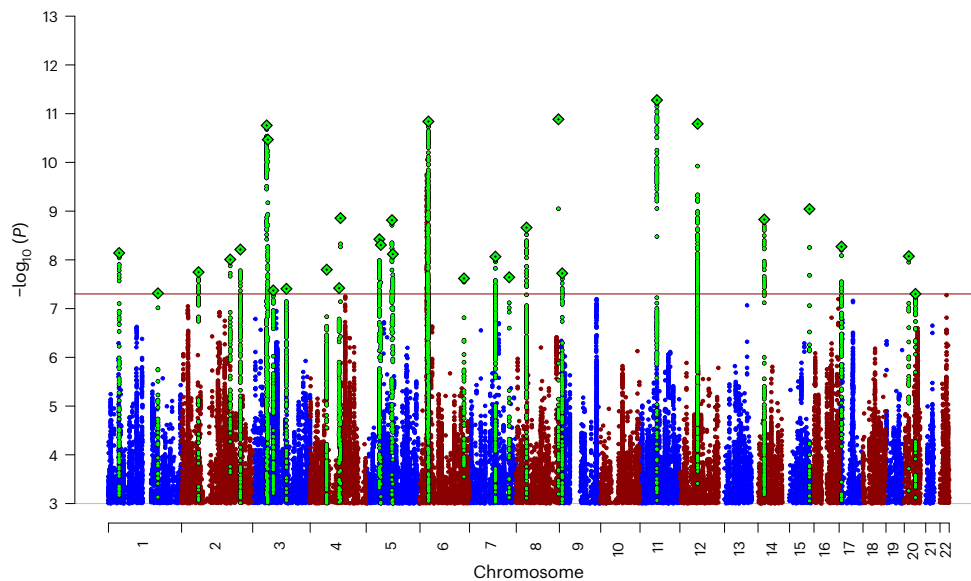


Fig. 1 | Manhattan plot of OCD GWAS meta-analysis. The y axis represents $-\log_{10}(P)$ values (two sided, not adjusted for multiple testing) for the association of variants with OCD using an inverse-variance-weighted fixed-effects model ($n_{\text{cases}} = 53,660$ and $n_{\text{controls}} = 2,044,417$). The x axis shows chromosomes 1–22.

The horizontal red line represents the threshold for genome-wide significance ($P = 5 \times 10^{-8}$). Index variants of genome-wide significant loci are highlighted as green diamonds.

varied across cohorts: OCD diagnosis was determined (1) by a health-care professional in a clinical setting (18 cohorts, $n = 9,089$ cases), (2) from health records or biobanks (seven cohorts, $n = 9,138$ cases), (3) in a clinical setting or from health records with the additional characteristic that all OCD cases were primarily collected for another psychiatric disorder (three cohorts, $n = 5,266$ cases) or (4) by self-reported diagnosis in a consumer-based setting (23andMe, Inc., $n = 30,167$ cases). Cohort details, including phenotypic assessment, quality control and individual cohort GWAS analyses, are described in Supplementary Note 2 and Supplementary Table 1. We identified 30 independent (defined in Supplementary Note 3) loci among the 1,672 SNPs that exceeded the genome-wide threshold for significance ($P < 5 \times 10^{-8}$; Manhattan plot in Fig. 1, regional association plots and forest plots in Supplementary Figs. 2–31 and a list of all independent genome-wide significant SNPs in Table 1 with additional details in Supplementary Tables 2 and 3). The independence of the 30 lead SNPs was subsequently validated using conditional and joint analysis (GCTA-COJO)¹⁴ (Supplementary Table 4). Analysis of the X chromosome, conducted in a subset of the data for which this information was available (23andMe), yielded no significant associations (Supplementary Note 4 and Supplementary Fig. 37e). Of the 15 genome-wide significant loci previously reported in preprints^{12,13}, 13 were genome-wide significant in the current GWAS, with the remaining two showing suggestive significance ($P = 5.23 \times 10^{-8}$ and $P = 2.2 \times 10^{-7}$; Supplementary Table 5). Using MiXeR¹⁵, we estimated that approximately 11,500 (standard error of the effect estimate (s.e.) = 607) causal variants account for 90% of the OCD SNP-based heritability.

No statistically significant heterogeneity was observed across individual cohorts for the 30 genome-wide significant loci, as assessed with Cochran's Q -test (Supplementary Fig. 32), the I^2 statistic and the genomic structural equation modeling (GenomicSEM) Q_{SNP} statistic¹⁶ (Supplementary Table 2). Genome-wide analyses of samples grouped by clinical, comorbid, biobank and 23andMe information (Supplementary Table 3 and Supplementary Figs. 33–37) showed evidence that sample ascertainment impacted results at a genome-wide scale, although not beyond what is observed with closely related psychiatric disorders^{17,18}. We observed moderate to high genetic correlations across the subgroups (between 0.63, s.e. = 0.11 for biobanks and comorbid

information and 0.92, s.e. = 0.07 for 23andMe and comorbid information; Supplementary Table 7) and a satisfactory fit for a one-factor GenomicSEM model (Supplementary Table 8 and Supplementary Fig. 39). A common factor GWAS based on the one-factor GenomicSEM model resulted in 20 significant loci, all of which were also significant in the primary GWAS (Supplementary Table 8 and Supplementary Fig. 40; analysis details in Supplementary Note 5). SNP heritability (assuming a 1% population prevalence) was 6.7% (s.e. = 0.3%), with slightly higher estimates for the clinical ($h^2_{\text{SNP}} = 16.4\%$, s.e. = 1.5%) and comorbid ($h^2_{\text{SNP}} = 13.3\%$, s.e. = 1.7%) subgroups (Supplementary Table 1).

Gene-based findings

We prioritized putative risk genes for OCD using six positional and functional QTL gene-based mapping approaches. Positional mapping was performed with mBAT-combo¹⁹. Functional expression quantitative trait locus (eQTL) mapping was performed with transcriptome-wide association study (TWAS)²⁰, using PsychENCODE gene expression weights²¹, and summary-based Mendelian randomization (SMR)²² using the whole-blood eQTLGen²³ and MetaBrain²⁴ datasets. Functional protein QTL mapping was done using a protein-wide association study (PWAS) of human brain protein expression panels²⁵. Finally, we used the psychiatric omnilocus prioritization score (PsyOPS)²⁶, which combines positional mapping with biological annotations, to further prioritize risk genes within genome-wide significant loci. We identified 207 significant genes (Bonferroni correction, $P < 2.67 \times 10^{-6}$) with mBAT-combo and 24 genes using TWAS ($P < 4.76 \times 10^{-6}$), 14 of which were conditionally independent. The SMR-eQTLGen analysis identified 39 significant risk genes ($P < 4.28 \times 10^{-6}$), and the SMR-MetaBrain analysis identified 14 risk genes ($P < 9.23 \times 10^{-6}$). The PWAS identified three significant genes ($P < 3.39 \times 10^{-5}$), while PsyOPS prioritized 29 genes. In total, 251 genes were significantly associated with OCD through at least one gene-based approach, and 48 were implicated by at least two methods (Methods, Supplementary Note 7 and Supplementary Tables 9–14).

From the 48 genes implicated by at least two approaches, we prioritized likely causal genes for OCD using colocalization (TWAS-COLOC)^{27,28} and SMR-heterogeneity in dependent instruments (SMR-HEIDI)²² tests. Colocalization was used to identify significant TWAS associations for which the underlying GWAS and eQTL summary

Table 1 | Genome-wide-significant loci associated with OCD

SNP	Position	P value	OR	s.e.	A ₁ /A ₂	FRQCA	FRQCO	INFO	n genes	Selection of other significant traits
rs78587207 ^a	11q12.1	5.28 × 10 ⁻¹²	0.9522	0.0071	T/G	0.681	0.677	0.987	11	SCZ, well-being, neuroticism, educational attainment
rs13262595	8q24.3	1.31 × 10 ⁻¹¹	0.9566	0.0066	G/A	0.552	0.561	0.989	2	SCZ, well-being, neuroticism, educational attainment
rs4990036 ^a	6p21.33	1.45 × 10 ⁻¹¹	0.9299	0.0108	T/C	0.119	0.123	0.985	118	SCZ, depression, blood cell count, lung function
rs10877425 ^a	12q14.1	1.62 × 10 ⁻¹¹	0.9526	0.0072	G/A	0.513	0.516	0.985	0	–
rs7626445 ^a	3p21.31	1.74 × 10 ⁻¹¹	0.9551	0.0068	T/C	0.647	0.654	0.994	32	Neuroticism, smoking, blood cell count, height
rs2564930 ^a	3p21.1	3.41 × 10 ⁻¹¹	0.9546	0.007	T/C	0.339	0.345	0.988	12	SCZ, neuroticism, blood cell count, BMI
rs4702 ^a	15q26.1	9.07 × 10 ⁻¹⁰	1.0414	0.0066	G/A	0.455	0.449	0.984	5	SCZ, BP, MDD, risk-taking behavior
rs35518360 ^a	4q24	1.39 × 10 ⁻⁹	1.0757	0.0121	T/A	0.0828	0.0756	0.947	1	SCZ, neuroticism, educational attainment, BMI
rs4904738	14q21.1	1.48 × 10 ⁻⁹	0.9605	0.0067	T/C	0.558	0.553	0.984	1	MDD
rs254779	5q14.3	1.53 × 10 ⁻⁹	0.9606	0.0067	T/C	0.419	0.421	0.988	4	Educational attainment, cognition, fat mass
rs2198140 ^a	8p12	2.18 × 10 ⁻⁹	0.9590	0.007	T/C	0.496	0.513	0.979	1	Cognition
rs12516488 ^a	5p12	3.79 × 10 ⁻⁹	1.0531	0.0088	G/A	0.825	0.826	0.994	1	Neuroticism, age at first birth, age at first sexual intercourse
rs3899258 ^a	5q11.1	4.94 × 10 ⁻⁹	1.0509	0.0085	G/A	0.782	0.792	0.989	2	–
rs3027160	17p13.1	5.35 × 10 ⁻⁹	1.0497	0.0083	T/C	0.775	0.782	0.996	19	Sleep, height
rs203768	2q33.1	6.14 × 10 ⁻⁹	0.9513	0.0086	T/C	0.824	0.817	0.992	5	SCZ
rs11263940	1p34.3	7.23 × 10 ⁻⁹	0.9578	0.0074	T/C	0.689	0.69	0.991	0	Neuroticism, well-being
rs67839857 ^a	5q14.3	7.63 × 10 ⁻⁹	1.0423	0.0072	G/A	0.692	0.691	0.994	0	–
rs1555466	20p11.23	8.42 × 10 ⁻⁹	1.0490	0.0083	T/C	0.218	0.224	0.996	0	Ease of skin tanning
rs9886111	7q21.13	8.59 × 10 ⁻⁹	0.9598	0.0071	G/C	0.701	0.711	0.992	2	–
rs9287859	2q24.3	9.83 × 10 ⁻⁹	0.9595	0.0072	G/A	0.39	0.39	0.994	1	–
rs2087319	4q12	1.59 × 10 ⁻⁸	0.9579	0.0076	C/A	0.74	0.744	0.968	6	Height, blood pressure
rs11125759 ^a	2p16.1	1.79 × 10 ⁻⁸	0.9690	0.0071	G/A	0.569	0.556	0.991	1	BMI, sleep
rs6474628	9p23	1.89 × 10 ⁻⁸	1.0380	0.0066	T/G	0.579	0.585	0.999	0	–
rs11768238	7q33	2.28 × 10 ⁻⁸	0.9601	0.0073	G/A	0.661	0.661	0.998	2	Educational attainment, age at first sexual intercourse
rs9479138	6q25.1	2.41 × 10 ⁻⁸	1.0397	0.007	T/G	0.339	0.34	0.975	1	Educational attainment, age at first sexual intercourse, age at first birth, lung function
rs1567288	4q22.3	3.80 × 10 ⁻⁸	0.9643	0.0066	G/A	0.548	0.55	0.981	1	–
rs4831130	3q13.31	3.93 × 10 ⁻⁸	1.0427	0.0076	T/G	0.753	0.74	0.984	2	–
rs17718444	3p13	4.25 × 10 ⁻⁸	0.9622	0.007	T/C	0.323	0.32	0.984	2	Educational attainment, lung function, use of sun/UV protection
rs6660196 ^a	1q24.1	4.86 × 10 ⁻⁸	1.0403	0.0072	T/G	0.638	0.642	0.989	1	Blood cell count
rs4931	20q13.12	5.00 × 10 ⁻⁸	0.9609	0.0073	C/A	0.278	0.279	0.993	5	Blood cell count, height, BMI

Shown are the lead SNP, the chromosome, base pair position on the genome, *P* value, effect estimate as an odds ratio (OR), s.e., effect allele and non-effect allele (A₁ and A₂), frequency of A₁ in cases (FRQCA) and in controls (FRQCO), imputation quality score (INFO), number of genes in a region of 6.5 kb around the SNP (*n* genes) and a curated list of phenotypes that also showed a genome-wide significant association with this SNP (in one or more of the following four databases: CAUSALdb⁵⁰, GenomeAtlas⁵², the NHGRI-EBI GWAS Catalog⁵¹, the IEU Open GWAS project⁵³). If fewer than four traits are significant across all four databases, all four traits are shown. If more than five traits are significant across the databases, neuropsychiatric traits are prioritized (closely related traits are summarized into one trait category). For a full list of associations in the four databases, see Supplementary Table 18a–d. A more detailed list of the significant loci can be found in Supplementary Table 2. Abbreviations in the last column are SCZ, schizophrenia; BP, bipolar disorder; MDD, major depressive disorder; UV, ultraviolet. ^aPreviously identified GWAS hits for OCD (or SNPs in high linkage disequilibrium with a previously identified SNP).

statistics are likely to share a single causal variant. Similarly, HEIDI was used to select SMR associations for which the same causal variant affects gene expression and trait variation. Of the 48 genes implicated by at least two gene-based tests, 25 were also significant in either the TWAS-COLOC or the SMR-HEIDI tests, suggesting causality (Fig. 2a). Only 2 of these 25 genes were prioritized by both TWAS-COLOC and SMR-HEIDI: *WDR6* (WD repeat domain 6) and *DALRD3* (DALR anticodon binding domain-containing 3). Another gene of interest, *CTNND1* (catenin δ1), was implicated by three of our five approaches (multivariate set-based association test (mBAT-combo), TWAS, PWAS) and showed evidence for colocalization. Only three genes were implicated in the PWAS; of these, *CTNND1* was the only gene also implicated

in the TWAS. In the PWAS, downregulation of *CTNND1* protein expression in the human dorsolateral prefrontal cortex (dlPFC) was significantly associated with OCD risk ($Z = -4.49$, $P = 7.11 \times 10^{-6}$; Supplementary Table 13), consistent with the downregulation of *CTNND1* gene expression in the prefrontal cortex seen in the TWAS ($Z = -6.86$, $P = 6.90 \times 10^{-12}$; Supplementary Table 10). For a discussion of the overlap between the gene findings with rare coding variants in OCD, see Supplementary Table 6 and Supplementary Note 7.

Tissue and cell type enrichment analysis

After mapping significantly associated SNPs from the GWAS meta-analysis to likely causal genes, we explored which tissues or cell

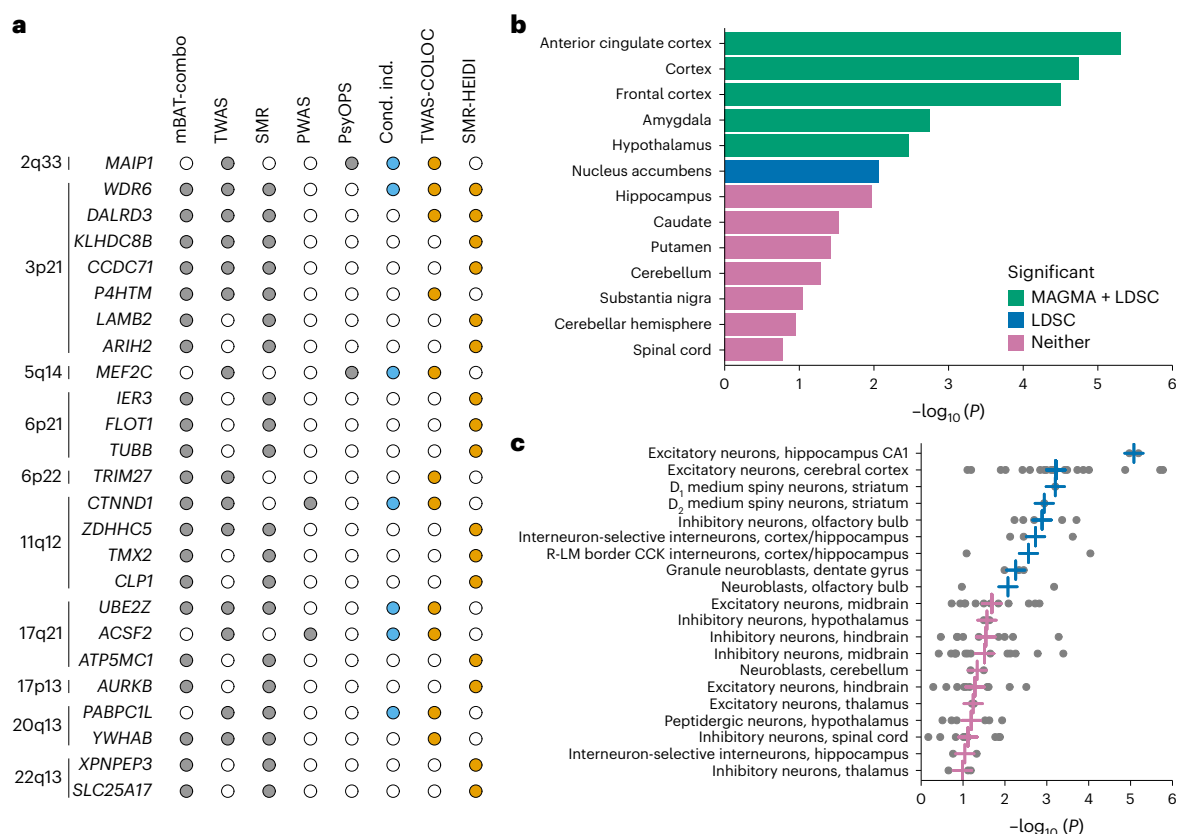


Fig. 2 | Gene-based, tissue and cell type enrichment analyses. a, List of 25 genes that were implicated in at least two of the five different gene-based tests (significance indicated by gray dots) and passed the TWAS colocalization and/or SMR-HEIDI filters (significance indicated by orange dots). Conditionally independent (cond. ind.) genes within each locus are indicated by blue dots. **b**, Enrichment of OCD GWAS signal in human brain-related tissues from GTEx (version 8). No significant enrichment was observed in the peripheral tissues (not included in the figure). The horizontal bar size represents the significance of the enrichment measured using the MAGMA gene set enrichment test or partitioned LDSC. **c**, Top 20 groups of brain cell types ($n = 35$ total tested) enriched with OCD

GWAS signal using MAGMA. Dots represent $-\log_{10}(P)$ values from MAGMA gene set enrichment tests of individual neuronal cell types from Zeisel et al.³⁰. Vertical crosses represent the mean $-\log_{10}(P)$ value observed for each brain cell type group. Blue crosses represent a significant enrichment of OCD GWAS signals (FDR across 35 groups, FDR < 0.05), while pink crosses indicate nonsignificant enrichment. Gray points represent the association ($-\log_{10}(P)$ value) for each single cell cluster ('level 5' analysis defined by Zeisel et al.³⁰) in a given cell type (for example, excitatory neurons, cerebral cortex). CCK, cholecystokinin-expressing; R-LM, stratum radiatum-stratum lacunosum-moleculare.

types showed enriched gene expression of OCD-associated genetic signals using a previously described approach²⁹ on published human gene expression datasets from bulk tissue RNA-seq data from the Genotype-Tissue Expression (GTEx) project and single-cell RNA-sequencing data from the adult mouse central and peripheral nervous systems³⁰. We found enrichment of OCD GWAS signals in six of 13 human brain tissue types in GTEx but no enrichment in human peripheral tissues (Fig. 2b and Supplementary Table 15). In the adult mouse central and peripheral nervous systems, we found enrichment of OCD GWAS signals in 41 of 166 tested specific single cell types using the MAGMA gene set enrichment test (Supplementary Table 16). When summarizing results of individual single cell types into groups of cell types defined by the same region or tissue and cell type, nine of 35 were enriched for OCD GWAS signals (top 20 shown in Fig. 2c). Strong enrichment of OCD GWAS signal was especially observed in excitatory neurons of the hippocampus and the cerebral cortex as well as in D₁ and D₂ medium spiny neurons (MSNs).

Genetic relationship of OCD with other phenotypes

Using genome-wide association analysis, we examined whether the 30 independent OCD-associated loci identified by our GWAS meta-analysis have previously been associated with other phenotypes (see Supplementary Tables 17a–d for lookups in four, partially overlapping GWAS databanks and Table 1 for highlighted associations). We found that 22 of the 30 loci were associated with other phenotypes, including

schizophrenia (seven loci), depression and major depressive disorder (two loci), bipolar disorder (one locus), neuroticism (seven loci), educational attainment (seven loci) and body fat mass or body mass index (eight loci).

We further used bivariate linkage disequilibrium score regression (LDSC)³¹ to investigate the extent of genetic correlations between OCD and 112 previously published GWASs encompassing psychiatric, substance use and neurological phenotypes, among others (Fig. 3). We found that 65 phenotypes were significantly correlated with OCD after correcting for multiple testing using the Benjamini–Hochberg³² procedure to control the false discovery rate (FDR) at a threshold of 0.05. OCD was significantly positively correlated with all tested psychiatric phenotypes; the highest correlations were with anxiety ($r_G = 0.70$), depression ($r_G = 0.60$), anorexia nervosa ($r_G = 0.52$), Tourette syndrome ($r_G = 0.47$) and post-traumatic stress disorder (PTSD; $r_G = 0.48$). Significant positive genetic correlations were also obtained for neuroticism ($r_G = 0.53$), in particular for the worry subcluster ($r_G = 0.64$), and all individual items in the worry subcluster, with slightly lower estimates for the depressive subcluster ($r_G = 0.35$). Suicide attempt ($r_G = 0.40$), history of childhood maltreatment ($r_G = 0.37$) and tiredness ($r_G = 0.36$) were also notable for strong positive associations with OCD. Of the assessed neurological disorders, OCD was only significantly correlated with migraine ($r_G = 0.15$). Some autoimmune disorders, such as Crohn's disease ($r_G = -0.13$), ulcerative

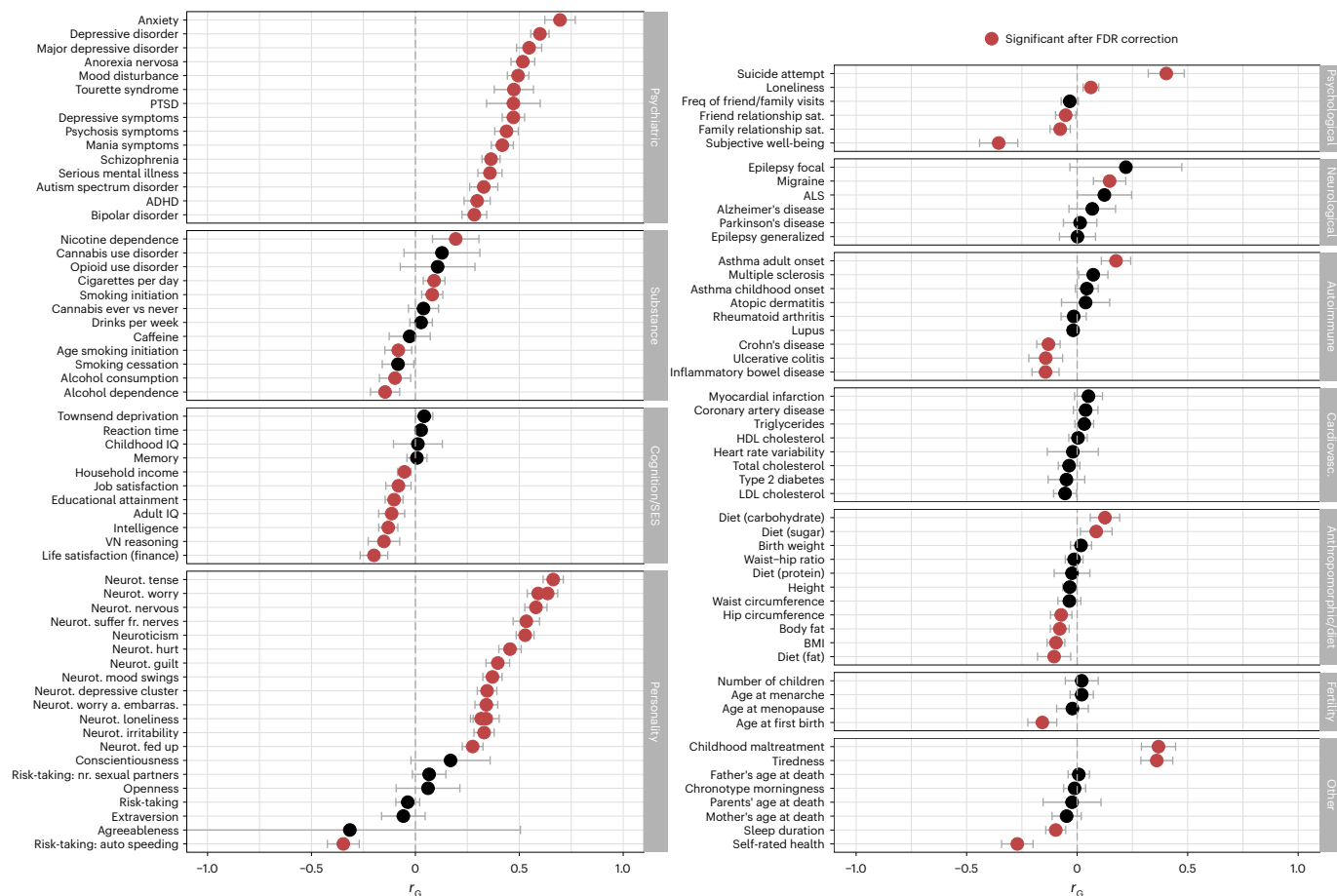


Fig. 3 | Genetic correlations (r_G) between OCD and 112 phenotypes.

This includes psychiatric, substance use, cognition–socioeconomic status (SES), personality, psychological, neurological, autoimmune, cardiovascular (cardiovasc.), anthropomorphic–diet, fertility and other phenotypes. References and sample sizes of the corresponding summary statistics of the GWAS studies can be found in Supplementary Table 18. The OCD summary statistics are of the main meta-analysis ($n_{\text{cases}} = 53,660$ and $n_{\text{controls}} = 2,044,417$). Error bars represent the 95% confidence intervals for the genetic correlation estimates (r_G). Red circles

indicate significant associations with a P value adjusted for multiple testing with the Benjamini–Hochberg procedure to control the FDR (<0.05). Black circles indicate associations that are not significant. a., after; ADHD, attention-deficit hyperactivity disorder; ALS, amyotrophic lateral sclerosis; BMI, body mass index; embarrass., embarrassment; freq, frequency; fr., from; HDL, high-density lipoprotein; IQ, intelligence quotient; LDL, low-density lipoprotein; neurot., neuroticism; nr., number; PTSD, post-traumatic stress disorder; sat., satisfaction; VN, verbal-numerical.

colitis ($r_G = -0.14$) and inflammatory bowel disease ($r_G = -0.14$), showed negative correlations with OCD (see Fig. 3 and Supplementary Table 18 for all genetic correlation estimates, 95% confidence intervals and P values, Supplementary Note 6 for a more in-depth discussion of all significant genetic correlations and Supplementary Table 19 and Supplementary Figs. 41 and 42 for subgroup-specific genetic correlation estimates).

Discussion

The OCD GWAS reported here, comprising over 53,000 cases, identified 30 independent genome-wide significant loci. Common SNPs explained 6.7% of the variation in OCD risk in our meta-analysis (LDSC with an assumed population prevalence of 1%), a significant reduction from the 28% reported previously¹⁰. However, differences in the assumed population prevalence (where a lower assumed prevalence for LDSC heritability calculation results in a lower heritability estimate) and an increase in sample heterogeneity likely contributed to this discrepancy. The reduction in SNP heritability is in line with previous observations for closely related psychiatric disorders such as attention deficit hyperactivity disorder (ADHD)^{33,34} or depression^{17,35–37}, where expanding the phenotype definition increased genetic heterogeneity, potentially accounting for the observed decrease in SNP heritability. This aligns with the fact that heritability estimates for more homogeneous OCD

subgroups were higher: 16.4% for the clinically ascertained subgroup and 13.3% for the comorbid subgroup (Supplementary Note 10). The current estimates are comparable to those of other psychiatric and substance use disorders, with SNP heritability estimates ranging between 9% and 28%³⁸.

The most significant SNP (**rs78587207** ($P = 5.28 \times 10^{-12}$)) identified in the GWAS is located on chr11q12.1 and has been previously associated with several traits, including neuropsychiatric phenotypes³⁹ such as depressive symptoms⁴⁰ and neuroticism⁴⁰. Gene-based analyses identified four putative causal genes within this locus. The closest gene to **rs78587207** is **CTNND1**, which encodes the cell adhesion molecule p120 catenin. This gene was associated with OCD using three gene-based tests (mBAT-combo, TWAS and PWAS), and we found strong evidence for colocalization of the TWAS signal for **CTNND1** in the dlPFC. The dlPFC has been consistently implicated in the neural circuitry of OCD as well as in compulsivity more broadly as part of the cortico–striatal–thalamo–cortical circuitry^{41,42}. The protein product of **CTNND1** is a regulator of cell–cell adhesion⁴³ and has a crucial role in gene transcription, Rho GTPase activity and cytoskeletal organization^{44–46}. Other credible causal genes in the locus include **CLP1** (cleavage factor polyribonucleotide kinase subunit 1), **TMX2** (thioredoxin-related transmembrane protein 2) and **ZDHHC5** (zinc finger DHHC type palmitoyltransferase 5). Rare genetic mutations in

CLPI are associated with pontocerebellar hypoplasia type 10, a very rare autosomal recessive neurodegenerative disease characterized by brain atrophy and delayed myelination resulting in intellectual disability⁴⁷. *TMX2* is associated with increased risk of neurodevelopmental disorders with microcephaly, cortical malformations, spasticity and congenital nervous system abnormalities⁴⁸. *ZDHHCS* is broadly expressed in the brain, including the frontal cortex. *ZDHHCS* has not been implicated in brain development but has been linked to lung acinar adenocarcinoma and lung papillary adenocarcinoma in prior studies⁴⁹.

Our finding that approximately 11,500 (s.e. = 607) causal variants account for 90% of the SNP-based heritability of OCD suggests that OCD is more polygenic than other complex traits such as height ($n_{\text{causal}} = 4,000$), schizophrenia ($n_{\text{causal}} = 9,600$) and ADHD ($n_{\text{causal}} = 5,600$) but less polygenic than major depression ($n_{\text{causal}} = 14,500$) and educational attainment ($n_{\text{causal}} = 13,200$)⁵⁰.

We identified a total of 25 credible causal genes based on robust evidence using multiple positional and functionally informed gene-based approaches. Notably, *DLGAP1*, which has been previously implicated in OCD pathogenesis^{10,51}, was not identified in either the GWAS or in the gene-based analyses. Of the 25 genes that were implicated, 15 were within 6.5 kb of a SNP that surpassed genome-wide significance in the meta-analysis. In addition to the four genes discussed above, several others are of particular interest, including *WDR6* and *DALRD3*, which had the strongest evidence from the gene-based analyses. These genes lie in a gene-rich region on chr3p21.31, which, in addition to harboring multiple genome-wide significant SNPs, has been previously associated with a broad range of psychiatric disorders and related traits, including schizophrenia³⁹, well-being⁵² and the worry subcluster of neuroticism⁵³.

WDR6 is broadly expressed in the brain, particularly the hypothalamus. Its protein product is involved in cell growth arrest⁵⁴, and recent studies have implicated it in anorexia nervosa⁵⁵ and Parkinson's disease⁵⁶. *DALRD3* is located on chromosome 3 in the same region as *WDR6*. *DALRD3*, when fully disrupted, is implicated in a form of epileptic encephalopathy with associated developmental delay⁵⁷. Finally, a third gene in the 3p21 locus, *CELSR3* (cadherin EGF LAG seven-pass G type receptor 3), encodes a protocadherin that is highly expressed in the developing basal ganglia⁵⁸. Multiple loss-of-function mutations in *CELSR3* have been associated with Tourette syndrome^{59,60}, which co-occurs with OCD in 10–20% of patients.

Four other genes identified through these analyses are located in the MHC locus, a region on chromosome 6 that has a major role in the adaptive immune system and has been repeatedly linked to major psychiatric disorders⁶¹. The newly identified MHC association for OCD is noteworthy given evidence linking OCD with autoimmune disorders^{62–64}. Genetic pleiotropy may underlie this connection, with variants predisposing individuals to both autoimmune conditions and OCD⁶⁵. Furthermore, some OCD subtypes, such as pediatric acute-onset neuropsychiatric disorders associated with *Streptococcus* and pediatric acute-onset neuropsychiatric syndrome, may have autoimmune origins^{66,67}. Nevertheless, we were surprised to discover several negative genetic correlations between OCD and autoimmune disorders such as Crohn's disease, ulcerative colitis and inflammatory bowel disease in our analyses, suggesting that there is heterogeneity (and perhaps pleiotropy) in the genetic relationships between autoimmune disorders and OCD.

Tissue and cell type enrichment analysis revealed significant enrichment of OCD SNP heritability in several tissues and cell types, with the strongest enrichment in excitatory neurons of the hippocampus and the cerebral cortex and in dopamine D₁ receptor (D₁R)-positive and dopamine D₂ receptor (D₂R)-positive MSNs in the striatum. These findings are in line with traditional neural circuitry models of OCD, which focus on frontal cortical–striatal pathways^{68,69}. These findings are consistent with and build on previous work linking various neuronal cell types to psychiatric and cognitive phenotypes⁷⁰.

Interestingly, the frontal and anterior cingulate cortices, which were enriched in our tissue-based analyses, as well as the hippocampus and the striatum, which were implicated in our cell type-based analyses, are among the regions that are consistently implicated in neuroimaging studies of OCD^{41,71–73}. Enrichment in MSNs in the striatum is consistent with their role in the observed aberrant circuitry in OCD, where the D₁ MSNs project to the globus pallidus interna and the substantia nigra in the direct pathway and the D₂ type MSNs project to the globus pallidus externa in the indirect pathway⁷⁴. However, MSNs are also enriched in major depressive disorder⁷⁵, schizophrenia⁷⁶ and intelligence⁷⁷, suggesting that the observed enrichment is not specific for OCD.

Our analyses of the shared genetic risk between OCD and other psychiatric disorders provides further insights into the etiology of OCD. In line with previous observations^{38,78}, OCD was significantly genetically correlated with multiple psychiatric disorders and traits. The strongest genetic correlations were observed for anxiety disorders, depression and anorexia nervosa, all of which are highly comorbid with OCD⁷⁹. This aligns with previous findings from cross-disorder analyses suggesting a shared genetic susceptibility among most psychiatric disorders^{38,80,81}. A notable exception is our finding that risk variants for OCD are protective for alcohol dependence⁸², which is at odds with epidemiological evidence strongly linking OCD and alcohol-related disorders⁸³ but in line with a recent paper⁷⁹ reporting a lower-than-expected lifetime comorbidity of substance use disorders in OCD. The observed pattern of correlations with other phenotypes can be thought of as falling into two categories: compulsivity–impulsivity and rumination–worry–neuroticism. In both categories, the patterns of genetic correlations appear to follow a gradient across disorders and traits. For example, in the compulsivity–impulsivity category, strong positive correlations are seen with anorexia nervosa and Tourette syndrome, which are disorders with strong compulsive features, with less positive associations seen with ADHD and negative correlations with alcohol dependence and risk-taking behaviors, which are all phenotypes characterized by impulsivity. A similar gradient is observed for the rumination–worry–neuroticism-related phenotypes, with strong positive correlations with anxiety and other ruminative phenotypes such as worry, transitioning to less strong correlations with individual depression-related items.

This study marks the transition from the flat (sample-building) phase of SNP discovery described for GWAS⁸⁴ (Supplementary Fig. 20), where few to no genome-wide significant loci are identified^{10,12,51,85}, to the linear phase of SNP discovery, where even relatively small increases in sample size identify additional genome-wide significant loci¹⁸. The strengths of the current study therefore include the marked increase in the number of OCD cases and the rigorous analytic methods, including two multivariate approaches (multi-trait analysis of GWAS (MTAG) and GenomicSEM) to control for potential overlapping study participants and to examine potential heterogeneity between the multiple ascertainment approaches. Potential weaknesses include the inability to document comorbid psychiatric disorders in the majority of cases that were not ascertained from clinical collections or electronic registries, the lack of inclusion of non-European ancestries and the limited availability of sex chromosome data. Owing to the nature of our study, imputation references used in the different cohorts were heterogeneous and did not allow for confident analysis of rare variant associations. Future larger-scale sequencing studies that are currently underway will be needed to identify associations in this allele frequency spectrum. We also note that the genetic correlation analyses are impacted by residual heterogeneity in genetic signals owing to the employment of heterogeneous ascertainment strategies.

In summary, this work substantially advances the field of OCD genetics by identifying new OCD genetic risk loci and multiple credible candidate causal genes, including those expressed in brain regions and cell types previously implicated in OCD⁸⁶. We have also shown that OCD is highly polygenic in nature, with many variants implicated

not only in OCD but also in commonly comorbid disorders or traits, in particular, anxiety, neuroticism, anorexia nervosa and depression. The observation that common variants explain only a modest amount of the phenotypic variation in OCD suggests that other types of genetic variation may also contribute to the etiology of OCD. Notably, whole-exome-sequencing studies have suggested that a substantial proportion of OCD cases (22%) may be influenced by rare de novo coding variants⁸⁷, especially in genes that are intolerant to loss of function⁸⁸. Similarly, rare potentially damaging copy number variations represent part of the risk architecture for OCD⁹. These findings emphasize the need for a comprehensive exploration of the contribution of both common and rare genetic factors as well as their interplay to OCD risk. Finally, with the implication of the MHC complex, we provide additional evidence for potential shared genetic influences underlying both OCD and increased liability to autoimmune processes, although the directionality of those relationships remains to be definitively elucidated. In addition to continuing to increase sample sizes, future studies will require ancestrally diverse samples to further facilitate the discovery of additional OCD risk variants. Similarly, sex-specific analyses and additional clinical phenotyping will allow for the further elucidation of genetic and clinical relationships between OCD and co-occurring disorders. Finally, with the emergence of drug databases describing the relations between drugs and molecular phenotypes⁸⁹, our results may be useful for drug repurposing (that is, identifying existing drugs targeting OCD risk genes), leading to new opportunities to find more effective treatments.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41588-025-02189-z>.

References

1. Fawcett, E. J., Power, H. & Fawcett, J. M. Women are at greater risk of OCD than men: a meta-analytic review of OCD prevalence worldwide. *J. Clin. Psychiatry* **81**, 19r13085 (2020).
2. World Health Organization. *The Global Burden of Disease: 2004 Update* (WHO Press, 2008).
3. Fernández de la Cruz, L. et al. Suicide in obsessive-compulsive disorder: a population-based study of 36 788 Swedish patients. *Mol. Psychiatry* **22**, 1626–1632 (2017).
4. Meier, S. M. et al. Mortality among persons with obsessive-compulsive disorder in Denmark. *JAMA Psychiatry* **73**, 268–274 (2016).
5. Blanco-Vieira, T. et al. The genetic epidemiology of obsessive-compulsive disorder: a systematic review and meta-analysis. *Transl. Psychiatry* **13**, 230 (2023).
6. Burton, C. L. et al. Heritability of obsessive-compulsive trait dimensions in youth from the general population. *Transl. Psychiatry* **8**, 191 (2018).
7. Pauls, D. L. The genetics of obsessive compulsive disorder: a review of the evidence. *Am. J. Med. Genet. C Semin. Med. Genet.* **148**, 133–139 (2008).
8. van Grootheest, D. S., Cath, D. C., Beekman, A. T. & Boomsma, D. I. Twin studies on obsessive-compulsive disorder: a review. *Twin Res. Hum. Genet.* **8**, 450–458 (2005).
9. Mahjani, B. et al. The genetic architecture of obsessive-compulsive disorder: contribution of liability to OCD from alleles across the frequency spectrum. *Am. J. Psychiatry* **179**, 216–225 (2022).
10. International Obsessive Compulsive Disorder Foundation Genetics Collaborative (IOCDF-GC) and OCD Collaborative Genetics Association Studies (OC GAS). Revealing the complex genetic architecture of obsessive-compulsive disorder using meta-analysis. *Mol. Psychiatry* **23**, 1181–1188 (2018).
11. Davis, L. K. et al. Partitioning the heritability of Tourette syndrome and obsessive compulsive disorder reveals differences in genetic architecture. *PLoS Genet.* **9**, e1003864 (2013).
12. Strom, N. I. et al. Genome-wide association study identifies new locus associated with OCD. Preprint at *medRxiv* <https://doi.org/10.1101/2021.10.13.21261078> (2021).
13. Strom, N. I. et al. Genome-wide association study identifies new loci associated with OCD. Preprint at *medRxiv* <https://doi.org/10.1101/2024.03.06.24303776> (2024).
14. 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).
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. Grotzinger, A. D. et al. Genomic structural equation modelling provides insights into the multivariate genetic architecture of complex traits. *Nat. Hum. Behav.* **3**, 513–525 (2019).
17. 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).
18. Strom, N. I. et al. Genome-wide association study of major anxiety disorders in 122,341 European-ancestry cases identifies 58 loci and highlights GABAergic signaling. Preprint at *medRxiv* <https://doi.org/10.1101/2024.07.03.24309466> (2024).
19. Li, A. et al. mBAT-combo: a more powerful test to detect gene-trait associations from GWAS data. *Am. J. Hum. Genet.* **110**, 30–43 (2023).
20. Gusev, A. et al. Integrative approaches for large-scale transcriptome-wide association studies. *Nat. Genet.* **48**, 245–252 (2016).
21. Gandal, M. J. et al. Shared molecular neuropathology across major psychiatric disorders parallels polygenic overlap. *Science* **359**, 693–697 (2018).
22. Zhu, Z. et al. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nat. Genet.* **48**, 481–487 (2016).
23. Vösa, U. et al. Large-scale cis- and trans-eQTL analyses identify thousands of genetic loci and polygenic scores that regulate blood gene expression. *Nat. Genet.* **53**, 1300–1310 (2021).
24. Qi, T. et al. Identifying gene targets for brain-related traits using transcriptomic and methylomic data from blood. *Nat. Commun.* **9**, 2282 (2018).
25. Wingo, T. S. et al. Brain proteome-wide association study implicates novel proteins in depression pathogenesis. *Nat. Neurosci.* **24**, 810–817 (2021).
26. Wainberg, M., Merico, D., Keller, M. C., Fauman, E. B. & Tripathy, S. J. Predicting causal genes from psychiatric genome-wide association studies using high-level etiological knowledge. *Mol. Psychiatry* **27**, 3095–3106 (2022).
27. Giambartolomei, C. et al. Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. *PLoS Genet.* **10**, e1004383 (2014).
28. Wallace, C. Eliciting priors and relaxing the single causal variant assumption in colocalisation analyses. *PLoS Genet.* **16**, e1008720 (2020).
29. Bryois, J. et al. Genetic identification of cell types underlying brain complex traits yields insights into the etiology of Parkinson's disease. *Nat. Genet.* **52**, 482–493 (2020).
30. Zeisel, A. et al. Molecular architecture of the mouse nervous system. *Cell* **174**, 999–1014 (2018).
31. Bulik-Sullivan, B. K. et al. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* **47**, 291–295 (2015).

32. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. B* **57**, 289–300 (1995).
33. Demontis, D. et al. Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. *Nat. Genet.* **51**, 63–75 (2019).
34. 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).
35. Hall, L. S. et al. Genome-wide meta-analyses of stratified depression in Generation Scotland and UK Biobank. *Transl. Psychiatry* **8**, 9 (2018).
36. Als, T. D. et al. Depression pathophysiology, risk prediction of recurrence and comorbid psychiatric disorders using genome-wide analyses. *Nat. Med.* **29**, 1832–1844 (2023).
37. Cai, N. et al. Minimal phenotyping yields genome-wide association signals of low specificity for major depression. *Nat. Genet.* **52**, 437–447 (2020).
38. Derks, E. M., Thorp, J. G. & Gerring, Z. F. Ten challenges for clinical translation in psychiatric genetics. *Nat. Genet.* **54**, 1457–1465 (2022).
39. Pardiñas, A. F. et al. Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. *Nat. Genet.* **50**, 381–389 (2018).
40. Baselmans, B. M. L. et al. Multivariate genome-wide analyses of the well-being spectrum. *Nat. Genet.* **51**, 445–451 (2019).
41. van den Heuvel, O. A. et al. Brain circuitry of compulsivity. *Eur. Neuropsychopharmacol.* **26**, 810–827 (2016).
42. Shephard, E. et al. Toward a neurocircuit-based taxonomy to guide treatment of obsessive–compulsive disorder. *Mol. Psychiatry* **26**, 4583–4604 (2021).
43. Davis, M. A., Ireton, R. C. & Reynolds, A. B. A core function for p120-catenin in cadherin turnover. *J. Cell Biol.* **163**, 525–534 (2003).
44. Yanagisawa, M. et al. A p120 catenin isoform switch affects Rho activity, induces tumor cell invasion, and predicts metastatic disease. *J. Biol. Chem.* **283**, 18344–18354 (2008).
45. Daniel, J. M. & Reynolds, A. B. The catenin p120^{ctn} interacts with Kaiso, a novel BTB/POZ domain zinc finger transcription factor. *Mol. Cell. Biol.* **19**, 3614–3623 (1999).
46. Ishiyama, N. et al. Dynamic and static interactions between p120 catenin and E-cadherin regulate the stability of cell–cell adhesion. *Cell* **141**, 117–128 (2010).
47. Schaffer, A. E. et al. *CLP1* founder mutation links tRNA splicing and maturation to cerebellar development and neurodegeneration. *Cell* **157**, 651–663 (2014).
48. Vandervore, L. V. et al. TMX2 is a crucial regulator of cellular redox state, and its dysfunction causes severe brain developmental abnormalities. *Am. J. Hum. Genet.* **105**, 1126–1147 (2019).
49. Zhang, Y. et al. Potential role of S-palmitoylation in cancer stem cells of lung adenocarcinoma. *Front. Cell Dev. Biol.* **9**, 734897 (2021).
50. Hindley, G. et al. Charting the landscape of genetic overlap between mental disorders and related traits beyond genetic correlation. *Am. J. Psychiatry* **179**, 833–843 (2022).
51. Mattheisen, M. et al. Genome-wide association study in obsessive–compulsive disorder: results from the OCGAS. *Mol. Psychiatry* **20**, 337–344 (2015).
52. Watanabe, K. et al. A global overview of pleiotropy and genetic architecture in complex traits. *Nat. Genet.* **51**, 1339–1348 (2019).
53. Nagel, M., Watanabe, K., Stringer, S., Posthuma, D. & van der Sluis, S. Item-level analyses reveal genetic heterogeneity in neuroticism. *Nat. Commun.* **9**, 905 (2018).
54. Xie, X., Wang, Z. & Chen, Y. Association of LKB1 with a WD-repeat protein WDR6 is implicated in cell growth arrest and p27^{Kip1} induction. *Mol. Cell. Biochem.* **301**, 115–122 (2007).
55. Adams, D. M., Reay, W. R. & Cairns, M. J. Multiomic prioritisation of risk genes for anorexia nervosa. *Psychol. Med.* **53**, 6754–6762 (2023).
56. Kia, D. A. et al. Identification of candidate Parkinson disease genes by integrating genome-wide association study, expression, and epigenetic data sets. *JAMA Neurol.* **78**, 464–472 (2021).
57. Lentini, J. M., Alsaif, H. S., Faqeh, E., Alkuraya, F. S. & Fu, D. *DALRD3* encodes a protein mutated in epileptic encephalopathy that targets arginine tRNAs for 3-methylcytosine modification. *Nat. Commun.* **11**, 2510 (2020).
58. Wu, J., Poppi, L. A. & Tischfield, M. A. Planar cell polarity and the pathogenesis of Tourette disorder: new hypotheses and perspectives. *Dev. Biol.* **489**, 14–20 (2022).
59. Willsey, A. J. et al. De novo coding variants are strongly associated with Tourette disorder. *Neuron* **94**, 486–499 (2017).
60. Zhao, X. et al. A whole-exome sequencing study of Tourette disorder in a Chinese population. *DNA Cell Biol.* **39**, 63–68 (2020).
61. Debnath, M., Berk, M., Leboyer, M. & Tamouza, R. The MHC/HLA gene complex in major psychiatric disorders: emerging roles and implications. *Curr. Behav. Neurosci. Rep.* **5**, 179–188 (2018).
62. Mataix-Cols, D. et al. A total-population multigenerational family clustering study of autoimmune diseases in obsessive–compulsive disorder and Tourette’s/chronic tic disorders. *Mol. Psychiatry* **23**, 1652–1658 (2018).
63. Tylee, D. S. et al. Genetic correlations among psychiatric and immune-related phenotypes based on genome-wide association data. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* **177**, 641–657 (2018).
64. Westwell-Roper, C. et al. Immune-related comorbidities in childhood-onset obsessive compulsive disorder: lifetime prevalence in the Obsessive Compulsive Disorder Collaborative Genetics Association Study. *J. Child Adolesc. Psychopharmacol.* **29**, 615–624 (2019).
65. Zhang, T. et al. Prenatal and early childhood infections and subsequent risk of obsessive–compulsive disorder and tic disorders: a nationwide, sibling-controlled study. *Biol. Psychiatry* **93**, 1023–1030 (2023).
66. Swedo, S. E. et al. Pediatric autoimmune neuropsychiatric disorders associated with streptococcal infections: clinical description of the first 50 cases. *Am. J. Psychiatry* **155**, 264–271 (1998).
67. Wilbur, C. et al. PANDAS/PANS in childhood: controversies and evidence. *Paediatr. Child Health* **24**, 85–91 (2019).
68. Ahmari, S. E. et al. Repeated cortico–striatal stimulation generates persistent OCD-like behavior. *Science* **340**, 1234–1239 (2013).
69. Ahmari, S. E. & Rauch, S. L. The prefrontal cortex and OCD. *Neuropsychopharmacology* **47**, 211–224 (2022).
70. Olsislagers, M., Rademaker, K., Adan, R. A. H., Lin, B. D. & Luykx, J. J. Comprehensive analyses of RNA-seq and genome-wide data point to enrichment of neuronal cell type subsets in neuropsychiatric disorders. *Mol. Psychiatry* **27**, 947–955 (2022).
71. Boedhoe, P. S. W. et al. Cortical abnormalities associated with pediatric and adult obsessive–compulsive disorder: findings from the ENIGMA Obsessive–Compulsive Disorder Working Group. *Am. J. Psychiatry* **175**, 453–462 (2018).
72. Bruin, W. B. et al. Structural neuroimaging biomarkers for obsessive–compulsive disorder in the ENIGMA-OCD consortium: medication matters. *Transl. Psychiatry* **10**, 342 (2020).
73. van den Heuvel, O. A. et al. An overview of the first 5 years of the ENIGMA obsessive–compulsive disorder working group: the power of worldwide collaboration. *Hum. Brain Mapp.* **43**, 23–36 (2022).
74. Haber, S. N. Corticostriatal circuitry. *Dialogues Clin. Neurosci.* **18**, 7–21 (2016).

75. Major Depressive Disorder Working Group of the Psychiatric Genomics Consortium. Trans-ancestry genome-wide study of depression identifies 697 associations implicating cell types and pharmacotherapies. *Cell* **188**, 640–652 (2025).
 76. Trubetskoy, V. et al. Mapping genomic loci implicates genes and synaptic biology in schizophrenia. *Nature* **604**, 502–508 (2022).
 77. Savage, J. E. et al. Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. *Nat. Genet.* **50**, 912–919 (2018).
 78. Romero, C. et al. Exploring the genetic overlap between twelve psychiatric disorders. *Nat. Genet.* **54**, 1795–1802 (2022).
 79. Sharma, E. et al. Comorbidities in obsessive–compulsive disorder across the lifespan: a systematic review and meta-analysis. *Front. Psychiatry* **12**, 703701 (2021).
 80. Lee, P. H. et al. Genomic relationships, novel loci, and pleiotropic mechanisms across eight psychiatric disorders. *Cell* **179**, 1469–1482 (2019).
 81. Grotzinger, A. D. Shared genetic architecture across psychiatric disorders. *Psychol. Med.* **51**, 2210–2216 (2021).
 82. Walters, R. K. et al. Transancestral GWAS of alcohol dependence reveals common genetic underpinnings with psychiatric disorders. *Nat. Neurosci.* **21**, 1656–1669 (2018).
 83. Virtanen, S. et al. Association of obsessive–compulsive disorder and obsessive–compulsive symptoms with substance misuse in 2 longitudinal cohorts in Sweden. *JAMA Netw. Open* **5**, e2214779 (2022).
 84. Sullivan, P. F. et al. Psychiatric genomics: an update and an agenda. *Am. J. Psychiatry* **175**, 15–27 (2018).
 85. Stewart, S. E. et al. Genome-wide association study of obsessive–compulsive disorder. *Mol. Psychiatry* **18**, 788–798 (2013).
 86. Piantadosi, S. C. et al. Transcriptome alterations are enriched for synapse-associated genes in the striatum of subjects with obsessive–compulsive disorder. *Transl. Psychiatry* **11**, 171 (2021).
 87. Cappi, C. et al. De novo damaging DNA coding mutations are associated with obsessive–compulsive disorder and overlap with Tourette’s disorder and autism. *Biol. Psychiatry* **87**, 1035–1044 (2020).
 88. Halvorsen, M. et al. Exome sequencing in obsessive–compulsive disorder reveals a burden of rare damaging coding variants. *Nat. Neurosci.* **24**, 1071–1076 (2021).
 89. Knox, C. et al. DrugBank 6.0: the DrugBank Knowledgebase for 2024. *Nucleic Acids Res.* **52**, D1265–D1275 (2024).
 90. Wang, J. et al. CAUSALdb: a database for disease/trait causal variants identified using summary statistics of genome-wide association studies. *Nucleic Acids Res.* **48**, D807–D816 (2020).
 91. Buniello, A. et al. The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res.* **47**, D1005–D1012 (2019).
 92. Elsworth, B. et al. The MRC IEU OpenGWAS data infrastructure. Preprint at *bioRxiv* <https://doi.org/10.1101/2020.08.10.244293> (2020).
- Publisher’s note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
- Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.
- © The Author(s) 2025

Nora I. Strom^{1,2,3,4,5,261}✉, Zachary F. Gerring^{6,7,261}, Marco Galimberti^{8,9,261}, Dongmei Yu^{10,11,261}, Matthew W. Halvorsen¹², Abdel Abdellaoui¹³, Cristina Rodriguez-Fontenla^{14,15}, Julia M. Sealock¹⁶, Tim Bigdeli^{17,18}, Jonathan R. Coleman^{19,20}, Behrang Mahjani^{21,22}, Jackson G. Thorp^{23,24}, Katharina Bey²⁵, Christie L. Burton²⁶, Jurjen J. Luykx^{27,28}, Gwyneth Zai^{29,30}, Silvia Alemany^{31,32,33}, Christine Andre³⁴, Kathleen D. Askland³⁵, Julia Bäckman³, Nerisa Banaj³⁶, Cristina Barlassina³⁷, Judith Becker Nissen^{38,39}, O. Joseph Bienvenu⁴⁰, Donald Black⁴¹, Michael H. Bloch⁴², Sigrid Børte^{43,44,45}, Rosa Bosch^{46,47}, Michael Breen^{21,48,49}, Brian P. Brennan^{34,50}, Helena Brentani⁵¹, Joseph D. Buxbaum²¹, Jonas Bybjerg-Grauholm⁵², Enda M. Byrne⁵³, Judit Cabana-Dominguez^{31,32,33}, Beatriz Camarena⁵⁴, Adrian Camarena⁵⁵, Carolina Cappi^{56,57}, Angel Carracedo^{58,59,60}, Miguel Casas^{61,62}, Maria Cristina Cavallini⁶³, Valentina Ciullo³⁶, Edwin H. Cook⁶⁴, Jesse Crosby^{34,50}, Bernadette A. Cullen^{65,66}, Elles J. De Schipper³, Richard Delorme⁶⁷, Srdjan Djurovic^{45,68}, Jason A. Elias^{69,70}, Xavier Estivill⁷¹, Martha J. Falkenstein^{34,50}, Bengt T. Fundin⁷², Lauryn Garner³⁴, Christina Gironde³⁴, Fernando S. Goes⁷³, Marco A. Grados⁷⁴, Jakob Grove^{4,75,76,77}, Wei Guo⁷⁸, Jan Haavik^{79,80}, Kristen Hagen^{81,82,83}, Kelly Harrington^{84,85}, Alexandra Havdahl^{86,87}, Kira D. Höffler^{68,88,89}, Ana G. Hounie⁵⁷, Donald Hucks⁹⁰, Christina Hultman²², Magdalena Janecka^{21,91}, Eric Jenike³⁴, Elinor K. Karlsson^{92,93}, Kara Kelley³⁴, Julia Klawohn^{1,94}, Janice E. Krasnow⁹⁵, Kristi Krebs⁹⁶, Christoph Lange^{97,98}, Nuria Lanzagorta⁹⁹, Daniel Levey^{100,101}, Kerstin Lindblad-Toh^{93,102,103}, Fabio Macciardi¹⁰⁴, Brion Maher¹⁰⁵, Brittany Mathes³⁴, Evonne McArthur¹⁰⁶, Nathaniel McGregor¹⁰⁷, Nicole C. McLaughlin^{108,109}, Sandra Meier¹¹⁰, Euripedes C. Miguel¹¹¹, Maureen Mulhern²⁵, Paul S. Nestadt¹¹², Erika L. Nurmi¹¹³, Kevin S. O’Connell^{114,115}, Lisa Osiecki^{116,117}, Olga Therese Ousdal^{79,118}, Teemu Palviainen¹¹⁹, Nancy L. Pedersen²², Fabrizio Piras³⁶, Federica Piras¹²⁰, Sriramya Potluri³⁴, Raquel Rabionet^{121,122,123}, Alfredo Ramirez^{124,125,126,127,128}, Scott Rauch¹²⁹, Abraham Reichenberg¹³⁰, Mark A. Riddle¹³¹, Stephan Ripke^{132,133,134}, Maria C. Rosário¹³⁵, Aline S. Sampaio¹³⁶, Miriam A. Schiele¹³⁷, Anne Heidi Skogholt¹³⁸, Laura G. Sloofman²¹, Jan Smit¹³⁹, Maria Soler Artigas^{31,32,33,140}, Laurent F. Thomas^{141,142,143,144}, Eric Tiffet³⁴, Homero Vallada^{51,145}, Nathaniel van Kirk^{146,147}, Jeremy Veenstra-VanderWeele^{148,149}, Nienke N. Vulink¹⁵⁰, Christopher P. Walker¹⁵¹, Ying Wang¹⁵², Jens R. Wendland¹⁵³,

Bendik S. Winsvold ^{43,154,155}, Yin Yao ¹⁵⁶, Hang Zhou ^{157,158,159}, Estonian Biobank*, 23andMe Inc.*, Arpana Agrawal ¹⁶⁰, Pino Alonso ^{161,162,163,164}, Götz Berberich ¹⁶⁵, Kathleen K. Bucholz ¹⁶⁶, Cynthia M. Bulik ^{22,167,168}, Danielle Cath ^{169,170}, Damiaan Denys ¹⁷¹, Valsamma Eapen ^{172,173}, Howard Edenberg ¹⁷⁴, Peter Falkai ^{175,176}, Thomas V. Fernandez ¹⁷⁷, Abby J. Fyer ^{178,179}, J. M. Gaziano ^{180,181}, Dan A. Geller ^{50,182}, Hans J. Grabe ¹⁸³, Benjamin D. Greenberg ^{184,185,186}, Gregory L. Hanna ¹⁸⁷, Ian B. Hickie ¹⁸⁸, David M. Hougaard ^{52,76}, Norbert Kathmann¹, James Kennedy ³⁰, Dongbing Lai ¹⁸⁹, Mikael Landén ^{22,190}, Stéphanie Le Hellard ^{68,80}, Marion Leboyer ¹⁹¹, Christine Lochner¹⁹², James T. McCracken¹¹³, Sarah E. Medland ¹⁹³, Preben B. Mortensen ^{76,194,195}, Benjamin M. Neale ^{133,196}, Humberto Nicolini^{197,198}, Merete Nordentoft^{199,200}, Michele Pato²⁰¹, Carlos Pato²⁰¹, David L. Pauls⁵⁰, John Piacentini²⁰², Christopher Pittenger²⁰³, Danielle Posthuma ^{204,205}, Josep Antoni Ramos-Quiroga ^{206,207,208,209}, Steven A. Rasmussen¹⁰⁸, Margaret A. Richter^{30,210}, David R. Rosenberg²¹¹, Stephan Ruhrmann ²¹², Jack F. Samuels²¹³, Sven Sandin ^{21,22}, Paul Sandor ³⁰, Gianfranco Spalletta^{36,214}, Dan J. Stein ²¹⁵, S. Evelyn Stewart^{216,217,218}, Eric A. Storch²¹⁹, Barbara E. Stranger ^{220,221}, Maurizio Turiel²²², Thomas Werge ^{223,224}, Ole A. Andreassen^{225,226}, Anders D. Børglum ^{4,227,228}, Susanne Walitza^{229,230,231}, Kristian Hveem^{155,232,233}, Bjarne K. Hansen^{82,234}, Christian Rück³, Nicholas G. Martin ²³⁵, Lili Milani ⁹⁶, Ole Mors²³⁶, Ted Reichborn-Kjennerud^{237,238}, Marta Ribasés ^{31,32,33,140}, Gerd Kvale^{83,239}, David Mataix-Cols³, Katharina Domschke ^{137,240}, Edna Grünblatt ^{229,230,231}, Michael Wagner^{25,241,242}, John-Anker Zwart ^{155,243}, Gerome Breen²⁴⁴, Gerald Nestadt ¹¹², Jaakko Kaprio ²⁴⁵, Paul D. Arnold^{246,247}, Dorothy E. Grice ²¹, James A. Knowles²⁴⁸, Helga Ask ^{249,250}, Karin J. Verweij ¹⁵⁰, Lea K. Davis ²⁵¹, Dirk J. Smit ²⁵², James J. Crowley ^{3,12,167}, Jeremiah M. Scharf^{116,133,262}, Murray B. Stein ^{253,254,262}, Joel Gelernter ^{255,256,262}, Carol A. Mathews^{257,262}, Eske M. Derks ^{258,262} & Manuel Mattheisen ^{2,4,259,262} ✉

¹Department of Psychology, Humboldt-Universität zu Berlin, Berlin, Germany. ²Department of Psychiatric Phenomics and Genomics (IPPG), Ludwig-Maximilians University Munich, Munich, Germany. ³Centre for Psychiatry Research, Department of Clinical Neuroscience, Karolinska Institutet and Stockholm Health Services, Region Stockholm, Stockholm, Sweden. ⁴Department of Biomedicine, Aarhus University, Aarhus, Denmark. ⁵University Hospital of Psychiatry and Psychotherapy, University of Bern, Bern, Switzerland. ⁶Department of Mental Health and Neuroscience, Translational Neurogenomics, QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia. ⁷Department of Population Health and Immunity, Healthy Development and Ageing, Walter and Eliza Hall Institute of Medical Research, Melbourne, Victoria, Australia. ⁸Department of Psychiatry, Human Genetics, Yale University, New Haven, CT, USA. ⁹VA Connecticut Healthcare System, West Haven, CT, USA. ¹⁰Department of Center for Genomic Medicine, Psychiatric and Neurodevelopmental Genetics Unit, Massachusetts General Hospital, Boston, MA, USA. ¹¹Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge, MA, USA. ¹²Department of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA. ¹³Department of Psychiatry, Amsterdam University Medical Centers, University of Amsterdam, Amsterdam, the Netherlands. ¹⁴Center for Research in Molecular Medicine and Chronic Diseases (CiMUS), Genomics and Bioinformatics, University of Santiago de Compostela, Santiago de Compostela, Spain. ¹⁵Grupo de Medicina Xenómica, Genetics, Instituto de Investigación Sanitaria de Santiago de Compostela (FIDIS), Santiago de Compostela, Spain. ¹⁶Vanderbilt Genetics Institute, Vanderbilt University, Nashville, TN, USA. ¹⁷Department of Psychiatry and Behavioral Sciences, SUNY Downstate Health Sciences University, Brooklyn, NY, USA. ¹⁸VA NY Harbor Healthcare System, Brooklyn, NY, USA. ¹⁹Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK. ²⁰National Institute for Health and Care Research Maudsley Biomedical Research Centre, South London and Maudsley NHS Trust, London, UK. ²¹Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, NY, USA. ²²Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden. ²³Mental Health and Neuroscience Program, QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia. ²⁴Faculty of Medicine, School of Biomedical Sciences, University of Queensland, Brisbane, Queensland, Australia. ²⁵Department of Psychiatry and Psychotherapy, University Hospital Bonn, Bonn, Germany. ²⁶Department of Neurosciences and Mental Health, Hospital for Sick Children, Toronto, Ontario, Canada. ²⁷Department of Psychiatry, Brain University Medical Center Utrecht, Utrecht, the Netherlands. ²⁸Second Opinion Outpatient Clinic, GGNet, Warnsveld, the Netherlands. ²⁹Molecular Brain Science Department, Campbell Family Mental Health Research Institute, Centre for Addiction and Mental Health, Toronto, Ontario, Canada. ³⁰Department of Psychiatry, University of Toronto, Toronto, Ontario, Canada. ³¹Psychiatric Genetics Unit, Group of Psychiatry, Mental Health and Addiction, Vall d'Hebron Research Institute, Universitat Autònoma de Barcelona, Barcelona, Spain. ³²Department of Mental Health, Hospital Universitari Vall d'Hebron, Barcelona, Spain. ³³Biomedical Network Research Centre on Mental Health (CIBERSAM), Madrid, Spain. ³⁴Obsessive-Compulsive Disorder Institute, McLean Hospital, Belmont, MA, USA. ³⁵Department of Psychiatry and Behavioural Neurosciences, McMaster University, Hamilton, Hamilton, Ontario, Canada. ³⁶Laboratory of Neuropsychiatry, IRCCS Santa Lucia Foundation, Rome, Italy. ³⁷Department of Health Sciences, University of Milano, Milano, Italy. ³⁸Department of Child and Adolescent Psychiatry, Aarhus University Hospital, Psychiatry, Aarhus, Denmark. ³⁹Institute of Clinical Medicine, Health, Aarhus University, Aarhus, Denmark. ⁴⁰Department of Psychiatry and Behavioral Sciences, General Hospital Psychiatry, Johns Hopkins University School of Medicine, Baltimore, MD, USA. ⁴¹Departments of Roy J. and Lucille A. Carver College of Medicine, University of Iowa, Iowa City, IA, USA. ⁴²Department of Child Study Center and Psychiatry, Yale University, New Haven, CT, USA. ⁴³Department of Research and Innovation, Division of Clinical Neuroscience, Oslo University Hospital, Oslo, Norway. ⁴⁴Department of Public Health and Nursing, Faculty of Medicine and Health Sciences, HUNT Center for Molecular and Clinical Epidemiology, Norwegian University of Science and Technology (NTNU), Trondheim, Norway. ⁴⁵Department of Medical Genetics, Oslo University Hospital, Oslo, Norway. ⁴⁶Department of Child and Adolescent Mental Health, Hospital Sant Joan de Déu, Esplugues de Llobregat, Spain. ⁴⁷Instituto de Salud Carlos III, Centro de Investigación Biomédica en Red de Salud Mental (CIBERSAM), Madrid, Spain. ⁴⁸Seaver Autism Center for Research and Treatment, Icahn School of Medicine at Mount Sinai, New York, NY, USA. ⁴⁹The Mindich Child Health and Development Institute, Icahn School of Medicine at Mount Sinai, New York, NY, USA. ⁵⁰Department of Psychiatry, Harvard Medical School, Boston, MA, USA. ⁵¹Department of Psychiatry, Universidade de São Paulo, São Paulo, Brazil. ⁵²Department for Congenital Disorders, Statens Serum Institut, Copenhagen, Denmark. ⁵³Child Health Research Centre, University of Queensland, Brisbane, Queensland, Australia. ⁵⁴Pharmacogenetics Department, Investigaciones Clínicas, Instituto Nacional de Psiquiatría Ramon de la Fuente Muñiz, Mexico City, México. ⁵⁵Department of Surgery, Duke University, Durham, NC, USA. ⁵⁶Department of Psychiatry, Icahn School of Medicine at Mount Sinai Hospital, New York, NY, USA. ⁵⁷Department of Psychiatry, University of São Paulo, São Paulo, Brazil. ⁵⁸CiMUS, Genomics and

Bioinformatics Group, University of Santiago de Compostela, Santiago de Compostela, Spain. ⁵⁹Galician Foundation of Genomic Medicine, Grupo de Medicina Xenómica, Instituto de Investigación Sanitaria de Santiago (IDIS), Santiago de Compostela, Spain. ⁶⁰Medicina Genómica, Centro de Investigación Biomédica en Red, Enfermedades Raras (CIBERER), Santiago de Compostela, Spain. ⁶¹Programa MIND Escoles, Hospital Sant Joan de Déu, Esplugues de Llobregat, Spain. ⁶²Departamento de Psiquiatría y Medicina Legal, Universitat Autònoma de Barcelona, Bellaterra, Spain. ⁶³Department of Psychiatry, Ospedale San Raffaele, Milano, Italy. ⁶⁴Department of Psychiatry, University of Illinois Chicago, Chicago, IL, USA. ⁶⁵Department of Psychiatry and Behavioral Sciences, Johns Hopkins Medical Institutions, Baltimore, MD, USA. ⁶⁶Department of Mental Health, Bloomberg School of Public Health, Baltimore, MD, USA. ⁶⁷Child and Adolescent Psychiatry Department, APHP, Paris, France. ⁶⁸Department of Clinical Science, University of Bergen, Bergen, Norway. ⁶⁹Psychiatry, McLean Hospital OCDI, Harvard Medical School, Belmont, MA, USA. ⁷⁰Adult Psychological Services, CBTeam LLC, Lexington, MA, USA. ⁷¹qGenomics (Quantitative Genomics Laboratories), Esplugues de Llobregat, Spain. ⁷²Department of Medical Epidemiology and Biostatistics, Center for Eating Disorders Innovation, Karolinska Institutet, Stockholm, Sweden. ⁷³Department of Psychiatry, Johns Hopkins University, Baltimore, MD, USA. ⁷⁴Department of Psychiatry and Behavioral Sciences, Child and Adolescent Psychiatry, Johns Hopkins University, Baltimore, MD, USA. ⁷⁵The Lundbeck Foundation Initiative for Integrative Psychiatric Research, iPSYCH, Aarhus, Denmark. ⁷⁶Center for Genomics and Personalized Medicine, Aarhus, Denmark. ⁷⁷Bioinformatics Research Centre, Aarhus, Denmark. ⁷⁸Genetic Epidemiology Research Branch, National Institute of Mental Health, Bethesda, MD, USA. ⁷⁹Department of Biomedicine, University of Bergen, Bergen, Norway. ⁸⁰Bergen Center for Brain Plasticity, Division of Psychiatry, Haukeland University Hospital, Bergen, Norway. ⁸¹Department of Psychiatry, Møre og Romsdal Hospital Trust, Molde, Norway. ⁸²Bergen Center for Brain Plasticity, Psychiatry, Haukeland University Hospital, Bergen, Norway. ⁸³Department of Mental Health, Norwegian University for Science and Technology, Trondheim, Norway. ⁸⁴Million Veteran Program (MVP) Coordinating Center, VA Boston Healthcare System, Boston, MA, USA. ⁸⁵Department of Psychiatry, Boston University Chobanian and Avedisian School of Medicine, Boston, MA, USA. ⁸⁶PsychGen Centre for Genetic Epidemiology and Mental Health, Norwegian Institute of Public Health, Oslo, Norway. ⁸⁷Nic Waals Institute, Lovisenberg Diaconal Hospital, Oslo, Norway. ⁸⁸Bergen Center for Brain Plasticity, Haukeland University Hospital, Bergen, Norway. ⁸⁹Department of Medical Genetics, Dr. Einar Martens Research Group for Biological Psychiatry, Haukeland University Hospital, Bergen, Norway. ⁹⁰Department of Medicine, Genetic Medicine, Vanderbilt University Medical Center, Nashville, TN, USA. ⁹¹Department of Child and Adolescent Psychiatry, NYU Grossman School of Medicine, New York, NY, USA. ⁹²Department of Bioinformatics and Integrative Biology, University of Massachusetts Medical School, Worcester, MA, USA. ⁹³Department of Vertebrate Genomics, Broad Institute of MIT and Harvard, Cambridge, MA, USA. ⁹⁴Department of Medicine, MSB Medical School Berlin, Berlin, Germany. ⁹⁵Department of Psychiatry and Behavioral Sciences, Johns Hopkins University, Baltimore, MD, USA. ⁹⁶Estonian Genome Centre, Institute of Genomics, University of Tartu, Tartu, Estonia. ⁹⁷Department of Biostatistics, T.H. Chan School of Public Health, Boston, MA, USA. ⁹⁸Department of Medicine, Channing Division of Network Medicine, Brigham and Women's Hospital, Boston, MA, USA. ⁹⁹Research Department, Grupo Medico Carracci, Mexico City, Mexico. ¹⁰⁰Department of Psychiatry, Yale University, West Haven, CT, USA. ¹⁰¹Office of Research and Development, United States Department of Veterans Affairs, West Haven, CT, USA. ¹⁰²Department of Medical Biochemistry and Microbiology, Uppsala University, Uppsala, Sweden. ¹⁰³SciLifeLab, Uppsala University, Uppsala, Sweden. ¹⁰⁴Department of Psychiatry, University of California, Irvine, Irvine, CA, USA. ¹⁰⁵Department of Mental Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA. ¹⁰⁶Vanderbilt Genetics Institute, Vanderbilt University Medical Center, Nashville, TN, USA. ¹⁰⁷COBRE Center for Neuromodulation, Butler Hospital, Providence, RI, USA. ¹⁰⁸Department of Psychiatry and Human Behavior, Alpert Medical School, Brown University, Providence, RI, USA. ¹⁰⁹Butler Hospital, Providence, RI, USA. ¹¹⁰Department of Psychiatry, Dalhousie University, Halifax, Nova Scotia, Canada. ¹¹¹Department of Psychiatry, Faculdade de Medicina, Universidade de São Paulo, São Paulo, Brazil. ¹¹²Department of Psychiatry and Behavioral Science, Johns Hopkins University, Baltimore, MD, USA. ¹¹³Department of Psychiatry and Biobehavioral Sciences, Division of Child and Adolescent Psychiatry, University of California, Los Angeles, Los Angeles, CA, USA. ¹¹⁴Department of Clinical Medicine, Division of Mental Health and Addiction, Oslo University Hospital, Oslo, Norway. ¹¹⁵NORMENT, University of Oslo, Oslo, Norway. ¹¹⁶Department of Psychiatry, Massachusetts General Hospital, Boston, MA, USA. ¹¹⁷Psychiatric and Neurodevelopmental Genetics Unit, Harvard Medical School, Boston, MA, USA. ¹¹⁸Department of Biomedicine, Haukeland University Hospital, Bergen, Norway. ¹¹⁹Institute for Molecular Medicine Finland (FIMM), University of Helsinki, Helsinki, Finland. ¹²⁰Department of Clinical Neuroscience and Neurorehabilitation, Neuropsychiatry Laboratory, IRCCS Santa Lucia Foundation, Rome, Italy. ¹²¹Department of Genetics, Microbiology and Statistics, IBUB, Universitat de Barcelona, Barcelona, Spain. ¹²²CIBERER, Centro de Investigación Biomédica en Red, Madrid, Spain. ¹²³Department of Human Molecular Genetics, Institut de Recerca Sant Joan de Déu, Barcelona, Spain. ¹²⁴Department of Psychiatry and Psychotherapy, Division of Neurogenetics and Molecular Psychiatry, Faculty of Medicine and University Hospital Cologne, University of Cologne, Cologne, Germany. ¹²⁵Department of Neurodegenerative Diseases and Geriatric Psychiatry, Medical Faculty, University Hospital Bonn, Bonn, Germany. ¹²⁶DZNE Bonn, German Center for Neurodegenerative Diseases (DZNE), Bonn, Germany. ¹²⁷Department of Psychiatry and Glenn Biggs Institute for Alzheimer's and Neurodegenerative Diseases, UT Health San Antonio, San Antonio, TX, USA. ¹²⁸Cologne Excellence Cluster for Stress Responses in Ageing-associated Diseases (CECAD), University of Cologne, Cologne, Germany. ¹²⁹Department of Psychiatry, McLean Hospital, Harvard Medical School, Belmont, MA, USA. ¹³⁰Department of Mental Disorders, Norwegian Institute of Public Health, New York, NY, USA. ¹³¹Department of Psychiatry and Behavioral Sciences, Child and Adolescent, Johns Hopkins University School of Medicine, Baltimore, MD, USA. ¹³²Department of Psychiatry and Psychotherapy, Charité—Universitätsmedizin, Berlin, Germany. ¹³³Stanley Center for Psychiatric Research, Broad Institute of Harvard and MIT, Cambridge, MA, USA. ¹³⁴Site Berlin–Potsdam, German Center for Mental Health (DZPG), Berlin, Germany. ¹³⁵Department of Psychiatry, Child and Adolescent Psychiatry Unit (UPIA), Federal University of São Paulo (UNIFESP), São Paulo, Brazil. ¹³⁶Department of Neurosciences and Mental Health, Medical School, Federal University of Bahia, Salvador, Brazil. ¹³⁷Department of Psychiatry and Psychotherapy, Faculty of Medicine, University of Freiburg, Medical Center—University of Freiburg, Freiburg, Germany. ¹³⁸Department of Public Health and Nursing, HUNT Center for Molecular and Clinical Epidemiology, Trondheim, Norway. ¹³⁹Department of Psychiatry, Faculty of Medicine, Locaion VUmc, Amsterdam UMC, University of Amsterdam, Amsterdam, the Netherlands. ¹⁴⁰Department of Genetics, Microbiology, and Statistics, Faculty of Biology, Universitat de Barcelona, Barcelona, Spain. ¹⁴¹Department of Clinical and Molecular Medicine, NTNU, Trondheim, Norway. ¹⁴²Department of Public Health and Nursing, K.G. Jebsen Center for Genetic Epidemiology, NTNU, Trondheim, Norway. ¹⁴³BioCore, Bioinformatics Core Facility, NTNU, Trondheim, Norway. ¹⁴⁴Clinic of Laboratory Medicine, St. Olavs Hospital, Trondheim University Hospital, Trondheim, Norway. ¹⁴⁵Department of Molecular Medicine and Surgery, CMM, Karolinska Institutet, Stockholm, Sweden. ¹⁴⁶OCD Institute, Division of Depression and Anxiety, McLean Hospital, Belmont, MA, USA. ¹⁴⁷Department of Psychiatry, Harvard Medical School, Belmont, MA, USA. ¹⁴⁸Department of Psychiatry, Division of Child and Adolescent Psychiatry, Columbia University, New York, NY, USA. ¹⁴⁹Department of Child and Adolescent Psychiatry, New York State Psychiatric Institute, New York, NY, USA. ¹⁵⁰Department of Psychiatry, Amsterdam UMC, University of Amsterdam, Amsterdam, the Netherlands. ¹⁵¹Human Genetics Institute of New Jersey, Rutgers University, Piscataway, NJ, USA. ¹⁵²Department of Neurology, the Johns Hopkins University School of Medicine, Baltimore, MD, USA. ¹⁵³Laboratory of Clinical Science, NIMH Intramural Research Program, Bethesda, MD, USA. ¹⁵⁴Department of Neurology, Oslo University Hospital, Oslo, Norway. ¹⁵⁵HUNT Center for Molecular and Clinical Epidemiology, Department of Public Health and Nursing, Faculty of Medicine and Health Sciences, NTNU, Trondheim, Norway. ¹⁵⁶Department of Computational Biology,

Institute of Life Science, Fudan University, Fudan, China. ¹⁵⁷Department of Psychiatry, Yale School of Medicine, New Haven, CT, USA. ¹⁵⁸Department of Psychiatry, Veterans Affairs Connecticut Healthcare System, West Haven, CT, USA. ¹⁵⁹Section of Biomedical Informatics and Data Science, Yale School of Medicine, New Haven, CT, USA. ¹⁶⁰Department of Psychiatry, Washington University School of Medicine, Saint Louis, MO, USA. ¹⁶¹Department of Psychiatry, OCD Clinical and Research Unit, Bellvitge Hospital, Barcelona, Spain. ¹⁶²Department of Clinical Sciences, University of Barcelona, Barcelona, Spain. ¹⁶³Department of Psychiatry and Mental Health, Bellvitge Biomedical Research Institute IDIBELL, Barcelona, Spain. ¹⁶⁴CIBERSAM, Mental Health Network Biomedical Research Center, Madrid, Spain. ¹⁶⁵Psychosomatic Department, Windach Hospital of Neurobehavioural Research and Therapy, Windach, Germany. ¹⁶⁶Department of Psychiatry, Washington University School of Medicine, St Louis, MO, USA. ¹⁶⁷Department of Psychiatry, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA. ¹⁶⁸Department of Nutrition, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA. ¹⁶⁹Departments of Rijksuniversiteit Groningen and Psychiatry, University Medical Center Groningen, Groningen, the Netherlands. ¹⁷⁰Department of Specialized Training, Drenthe Mental Health Care Institute, Groningen, the Netherlands. ¹⁷¹Department of Psychiatry, Institute of the Royal Netherlands Academy of Arts and Sciences (NIN-KNAW), Amsterdam, the Netherlands. ¹⁷²Discipline of Psychiatry and Mental Health, School of Clinical Medicine, UNSW, Sydney, New South Wales, Australia. ¹⁷³Academic Unit of Child Psychiatry South-West Sydney, South-West Sydney Clinical School, SWSLHD and Ingham Institute, Sydney, New South Wales, Australia. ¹⁷⁴Department of Biochemistry and Molecular Biology, Indiana University School of Medicine, Indianapolis, IN, USA. ¹⁷⁵Department of Psychiatry and Psychotherapy, University Hospital LMU, Munich, Germany. ¹⁷⁶Department of Psychiatry, Max Planck Institute, Munich, Germany. ¹⁷⁷Child Study Center and Psychiatry, Yale School of Medicine, New Haven, CT, USA. ¹⁷⁸Department of Psychiatry, New York State Psychiatric Institute, New York, NY, USA. ¹⁷⁹Department of Psychiatry, Columbia University Medical Center, New York, NY, USA. ¹⁸⁰Department of Medicine, VA Boston Healthcare System, Boston, MA, USA. ¹⁸¹Department of Medicine, Mass General Brigham, Boston, MA, USA. ¹⁸²Department of Psychiatry, Child Psychiatry, Massachusetts General Hospital, Boston, MA, USA. ¹⁸³Department of Psychiatry and Psychotherapy, University Medicine Greifswald, Greifswald, Germany. ¹⁸⁴COBRE Center on Neuromodulation, Butler Hospital, Providence, RI, USA. ¹⁸⁵Center for Neurorestoration and Neurotechnology, VA Providence Healthcare System, Providence, RI, USA. ¹⁸⁶Department of Psychiatry and Human Behavior, Alpert Medical School, Brown University, Providence, RI, USA. ¹⁸⁷Department of Psychiatry, Child and Adolescent Psychiatry, University of Michigan, Ann Arbor, MI, USA. ¹⁸⁸Brain and Mind Centre, the University of Sydney, Sydney, New South Wales, Australia. ¹⁸⁹Department of Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, IN, USA. ¹⁹⁰Institute of Neuroscience and Physiology, Department of Psychiatry and Neurochemistry, University of Gothenburg, Gothenburg, Sweden. ¹⁹¹Department of Addictology and Psychiatry, Université Paris-Est Créteil, AP-HP, Inserm, Paris, France. ¹⁹²Department of Psychiatry, SA MRC Unit on Risk and Resilience in Mental Disorders, Stellenbosch University, Stellenbosch, South Africa. ¹⁹³Department of Mental Health, Psychiatric Genetics, QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia. ¹⁹⁴National Centre for Register-based Research, Aarhus University, Aarhus, Denmark. ¹⁹⁵Centre for Integrated Register-based Research, Aarhus University, Aarhus, Denmark. ¹⁹⁶Analytic and Translational Genetics Unit, Massachusetts General Hospital, Boston, MA, USA. ¹⁹⁷Department of Psychiatry, Psychiatry, Carracci Medical Group, Mexico City, México. ¹⁹⁸Psiquiatría, Instituto Nacional de Medicina Genómica, Mexico City, México. ¹⁹⁹Mental Health Center Copenhagen, Copenhagen Research Center for Mental Health, Mental Health Services in the Capital Region of Denmark, Copenhagen, Denmark. ²⁰⁰Faculty of Health and Medical Sciences, Department of Clinical Medicine, University of Copenhagen, Copenhagen, Denmark. ²⁰¹Department of Psychiatry, Rutgers University, Piscataway, NJ, USA. ²⁰²Department of Psychiatry and Biobehavioral Sciences, Child and Adolescent Psychiatry, UCLA Semel Institute for Neuroscience and Human Behavior, Los Angeles, CA, USA. ²⁰³Department of Psychiatry, Yale University, New Haven, CT, USA. ²⁰⁴Department of Complex Trait Genetics, Vrije Universiteit Amsterdam, Center for Neurogenomics and Cognitive Research, Amsterdam Neuroscience, Amsterdam, the Netherlands. ²⁰⁵Department of Child and Adolescent Psychiatric, Section Complex Trait Genetics, VU Medical Center Amsterdam, Amsterdam, the Netherlands. ²⁰⁶Department of Psychiatry, Hospital Universitari Vall d'Hebron, Barcelona, Spain. ²⁰⁷Group of Psychiatry, Mental Health and Addictions, Psychiatric Genetics Unit, Vall d'Hebron Research Institute, Barcelona, Spain. ²⁰⁸CIBERSAM, Barcelona, Spain. ²⁰⁹Department of Psychiatry and Forensic Medicine, Universitat Autònoma de Barcelona, Barcelona, Spain. ²¹⁰Department of Psychiatry, Sunnybrook Health Sciences Centre, Toronto, Ontario, Canada. ²¹¹Department of Psychiatry and Behavioral Neurosciences, Child and Adolescent Psychiatry, Wayne State University School of Medicine, Detroit, MI, USA. ²¹²Department of Psychiatry and Psychotherapy, University of Cologne, Cologne, Germany. ²¹³Department of Psychiatry and Behavioral Sciences, the Johns Hopkins University School of Medicine, Baltimore, MD, USA. ²¹⁴Department of Psychiatry and Behavioral Sciences, Division of Neuropsychiatry, Baylor College of Medicine, Houston, TX, USA. ²¹⁵Department of Psychiatry and Neuroscience Institute, SAMRC Unit on Risk and Resilience in Mental Disorders, University of Cape Town, Cape Town, South Africa. ²¹⁶Department of Psychiatry, University of British Columbia, Vancouver, British Columbia, Canada. ²¹⁷British Columbia Children's Hospital Research Institute, Vancouver, British Columbia, Canada. ²¹⁸British Columbia Mental Health and Substance Use Services Research Institute, Vancouver, British Columbia, Canada. ²¹⁹Department of Psychiatry and Behavioral Sciences, Baylor College of Medicine, Houston, TX, USA. ²²⁰Department of Pharmacology, Northwestern University Feinberg School of Medicine, Chicago, IL, USA. ²²¹Center for Genetic Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL, USA. ²²²Department of Cardiology, University of Milan, Milan, Italy. ²²³Institute of Biological Psychiatry, Mental Health Center Sct. Hans, Copenhagen University Hospital, Mental Health Services (RHP), Copenhagen, Denmark. ²²⁴Institute of Clinical Medicine, University of Copenhagen, Copenhagen, Denmark. ²²⁵Institute of Clinical Medicine, NORMENT Centre, University of Oslo, Oslo, Norway. ²²⁶Division of Mental Health and Addiction, Center for Precision Psychiatry, Oslo University Hospital, Oslo, Norway. ²²⁷The Lundbeck Foundation Initiative for Integrative Psychiatric Research, iPSYCH, Aarhus University, Aarhus, Denmark. ²²⁸Center for Genomics and Personalized Medicine, Aarhus University, Aarhus, Denmark. ²²⁹Department of Child and Adolescent Psychiatry and Psychotherapy, Psychiatric University Hospital Zurich (PUK), University of Zurich, Zürich, Switzerland. ²³⁰Neuroscience Center Zurich, University of Zurich and the ETH Zurich, Zürich, Switzerland. ²³¹Zurich Center for Integrative Human Physiology, University of Zurich, Zürich, Switzerland. ²³²HUNT Research Center, Department of Public Health and Nursing, Faculty of Medicine and Health Sciences, NTNU, Trondheim, Norway. ²³³Department of Research, Innovation and Education, St. Olavs Hospital, Trondheim University Hospital, Trondheim, Norway. ²³⁴Centre for Crisis Psychology, Psychology, University of Bergen, Bergen, Norway. ²³⁵Department of Genetic Epidemiology, QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia. ²³⁶Psychosis Research Unit, Psychiatry, Aarhus University Hospital, Aarhus, Denmark. ²³⁷Department of Mental Disorders, Norwegian Institute of Public Health, Oslo, Norway. ²³⁸Institute of Clinical Medicine, University of Oslo, Oslo, Norway. ²³⁹Department of Clinical Psychology, Faculty of Psychology, University of Bergen, Bergen, Norway. ²⁴⁰Partner Site Berlin, DZPG, Berlin, Germany. ²⁴¹Department of Neurodegenerative Diseases and Geriatric Psychiatry, University Hospital Bonn, Bonn, Germany. ²⁴²DZNE, Bonn, Germany. ²⁴³Department of Research and Innovation, Clinical Neuroscience, Oslo University Hospital and University of Oslo, Oslo, Norway. ²⁴⁴Social, Genetic and Developmental Psychiatric Centre, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK. ²⁴⁵FIMM, University of Helsinki, Helsinki, Finland. ²⁴⁶Department of Psychiatry, the Mathison Centre for Mental Health Research and Education, Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada. ²⁴⁷Program in Genetics and Genome Biology, Hospital for Sick Children, Toronto, Ontario, Canada. ²⁴⁸Department of Genetics, Human Genetics Institute of New Jersey, Rutgers University, Piscataway, NJ, USA. ²⁴⁹PsychGen Center for Genetic Epidemiology, Norwegian

Institute of Public Health, Oslo, Norway. ²⁵⁰PROMENTA Research Center, Department of Psychology, University of Oslo, Oslo, Norway. ²⁵¹Department of Medicine, Division of Genetic Medicine, Vanderbilt University Medical Center, Nashville, TN, USA. ²⁵²Department of Psychiatry, Amsterdam UMC location AMC, Amsterdam, the Netherlands. ²⁵³Psychiatry Service, VA San Diego Healthcare System, San Diego, CA, USA. ²⁵⁴Department of Psychiatry and School of Public Health, University of California San Diego, La Jolla, CA, USA. ²⁵⁵Department of Psychiatry, Human Genetics (Psychiatry), Yale University School of Medicine, West Haven, CT, USA. ²⁵⁶Department of Psychiatry, Veterans Affairs Connecticut Healthcare Center, West Haven, CT, USA. ²⁵⁷Psychiatry and Genetics Institute, Evelyn F. and William L. Mc Knight Brain Institute, Center for OCD, Anxiety and Related Disorders, University of Florida, Gainesville, FL, USA. ²⁵⁸Department of Mental Health and Neuroscience, QIMR Berghofer, Brisbane, Queensland, Australia. ²⁵⁹Department of Community Health and Epidemiology and Faculty of Computer Science, Dalhousie University, Halifax, Nova Scotia, Canada. ²⁶¹These authors contributed equally: Nora I. Strom, Zachary F. Gerring, Marco Galimberti, Dongmei Yu. ²⁶²These authors jointly supervised this work: Jeremiah M. Scharf, Murray B. Stein, Joel Gelernter, Carol A. Mathews, Eske M. Derks, Manuel Mattheisen. *Lists of authors and their affiliations appear at the end of the paper.

✉ e-mail: nora.strom@hu-berlin.de; manuel.mattheisen@dal.ca

Estonian Biobank

Andres Metspalu⁹⁶, Tõnu Esko⁹⁶, Reedik Mägi⁹⁶, Mari Nelis⁹⁶ & Georgi Hudjashov⁹⁶

23andMe Inc.

Chris German²⁶⁰

²⁶⁰23andMe, Inc., Sunnyvale, CA, USA.

Methods

Ethics

All relevant ethics approvals have been obtained by the respective cohort's institutions, and a list of all respective approvals can be found in Supplementary Note 2.

Study participants

We analyzed genomic data from 28 OCD case-control cohorts including 53,660 OCD cases and 2,044,417 controls of European ancestry. Supplementary Table 1 provides an overview of the individual cohorts. A subset of the cases and controls have been included in previous studies^{10,51,85} and preprints^{12,13}, as described in Supplementary Note 2. Among all included individuals, 323 cases were part of a parent-proband trio; in these cases, parents were used as pseudocontrols. A total of 20,427 cases met DSM-5 (ref. 93) or ICD-10 (<https://icd.who.int/>) criteria for OCD as assessed by a healthcare professional or derived from (electronic) health records, while the remaining 32,233 cases were based on self-reported OCD diagnosis (23andMe, AGDS and parts of UKBB). Cohort-specific sample and analytic details can be found in Supplementary Note 2. Data collections were approved by the relevant institutional review boards at all participating sites, and all participants provided written informed consent.

Individual GWAS analyses and harmonizing of results

First, the data of each participating cohort were analyzed individually (see Supplementary Note 2 for details). Genetic data were imputed using either the Haplotype Reference Consortium (HRC)⁹⁴ or 1000 Genomes Project Phase 3 reference panels⁹⁵. The resulting GWAS summary statistics were then harmonized before a conjoint meta-analysis of all autosomes was conducted. Each summary statistic dataset was transformed to the 'daner' file format following RICOPILI⁹⁶ specifications. All variants had to meet the following criteria for inclusion: minor allele frequency (MAF) > 1% in cases and controls, INFO score > 0.8 and < 1.2. If the effect measure, *P* value or s.e. was missing or was out of bounds (infinite), the SNP was removed. Once cleaned summary statistics were produced, all datasets were aligned to the HRC reference panel. If variants were reported on different strands, they were flipped to the orientation in the HRC reference. Furthermore, strand-ambiguous A/T and C/G SNPs were removed if their MAF was > 0.4. In the case that A/T and C/G SNPs showed a MAF < 0.4, allele frequencies were compared to frequencies in the HRC reference. If an allele frequency match was found, that is, minor alleles were the same in the summary statistics and the HRC reference, the same strand orientation was assumed. If an allele mismatch was found, that is, the allele had a frequency > 0.5 in the HRC reference, it was assumed that alleles were reported on different strands, and alleles were flipped subsequently. Marker names were uniformly switched to those present in the HRC reference. If a variant did not overlap with the variants in the HRC reference, it was removed.

GWAS meta-analysis

Inverse-variance-weighted meta-analysis was conducted on 28 European cohorts using METAL⁹⁷. Weighting was based on standard error primarily to account for the large case-control imbalances in cohorts that used linear mixed model approaches in their primary GWAS. Heterogeneity was assessed with Cochran's *Q* statistic and the *I*² statistic^{98,99} (see Supplementary Note 5 for details). The genomic control factor λ was calculated for each individual GWAS and for the overall meta-analysis to identify residual population stratification or systematic technical artifacts. GWAS summary statistics were subjected to LDSC analyses on high-quality common SNPs (INFO score > 0.9) to examine the LDSC intercept to distinguish polygenicity from other types of inflation and to estimate the genetic heritability from the meta-analysis and genetic correlations between cohorts. The genomic inflation factor λ was estimated at 1.330 with a λ_{1000} of 1.033,

while the LDSC intercept was 1.0155 (s.e. = 0.0085), indicating that the inflation was mostly due to polygenic signal and unlikely to be substantially confounded by population structure. The genome-wide significance threshold for the GWAS was set at a *P* value of 5.0×10^{-8} . The 23andMe data included information on the X chromosome; as this information was not present for all other cohorts, analysis of the X chromosome was only conducted in this subcohort (see Supplementary Note 4 for details).

We further conducted GWAS meta-analyses on the following four subgroups, defined by differences in their sample ascertainment: (1) clinical OCD cases diagnosed by a healthcare professional in a clinical setting ($n_{\text{cases}} = 9,089$, $n_{\text{controls}} = 21,077$; including IOCDF, IOCDF_trio, EPOC, NORDiC-nor, NORDiC-swe, EGOS, OCGAS, OCGAS-ab, OCGAS-gh, OCGAS-nes, Psych_Broad, WWF, MVP, Michigan/Toronto IGS, YalePenn, Chop, CoGa), (2) comorbid individuals who were primarily ascertained for another comorbid psychiatric disorder ($n_{\text{cases}} = 5,266$, $n_{\text{controls}} = 43,760$; AGDS, iPSYCH), (3) biobank data from large-scale biobanks or registries with ICD or DSM codes ($n_{\text{cases}} = 9,138$, $n_{\text{controls}} = 1,049,776$; BioVU, EstBB, FinnGen, HUNT, MoBa, UKBB) or (4) 23andMe data ($n_{\text{cases}} = 30,167$, $n_{\text{controls}} = 929,804$). While these groups are not exclusive (for example, diagnoses in health records were originally given in a clinical setting or comorbid cases were also assessed in a clinical setting or derived from health records), we defined these groups by the cohort's primary characteristic. We also conducted one meta-analysis including all clinical, comorbid and biobank subgroups, while excluding the 23andMe data, resulting in 23,493 cases and 1,114,613 controls. As 23andMe is the only consumer-based dataset, we intended to compare this dataset to all others.

Number of trait-specific causal variants (MiXeR analysis)

We applied MiXeR version 1.3 (ref. 15) to quantify the polygenicity of OCD (that is, estimate the total number of trait-influencing genetic variants). MiXeR fits a Gaussian mixture model assuming that common genetic effects on a trait are a mixture of causal variants and noncausal variants. Polygenicity is reported as the number of causal variants that explain 90% of SNP heritability of OCD (to avoid extrapolating model parameters into the area of infinitesimally small effects).

SNP-based fine-mapping (GCTA-COJO)

We performed a conditional and joint analysis (GCTA-COJO)¹⁴ to identify independent signals within significant OCD loci. This approach performs a conditional and joint analysis on the basis of conditional *P* values before calculating the joint effects of all selected SNPs. We used the stepwise model selection procedure to select independently associated SNPs. The linkage disequilibrium reference sample was created from 73,005 individuals from the QIMR Berghofer Medical Research Institute genetic epidemiology cohort. The distance assumed for complete linkage disequilibrium was 10 Mb, and we used the default *P*-value threshold of 5×10^{-8} to define a genome-wide significant hit.

Multi-trait analysis of ascertainment subgroups

We used MTAG¹⁰⁰ to conduct multivariable GWAS analyses, reporting GWAS results for each of the ascertainment-specific subgroups. Through this approach, we aimed to address potential concerns about heterogeneity in genetic liability for individual subgroups following different ascertainment strategies. MTAG is a multi-trait analysis that is usually used to combine different but related traits into one meta-analysis by leveraging the shared heritability among the different traits and thereby gaining power. In this case, our aim was to generate ascertainment-specific estimates, while boosting power by leveraging the high shared heritability between the subgroups. The MTAG analysis resulted in four different GWAS summary statistics, one for each subgroup (clinical, comorbid, biobanks, 23andMe). We performed maxFDR analyses to approximate the upper bound on the FDR of MTAG results.

GenomicSEM

Similarly, we used GenomicSEM¹⁶ to model the joint genetic architecture of the four subgroups. First, we ran a common factor model without individual SNP effects, following the tutorial ‘Models without individual SNP effects’ on the GenomicSEM GitHub website (Code availability). Second, we ran a multivariate GWAS of the common factor (see Supplementary Note 5 for details). We specified the model using unit variance identification, for which the latent factor variance is fixed to 1 and the loadings of the traits are estimated freely. This ensures that we capture how much of each subgroup contributes to the latent factor. GenomicSEM also generates Q_{SNP} values, which indicate possible heterogeneous effects across the subgroups. The Q_{SNP} statistic is mathematically similar to the Q statistic from standard meta-analysis and is a χ^2 -distributed test statistic, with larger values indexing a violation of the null hypothesis that the SNP acts entirely through the common factor.

SNP heritability estimation

The proportion of the phenotypic variance that could be explained by the aggregated effect of all included SNPs (SNP-based heritability, h^2_{SNP}) was estimated using LDSC³¹. The analysis was performed using precomputed linkage disequilibrium scores from samples restricted to European ancestry in the 1000 Genomes Project⁹⁵, filtered for SNPs included in the HapMap 3 reference panel¹⁰¹. SNP heritability was estimated based on the slope of the LDSC, with heritability on the liability scale calculated assuming a 1% population prevalence of OCD¹. To omit a downward bias in our estimates of liability-scale heritability, following Grotzinger et al.¹⁰², we accounted for varying levels of ascertainment across cohorts in our meta-analysis by summing the effective sample sizes across the contributing cohorts and using that as the input sample size for LDSC. For conversion to the liability scale (1%), the sample prevalence was then specified as 0.5. The SNP heritability was calculated for the whole OCD sample as well as for ascertainment-specific subgroups.

Genetic correlations

We used cross-trait LDSC³¹, a method that computes genetic correlations between GWASs without bias from ancestry differences or sample overlap to calculate genetic correlations between the primary OCD meta-analysis and other phenotypes of interest. The selection of traits was based on phenotypic relevance and/or prior report of a genetic relationship with OCD. The genetic correlation between traits was based on the estimated slope from the regression of the product of Z scores from two GWASs on the linkage disequilibrium score and represents the genetic covariation between two traits based on all polygenic effects captured by the included SNPs. The genome-wide linkage disequilibrium information used by these methods was based on European populations from the HapMap 3 reference panel¹⁰¹, and GWAS summary statistics were filtered to only include SNPs that were part of the 1,290,028 HapMap 3 SNPs.

To ensure the internal consistency of the datasets included in our meta-analysis, we calculated genetic correlations between all cohorts we considered to have a sample size large enough for LDSC (effective sample size of $\geq 1,000$) and between the four ascertainment-specific subgroups.

We further calculated genetic correlations between OCD and 112 other disorders and traits. The source studies of the GWAS summary statistics can be found in Supplementary Table 18. As a follow-up, we also calculated genetic correlations between the 112 phenotypes and each ascertainment-specific subcohort and compared the genetic correlation patterns between the four groups. For all cross-phenotype genetic correlation analyses, we adjusted P values for multiple testing using the Benjamini–Hochberg procedure to control for the FDR (< 0.05).

Gene-based analyses

To match the significant SNPs to the genes for which they likely influence function, we conducted a series of positional and functional gene-mapping analyses. The positional mapping employed MBAT-combo¹⁹, while the functional mapping tested whether genetic variants associated with OCD were also associated with differential expression of nearby genes (within a 1-Mb window) using (1) TWAS²⁰ using PsychENCODE data and included colocalization with COLOC^{27,28}, and (2) SMR²² using whole-blood eQTL information and brain tissues from MetaBrain, alongside the HEIDI test, which tests for heterogeneity in GWAS signal and eQTL association. Furthermore, a PWAS was conducted. As a final step, genes within each locus were prioritized using PsyOPS²⁶, which integrates both positional and functional information. The details of each method are described below.

Positional gene mapping (MBAT-combo)

A gene-based analysis was conducted using mBAT-combo¹⁹ within GCTA version 1.94.1 (ref. 14). The European subsample ($n = 503$ individuals) from phase 3 of the 1000 Genomes Project⁹⁵ was used as the linkage disequilibrium reference panel with the fastBAT default linkage disequilibrium cutoff of 0.9 applied. After filtering SNPs with MAF > 0.01 , there were 6,629,124 SNPs for analysis in our sample. A gene list consisting of 19,899 protein-coding genes was used to map the base pair position of genes using genome build hg19 (see Supplementary Note 7 for details).

Functional gene mapping

Transcriptome-wide association study. We used TWAS FUSION²⁰ to perform a TWAS of OCD. We used brain gene expression weights from the PsychENCODE¹⁰³ and linkage disequilibrium information from the 1000 Genomes Project Phase 3 (ref. 95). TWAS FUSION uses reference linkage disequilibrium and reference gene expression panels with GWAS summary statistics to estimate the association between gene expression and OCD risk. These data were processed with the test statistics from the OCD GWAS to estimate the expression–GWAS association statistic. We corrected for multiple testing using Bonferroni correction.

We performed colocalization analyses using the COLOC R function^{27,28} implemented in TWAS FUSION. Colocalization is a Bayesian method used to calculate the posterior probabilities (PP) that individual lead SNPs within a significant TWAS locus are (1) independent (for example, two causal SNPs in linkage disequilibrium, one affecting transcription and one affecting OCD; PP₃) or (2) share the same associated variant (for example, a single causal SNP affects both transcription and OCD (PP₄)). We also performed a conditional analysis to determine whether identified associations represented independent associations. This was performed using the FUSION software, which jointly estimates the effect of all significant features within each locus by using residual SNP associations with OCD after accounting for the predicted expression of other features.

Summary-based Mendelian randomization. SMR²² was performed using default settings and eQTL meta-analysis summary statistics from European populations for whole blood from eQTLGen²³ and all five nervous system tissues from MetaBrain (basal ganglia, cerebellum, cortex, hippocampus and spinal cord)¹⁰⁴. The HEIDI test was performed alongside SMR to test for effect size heterogeneity between the GWAS and eQTL summary statistics. Both SMR and TWAS have a number of important assumptions and limitations, which we discuss in Supplementary Note 9.

Psychiatric omnilocus prioritization score

We used the gene prioritization method PsyOPS²⁶ to rank genes within genome-wide significant loci. This supervised approach integrates biological annotations on mutational intolerance, brain-specific expression and involvement in neurodevelopmental disorder for genes within significant loci. Genes with the top PsyOPS score within each locus

were used for further gene prioritization (Gene prioritization). In the instance where two genes in the same locus had the same PsyOPS score, the gene nearest the index SNP was prioritized.

Protein-wide association study

We performed a PWAS using protein expression data from human brain samples. Human brain proteome reference weight data were obtained using the Religious Orders Study and Rush Memory and Aging Project (ROS/MAP) and the Banner Sun Health Research Institute (Banner) study. The ROS/MAP proteomes were generated from the dIPFC of 376 participants of European ancestry and included 1,476 proteins with significant SNP-based heritability ($P < 0.01$). The Banner PWAS weights were generated from 152 individuals of European ancestry and included 1,147 proteins with significant SNP-based heritability. The PWAS was performed using the TWAS FUSION software²⁰ with linkage disequilibrium reference information from the 1000 Genomes Project Phase 3 (ref. 95). We corrected for multiple testing using Bonferroni correction.

Gene prioritization

We created a list of prioritized genes using both gene-based tests and colocalization–HEIDI filters. Results from each gene-based test were first restricted to protein-coding genes with unique gene identifiers based on the release from GENCODE (version 40) for hg19. The following criteria were then used to prioritize genes: (1) a significant (Bonferroni-corrected) association from at least two gene-based tests (mBAT-combo, TWAS FUSION, SMR or PsyOPS) and (2) evidence of colocalization ($\text{COLOC PP}_4 > 0.8$) and/or significant SMR association with HEIDI $P > 0.05$. Joint–conditional tests of association and significant PWAS associations were used as ancillary approaches to further annotate the prioritized gene list.

Tissue and cell type enrichment analysis

An analysis of tissue and cell type enrichment of OCD GWAS association signals was conducted using MAGMA (version 1.08)¹⁰⁵ and partitioned LDSC¹⁰⁶. We used the previously described approach²⁹ to determine gene expression specificity in bulk tissue RNA-seq data from 37 tissues in GTEx (version 8) and single-cell RNA-sequencing data from 19 regions in the mouse central and peripheral nervous systems³⁰. The analysis was limited to protein-coding genes with 1:1 orthologs between mice and humans. Gene expression in each tissue or cell type was calculated relative to total expression across all tissues or cell types. Enrichment analysis was performed on genes with the top 10% specificity values in each tissue or cell type, as previously defined²⁹.

To evaluate the enrichment of tissue- and cell type-specific genes in OCD genetic association signals, we applied MAGMA and partitioned LDSC. We restricted the analysis to summary statistics for SNPs with a high INFO score (> 0.6) and frequency in the entire cohort ($\text{MAF} > 0.01$). Using MAGMA (version 1.08), we tested whether genes with the top 10% specificity in a tissue or cell type showed enrichment in gene-level genetic associations for OCD, with the 1000 Genomes Phase 3 European sample genotypes serving as the linkage disequilibrium reference panel. We used standard gene boundaries (35 kb upstream of the transcription start site to 10 kb downstream of the transcription stop site). Partitioned LDSC was used to examine whether SNPs within 100-kb regions of the top 10% specifically expressed genes were enriched for SNP-based heritability for OCD. All results were corrected for multiple testing with an FDR threshold of 0.05.

SNP and gene findings in the context of previous analyses

Previously reported associations for significant SNPs (PheWAS).

Multiple resources were used to identify previously reported associations of our 30 significant SNPs with other phenotypes. We used the IEU Open GWAS project⁹², PheWAS analysis of GWAS ATLAS⁵² and the NHGRI-EBI GWAS Catalog⁹¹ and identified credible SNPs

through CAUSALdb⁹⁰. CAUSALdb estimates causal probabilities of all genetic variants in GWAS significant loci using three state-of-the-art fine-mapping tools including PAINTOR, CAVIARBF and FINEMAP^{107–110}. We used default settings for our CAUSALdb queries.

Lookup of previous OCD GWAS findings. We performed a lookup of SNPs identified to be significantly associated with OCD-related phenotypes in previous GWASs. Note that this is not an independent replication, as previous studies partially overlap with the cohorts included in this GWAS.

Overlap of previous rare coding variants in OCD and GWAS gene findings. We performed a bidirectional lookup, assessing (1) whether gene findings from our GWAS showed evidence for rare variant involvement and (2) vice versa, whether findings from rare variant testing showed evidence of common variant association in our GWAS.

First, we comprehensively assessed the overlap between 251 genes that we highlighted in our study as carrying common risk variation for OCD (Supplementary Table 14) and current gene-based summary statistics from OCD exome-sequencing data. We used data from Halvorsen et al.⁸⁸ because it is the largest published exome-sequencing study of OCD presently. The supplementary materials from that paper include de novo variant calls from 771 case trios and 1,911 controls (supplementary table 14 in ref. 88). We compared the burden of de novo variants, partitioned by variant annotation (synonymous, missense, loss of function) in trio cases versus trio controls within these 251 GWAS-prioritized genes. As described previously⁸⁸, we only included de novo variants that were in loci well covered in both case and control data ($\text{In_Jointly_Covered_Loci} = \text{TRUE}$). We also excluded all calls from quartet samples in ref. 88 ($\text{Cohort} = \text{"OCD_JHU_quartets"}$). For each of the four variant annotation classes, we compared the proportion of cases that had at least one qualifying de novo variant to the proportion of controls using a two-sided Fisher's exact test.

Second, as Halvorsen et al.⁸⁸ describe an overall excess of loss-of-function variants in OCD cases relative to controls specifically within loss-of-function intolerant genes (supplementary table 13 in ref. 88), we analyzed the overlap between those genes and our GWAS-derived genes. We looked up 200 genes with a probability of loss-of-function intolerance > 0.995 (derived from ref. 111) and effect size estimate > 1 . We further tested for a difference in the proportion of these pLI > 0.995 genes with effect size estimate > 1 versus ≤ 1 within the set of genes highlighted in the OCD GWAS ($n = 251$) versus outside this set using a two-sided Fisher's exact test.

Reporting summary

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

Data availability

The meta-analyzed summary statistics (not including 23andMe data) are available from the Psychiatric Genomics Consortium Download page (<https://www.med.unc.edu/pgc/download-results/>). In line with 23andMe regulations, 10,000 SNPs from the full GWAS including 23andMe are also being made available at <https://www.med.unc.edu/pgc/download-results/>. The full GWAS summary statistics for the 23andMe discovery dataset will be made available through 23andMe to qualified researchers under an agreement with 23andMe that protects the privacy of the 23andMe participants. Datasets will be made available at no cost for academic use. Please visit <https://research.23andme.com/collaborate/#dataset-access/> for more information and to apply to access the data. MVP summary statistics are made available through dbGAP request under accession https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs001672.v12.p1 phs001672.v12.p1.

Code availability

Core analysis code for RICOPILI can be found at <https://sites.google.com/a/broadinstitute.org/ricopili/>. This includes PLINK (<https://www.cog-genomics.org/plink2/>), EIGENSOFT (<https://www.hsph.harvard.edu/alkes-price/software/>), Eagle2 (<https://alkesgroup.broadinstitute.org/Eagle/>), Minimac3 (<https://genome.sph.umich.edu/wiki/Minimac3>), SHAPEIT3 (https://mathgen.stats.ox.ac.uk/genetics_software/shapeit/shapeit.html), METAL (https://genome.sph.umich.edu/wiki/METAL_Documentation) and LDSR (<https://github.com/bulik/ldsc>). MAGMA can be found at <https://ctg.cncr.nl/software/magma>. GenomicSEM, specifically the tutorial 'Models without Individual SNP effects' can be found here: <https://github.com/GenomicSEM/GenomicSEM/wiki/3.-Models-without-Individual-SNP-effects>. TWAS FUSION can be found at <http://gusevlab.org/projects/fusion/>. PWAS: for access to the protein weights, see <https://www.synapse.org/#!Synapse:syn24872746>. GCTA (mBAT-combo and COJO) can be found at <https://yanglab.westlake.edu.cn/software/gcta/#Overview>. LDSC and partitioned heritability can be found at <https://github.com/bulik/ldsc>. Additional code for data processing (for example, harmonization of summary statistics) can be found at <https://doi.org/10.6084/m9.figshare.28451894> (ref. 112).

References

93. DSM-5. *Diagnostic and Statistical Manual of Mental Disorders: DSM-5* 237–242 (American Psychiatric Association, 2013).
94. McCarthy, S. et al. A reference panel of 64,976 haplotypes for genotype imputation. *Nat. Genet.* **48**, 1279–1283 (2016).
95. The 1000 Genomes Project Consortium. A global reference for human genetic variation. *Nature* **526**, 68–74 (2015).
96. Lam, M. et al. RICOPILI: Rapid Imputation for Consortias Pipeline. *Bioinformatics* **36**, 930–933 (2020).
97. Willer, C. J., Li, Y. & Abecasis, G. R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190–2191 (2010).
98. Higgins, J. P. T. & Thompson, S. G. Quantifying heterogeneity in a meta-analysis. *Stat. Med.* **21**, 1539–1558 (2002).
99. Higgins, J. P. T., Thompson, S. G., Deeks, J. J. & Altman, D. G. Measuring inconsistency in meta-analyses. *BMJ* **327**, 557–560 (2003).
100. Turley, P. et al. Multi-trait analysis of genome-wide association summary statistics using MTAG. *Nat. Genet.* **50**, 229–237 (2018).
101. Altshuler, D. M. et al. Integrating common and rare genetic variation in diverse human populations. *Nature* **467**, 52–58 (2010).
102. Grotzinger, A. D., de la Fuente, J., Privé, F., Nivard, M. G. & Tucker-Drob, E. M. Pervasive downward bias in estimates of liability-scale heritability in genome-wide association study meta-analysis: a simple solution. *Biol. Psychiatry* **93**, 29–36 (2023).
103. Wang, D. et al. Comprehensive functional genomic resource and integrative model for the human brain. *Science* **362**, eaat8464 (2018).
104. de Klein, N. et al. Brain expression quantitative trait locus and network analyses reveal downstream effects and putative drivers for brain-related diseases. *Nat. Genet.* **55**, 377–388 (2023).
105. 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).
106. Finucane, H. K. et al. Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nat. Genet.* **47**, 1228–1235 (2015).
107. Benner, C. et al. FINEMAP: efficient variable selection using summary data from genome-wide association studies. *Bioinformatics* **32**, 1493–1501 (2016).
108. Chen, W. et al. Fine mapping causal variants with an approximate Bayesian method using marginal test statistics. *Genetics* **200**, 719–736 (2015).
109. Kichaev, G. et al. Integrating functional data to prioritize causal variants in statistical fine-mapping studies. *PLoS Genet.* **10**, e1004722 (2014).
110. Kichaev, G. & Pasaniuc, B. Leveraging functional-annotation data in trans-ethnic fine-mapping studies. *Am. J. Hum. Genet.* **97**, 260–271 (2015).
111. Lek, M. et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **536**, 285–291 (2016).
112. Mattheisen, M. PGC OCD3 code. *Figshare* <https://doi.org/10.6084/m9.figshare.28451894.v1> (2025).

Acknowledgements

We thank the research participants and employees of all cohorts included in this study for making this work possible. A list of members of the 23andMe Research Team, HUNT, CoGa and the MVP who contributed to this study can be found in Supplementary Note 1. We thank the research participants and employees of 23andMe for making this work possible. The Trøndelag Health Study (HUNT) is a collaboration between the HUNT Research Centre (Faculty of Medicine and Health Sciences, NTNU), Trøndelag County Council, Central Norway Regional Health Authority and the Norwegian Institute of Public Health. Genotype quality control and imputation were conducted by the K.G. Jebsen Center for Genetic Epidemiology, Department of Public Health and Nursing, Faculty of Medicine and Health Sciences, NTNU. HUNT analyses were performed in digital laboratories at HUNT Cloud, HUNT Research Centre Department of Public Health and Nursing, Faculty of Medicine and Health Sciences, NTNU, Trondheim, Norway. NORDiC is funded by National Institute of Mental Health (NIMH) R01 MH110427 (PI J.J.C.), NIMH R01 MH105500 (PI J.J.C.) and the Swedish Research Council grant 2015–02271 (PI D.M.-C.). NORDiC was further supported by the Swedish Research Council (grants 2012-07111 and 2018-02487), Swedish Research Council for Health, Working Life and Welfare 2018-00221 and the Center for Innovative Medicine (CIMED). We are deeply grateful for the study participants contributing to the NORDiC research. We thank the collection team that worked to recruit them: A. Juréus, J. Pege, M. Rådström, R. Satgunathan-Dawoud, M. Krestelica and B. Ohlander, as well as the data manager B. Iliadou. We also thank the National Quality Registry for Eating Disorders (RIKSÄT) for help with recruiting patients. We finally thank the BBMRI.se and KI Biobank at Karolinska Institutet for professional biobank services. Grant support for the MoBa team was also provided from RCN (273291, 262656, 248778, 223273) and the KG Jebsen Stiftelsen. MoBa is supported by the Norwegian Ministry of Health and Care Services and the Ministry of Education and Research. We are grateful to all the participating families in Norway who take part in this ongoing cohort study. The AGDS was primarily funded by National Health and Medical Research Council (NHMRC) of Australia grant 1086683. This work was further supported by NHMRC grants 1145645, 1078901 and 1087889. We thank all the people who helped in the conception, implementation, beta testing, media campaign and data cleaning of the AGDS data. We specifically acknowledge D. Nyholt for advice on using the PBS for research; K. Kendler, P. Sullivan, A. McIntosh and C. Lewis for input on the questionnaire; L. Nunn, M. Ferguson, L. Winkler and N. Garden for data and sample collection; N. Zmicerevska, A. Nichles and C. Brennan for participant recruitment support; J. Davies, L. Lowrey and V. Antonini for support with IT aspects; and V. Morgan and K. Kirkby for help with the media campaign. We thank VIVA! Communications for their effort in promoting the study. We also acknowledge D. Whiteman and C. Olsen from QSkin. The work done by the EstBB team has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement 847776 (CoMorMent). Data analysis for EstBB was carried out in part at the High-Performance Computing Center of the University of Tartu. We thank participants, families and staff of primary and secondary schools who kindly

contribute to this research (M.R., Metal-Cat and INSchool). EGOS was supported by a grant from the Beatrice and Samuel A. Seaver Foundation to D.E.G. The genotyping of HUNT was financed by the National Institutes of Health (NIH), the University of Michigan, the Norwegian Research Council and Central Norway Regional Health Authority and the Faculty of Medicine and Health Sciences, NTNU. This research is based in part on data from the MVP, Office of Research and Development, Veterans Health Administration and was supported by awards CSP575b, IO1CX001849-01 and 1P1HX002375 and the National Center for PTSD Research. The MVP was supported by funding from the Department of Veterans Affairs Office of Research and Development, USVA, grants CSP575B and IO1CX001849, MVP-025 and the VA Cooperative Studies Program study no. 575B, the VA National Center for PTSD Research and the West Haven VA Mental Illness Research, Education and Clinical Center and by NIH grant R01AA026364 (J. Gelernter). This publication does not represent the views of the Department of Veterans Affairs or the United States government. The EPOC study was funded by the Deutsche Forschungsgemeinschaft (DFG; KA815/6-1 and WA731/10-1). LifeGene was supported by the Swedish Research Council, Karolinska Institutet–Stockholm County Council research grants, AFA Insurance and the Torsten and Ragnar Söderbergs Foundation. GENOS was supported by the German Research Foundation (GR 1912/1-1). The OCD Collaborative Genetics Association Study (OC GAS) is a collaborative research study and was funded by the following NIMH grant numbers: MH071507, MH079489, MH079487, MH079488 and MH079494. This work (OC GAS and IOCDF) is supported by the Netherlands Organization for Scientific Research—Gravitation project ‘BRAINS CAPES: a Roadmap from Neurogenetics to Neurobiology’ (024.004.012) and the European Research Council advanced grant ‘From GWAS to Function’ (2018-ADG 834057). The OC GAS and IOCDF samples are supported through NIMH grant numbers MH071507 (G.N.), MH079489 (D.A.G.), MH079487 (J.T.M.), MH079488 (A.J.F.) and MH079494 (J.A.K.). The iPSYCH team was supported by grants from the Lundbeck Foundation (R102-A9118, R155-2014-1724 and R248-2017-2003), the NIH/NIMH (1R01MH124851-01 to A.D.B.) and the universities and university hospitals of Aarhus and Copenhagen. The Danish National Biobank resource was supported by the Novo Nordisk Foundation. High-performance computer capacity for handling and statistical analysis of iPSYCH data in the GenomeDK HPC facility was provided by the Center for Genomics and Personalized Medicine and the Centre for Integrative Sequencing, iSEQ, Aarhus University, Denmark (grant to A.D.B.). A.D.B. was also supported by the EU’s HORIZON-HLTH-2021-STAYHLTH-01 program, project number 101057385: Risk and Resilience in Developmental Diversity and Mental Health (R2D2-MH). Metal-Cat and INSchool were supported by the Agència de Gestió d’Ajuts Universitaris i de Recerca (2017SGR-1461, 2021SGR-00840), the Instituto de Salud Carlos III (PI20/00041, PI23/00404 and PI23/00026), the European Regional Development Fund and the ECNP Network (‘ADHD across the Lifespan’ and ‘la Marató de TV3’, 202228-30 and 202228-31). BioVU: CTSA (S.D., Vanderbilt Resources) was supported by the National Center for Research Resources, grant UL1 RR024975-01 and is now at the National Center for Advancing Translational Sciences (grant 2 UL1 TR000445-06). The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH. The dataset(s) used for the analyses (BioVU) described were obtained from Vanderbilt University Medical Center’s BioVU, which is supported by numerous sources: institutional funding, private agencies and federal grants. These include the NIH-funded Shared Instrumentation Grant S1ORR025141 and CTSA grants UL1 TR002243, UL1 TR000445 and UL1 RR024975. Genomic data are also supported by investigator-led projects that include U01HG004798, R01NS032830, RC2GM092618, P50GM115305, U01HG006378, U19HL065962 and R01HD074711 and additional funding sources listed at

<https://victor.vumc.org/biovu-funding/>. S.A. acknowledges a Miguel Servet contract (CP22/00026) awarded by the Instituto de Salud Carlos III and cofunded by the European Union Fund: Fondo Social Europeo Plus, FSE+. HYPERGENES and InterOmics cohorts provided controls of Italian origin for the present study. J.C.-D. acknowledges her contract from the Network Center for Biomedical Research (CIBER). R.D. acknowledges the Clinical Investigation Centre, Robert Debré Hospital. INSERM at APHP granted the study. B.T.F. acknowledges the Anorexia Nervosa Genetics Initiative, an initiative of the Klarman Family Foundation. J.H. acknowledges the Trond Mohn Foundation, Bergen, Norway. C. Lochner acknowledges the South African Medical Research Council and the National Research Foundation for their support. T.V.F.: research reported in this publication was supported by the NIMH of the NIH under award number R01MH114927. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH. N.G.M. has received funding from a project grant from the Australian NHMRC. The Research Council of Norway supported H.A., A.H. and T.R.-K. (274611). A. Havdahl was also supported by South East Norway Health Authority (2020022). Z.F.G. is supported by an Australian NHMRC EL1 Investigator Grant (2034743) and NIH/NIA grant AG068026. M.G. received support from the following grants (J. Gelernter): CSP575b, IO1CX001849-01, 1P1HX002375, the National Center for PTSD Research, 5R01DA054869-01. A. Abdellaoui was supported by the Foundation Volksbond Rotterdam. T.B. is supported by NIMH grant 7R01MH103657 (GPC-OC D). J.R.C.: this study represents independent research partly funded by the National Institute for Health Research Maudsley Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King’s College London. The views expressed are those of the authors and not necessarily those of the NHS, the National Institute for Health Research or the Department of Health and Social Care. C.B. was supported by grant EU FP7-HEALTH-2007-A-201550 and grant MIUR-CNR PBO5. E.M.B. was supported by NHMRC project grant 1145645 and the University of Queensland Health Research Accelerator Program. C.C. was supported by grant K99MH128540-01A1. V.C. was supported by the Italian Ministry of Health grant RC18-19-20-21/A. M.A.G. was supported by NIMH K23 MH066284. J.H. was supported by Stiftelsen KG Jebsen (SKGJ MED-02). K. Hagen was supported by the Trond Mohn Foundation. E.K.K. was supported by NIH R21 MH109938. P.S.N. was supported by R01MH071507. Fabrizio and Federica Piras are supported by the Italian Ministry of Health RC18-19-20-21/A grant. This work was in part supported by the German Research Foundation (DFG) grants RA1971/8-1 and RA1971/7-1 and by the Bundesministerium für Bildung und Forschung grant 01ED2007A to A. Ramirez. S. Ripke was supported by research grant 1R01MH124873-01. M.S.A. was supported by the Instituto de Salud Carlos III (P19/01224, PI22/00464 and CP22/00128) and the European Regional Development Fund. A. Agrawal was supported by grant U10AA008401. P.A. was supported by the Spanish Ministry of Science, Innovation and Universities (ISCIII PI22/00752) and Fundació La Marató 202201-30. C.M.B. was supported by R01 MH124871 (Sullivan and Bulik) PGC4. H.E. was supported by grant U10AA008401. D.A.G. was supported by the NIMH (OC GAS and OCGS). G.L.H. was supported by the NIMH (R01 MH58376, K20 MH01065, R01 MH101493, R01 MH085321). N.K. has received funding from DFG KA815/6-1. S.E.M. is supported by an Australian NHMRC Investigator Grant (APP1172917). B.M.N. is funded by grant R01MH124851. M.P. and C. Pato have received support from R01MH103657 and R01MH079494 from the NIMH and the Della Martin Foundation, Los Angeles, CA. J.P. has received support through the NIMH (R01MH50214: Collaborative OCD Genetics Study (G.N., Pi; J.T.M., UCLA Pi)). M.A. Richter was supported by funding from the Canadian Institutes for Health Research and the Ontario Mental Health Foundation. D.R.R. was supported by NIMH R01MH059299. J.F.S. was supported by NIMH grant number

MH071507. G.S. is supported by the Italian Ministry of Health RC18-19-20-21/A grant. E.A.S. collected data as part of the NIH grant 1R01MH093381. O.A.A. (MoBa) has received grant support from RCN (324499, 273291, 262656, 248778, 223273), KG Jebsen Stiftelsen and NordForsk 164218. J. Kaprio has been supported by the Academy of Finland (grant 336823). P.D.A. is supported by the Alberta Innovates Translational Health Chair in Child and Youth Mental Health. D.E.G. is supported by grant MH124679-01. J.A.K. is supported by the grants R01MH103657 and R01MH079494 from the NIMH and the Della Martin Foundation, Los Angeles, CA. K.J.V. is supported by the Foundation Volksbond Rotterdam. L.K.D. was supported by grants from the NIH including R01NS102371, R01MH113362, R01MH118223, R01NS105746 and R56MH120736. J.S. was supported by an NIH Training Grant in Human Genetics (2T32GM080178). J.J.C. was supported by NIH grants R01 MH105500 and R01 MH110427. M.B.S. has been funded by the US Veterans Affairs Administration.

Author contributions

J.M. Scharf, M.B.S., J. Gelernter, C.A.M., E.M.D. and M. Mattheisen designed the study. N.I.S., Z.F.G., M.G., D.Y. and M.W.H. conducted data analysis. N.I.S., Z.F.G., M.G., D.Y., M.W.H., A. Abdellaoui, C.R.-F., J.M. Sealock, T.B., J.R.C., B. Mahjani, J.G.T., K.B., C.L.B., J.J.L., G.Z., S.A., C.A., K.D.A., J.B., N.B., C.B., J.B.N., O.J.B., D.B., M.H.B., S.B., R.B., M.B., B.P.B., H.B., J.D.B., J.B.-G., E.M.B., J.C.-D., B.C., A. Camarena, C.C., A. Carracedo, M.C., M.C.C., V.C., E.H.C., J.C., B.A.C., E.J.D.S., R.D., S.D., J.A.E., X.E., M.J.F., B.T.F., L.G., C. Gironde, F.S.G., M.A.G., J. Grove, W.G., J.H., K. Hagen, K. Harrington, A.H., K.D.H., A.G.H., D.H., C.H., M.J., E.J., E.K.K., K. Kelley, J. Klawohn, J.E.K., K. Krebs, C. Lange, N.L., D. Levey, K.L.-T., F.M., B. Maher, B. Mathes, E.M., N.M., N.C.M., S.M., E.C.M., M. Mulhem, P.S.N., E.L.N., K.S.O'C., L.O., O.T.O., T.P., N.L.P., Fabrizio Piras, Federica Piras, S.P., R.R., A. Ramirez, S. Rauch, A. Reichenberg, M.A. Riddle, S. Ripke, M.C.R., A.S.S., M.A.S., A.H.S., L.G.S., J.S., M.S.A., L.F.T., E.T., H.V., N.v.K., J.V.-V., N.N.V., C.P.W., Y.W., J.R.W., B.S.W., Y.Y., H.Z., A. Agrawal, P.A., G. Berberich, K.K.B., C.M.B., D.C., D.D., V.E., H.E., P.F., T.V.F., A.J.F., J.M.G., D.A.G., H.J.G., B.D.G., G.L.H., I.B.H., D.M.H., N.K., J. Kennedy, D. Lai, M. Landén, S.L.H., M. Leboyer, C. Lochner, J.T.M., S.E.M., P.B.M., B.M.N., H.N., M. Nordentoft, M.P., C. Pato, D.L.P., J.P., C. Pittenger, D.P., J.A.R.-Q., S.A.R., M.A. Richter, D.R.R., S. Ruhrmann, J.F.S., S.S., P.S., G.S., D.J. Stein, S.E.S., E.A.S., B.E.S., M.T., T.W., O.A.A., A.D.B., S.W., K. Hveem, B.K.H., C.R., N.G.M., L.M., O.M., T.R.-K., M.R., G.K., D.M.-C., K.D., E.G., M.W., J.-A.Z., G. Breen, G.N., J. Kaprio, P.D.A., D.E.G., J.A.K., H.A., K.J.V., L.K.D., D.J. Smit, J.J.C., J.M. Scharf, M.B.S., J. Gelernter, C.A.M., E.M.D. and M. Mattheisen provided samples and/or processed individual cohort data. N.I.S., Z.F.G., M.G., D.Y., M.W.H., J.M. Scharf, M.B.S., J. Gelernter, C.A.M., E.M.D. and M. Mattheisen wrote the paper and formed the core revision group. J.M. Scharf, M.B.S., J. Gelernter, C.A.M., E.M.D. and M. Mattheisen supervised and directed the study. All authors discussed the results and approved the final version of the manuscript.

Funding

Open access funding provided by Humboldt-Universität zu Berlin.

Competing interests

C. German is employed by and holds stock or stock options in 23andMe. E.L.N. is on the scientific advisory board for Myriad Genetics and the medical advisory board for the Tourette Association of America and received clinical trial funding from Emalex and Octapharma Pharmaceuticals. J.V.-V. has served on advisory boards or consulted with Roche, Novartis and SynapDx; received research funding from Roche, Novartis, SynapDx, Seaside Therapeutics, Forest, Janssen, Acadia, Yamo and MapLight; and received stipends for editorial work from Wiley and Springer. J.R.W. is a current employee and shareholder of Takeda Pharmaceuticals and a past employee

and shareholder of F. Hoffmann-La Roche, Pfizer and Nestle Health Science. C.M.B. reports Pearson (author, royalty recipient). P.F. reports no conflict of interest regarding this study and reports having received financial support and served on the advisory board for Richter, Recordati, Boehringer Ingelheim, Otsuka, Janssen and Lundbeck. H.J.G. has received travel grants and speaker's honoraria from Fresenius Medical Care, Neuraxpharm, Servier and Janssen-Cilag as well as research funding from Fresenius Medical Care. I.B.H. is the co-director of health and policy at the Brain and Mind Centre of the University of Sydney, Australia. The Brain and Mind Centre operates early-intervention youth services at Camperdown under contract to headspace. I.B.H. has previously led community-based and pharmaceutical industry-supported (Wyeth, Eli Lilly, Servier, Pfizer, AstraZeneca, Janssen-Cilag) projects focused on the identification and better management of anxiety and depression. He is the chief scientific advisor to and a 3.2% equity shareholder in InnoWell, which aims to transform mental health services through the use of innovative technologies. B.M.N. is a member of the scientific advisory board at Deep Genomics and Neumora. C. Pittenger consults and/or receives research support from Biohaven Pharmaceuticals, Freedom Biosciences, Ceruvia Lifesciences, Transcend Therapeutics, UCB BioPharma and F-Prime Capital Partners. He owns equity in Alco Therapeutics. These relationships are not related to the current work. D.J. Stein has received consultancy honoraria from Discovery Vitality, Johnson & Johnson, Kanna, L'Oreal, Lundbeck, Orion, Sanofi, Servier, Takeda and Vistagen. E.A.S. reports receiving research funding to his institution from the Ream Foundation, the International OCD Foundation and the NIH. He was formerly a consultant for Brainsway and Biohaven Pharmaceuticals in the past 12 months. He owns stock less than \$5,000 in nView-Proem for distribution related to the YBOCS scales. He receives book royalties from Elsevier, Wiley, Oxford, the American Psychological Association, Guildford, Springer, Routledge and Jessica Kingsley. O.A.A. reports being a consultant to Cortechs.ai and Precision Health and speaker honoraria from Otsuka, Lundbeck, Sunovion and Janssen. A.D.B. has received a speaker fee from Lundbeck. D.M.-C. receives royalties for contributing articles to UpToDate and Wolters Kluwer Health and personal fees for editorial work from Elsevier, all unrelated to the current work. M.B.S. has in the past 3 years received consulting income from Acadia Pharmaceuticals, Big Health, Biogen, Bionomics, Boehringer Ingelheim, Clexio, Eisai, EmpowerPharm, Engrail Therapeutics, Janssen, Jazz Pharmaceuticals, NeuroTrauma Sciences, Otsuka, PureTech Health, Sage Therapeutics, Sumitomo Pharma and Roche-Genentech. M.B.S. has stock options in Oxeia Biopharmaceuticals and EpiVario. He has been paid for his editorial work on *Depression and Anxiety* (editor in chief), *Biological Psychiatry* (deputy editor) and UpToDate (co-editor in chief for psychiatry). J. Gelernter is paid for editorial work by the journal *Complex Psychiatry*. P.A. has received funding from Biohaven, Boston Scientific and Medtronic. All other authors report no conflicts of interest.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41588-025-02189-z>.

Correspondence and requests for materials should be addressed to Nora I. Strom or Manuel Mattheisen.

Peer review information *Nature Genetics* thanks Sarah Gagliano Taliun and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

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

Reporting Summary

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

Statistics

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

n/a Confirmed

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

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection No software was used for data collection. For individual studies please refer to their respective publications.

Data analysis This is fully described in the Online Methods and associated Supplementary Information. In brief: Additional quality control, a ssociation analyses, a nd polygenic risk scoring was done either using the Ricopili (v 1.118b) pipeline: <https://github.com/Nealelab/ricopili>, w hich relies on the following software: Eigensoft 6.0.1 (incl. smartPCA), P link 1.9, M ETAL 2011-03-25 or in other samples with comparable software tools and approaches (more details can be found in the Supplemental Table S1 and the Supplemental Text. SNP-based finemapping was performed using GCTA-COJO (as part of the GCTA v1.94.1 release), M TAG v1.08 was used for multivariable GWAS or the respective subgroups. Genomic SEM v 0.04 was used to run structural equation analyses, L DSC (v 1.0.1; <https://github.com/bulik/ldsc>) was used to calculate all heritability estimates and genetic correlations. Gene-based analyses were performed using MBAT-COMBO (within the GCTA package v1.94.1), TWAS Fusion (<http://gusevlab.org/projects/fusion/>; accessed Feb 2023), SMR v 1.3.1 (<https://yanglab.westlake.edu.cn/software/smr/#Overview>). Tissue and cell-type enrichment of OCD GWAS association signals was conducted using MAGMA (v1.08). R v3.4 was used in general for statistical analyses and plotting (<https://www.Rproject.org>) MiXER v1.2 was used to calculate the number of causal variants accounting for 90% of the SNP heritability in OCD. Analyses for PGC analyses were conducted on the SurfSara Lisa computing infrastructure and for all other analyses on the Bianca server as part of the Uppsala Multidisciplinary Center for Advanced Computational Science.

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

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

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

The meta-analyzed summary statistics (not including 23andMe data) will be made available via the Psychiatric Genomics Consortium Download page (<https://www.med.unc.edu/pgc/download-results/>). In line with 23andMe regulations, 10,000 SNPs from the full GWAS incl. 23andMe, will be made available. The full GWAS summary statistics for the 23andMe discovery data set will be made available through 23andMe to qualified researchers under an agreement with 23andMe that protects the privacy of the 23andMe participants. Datasets will be made available at no cost for academic use. Please visit <https://research.23andme.com/collaborate/#dataset-access/> for more information and to apply to access the data. MVP summary statistics are made available through dbGAP request under accession phs001672.v12.p1.

Research involving human participants, their data, or biological material

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

Reporting on sex and gender

Sex and Gender based analyses were not performed as a consequence of missing samples from the overall collection. We are currently running such analyses (sex-stratified analyses) in a subset of the overall sample and will report in an upcoming publication on the findings. For most samples the number of female and male cases (and controls) were not available and therefore not reported.

Reporting on race, ethnicity, or other socially relevant groupings

Analysis are limited to European ancestry samples, no race, ethnicity or other socially relevant groupings have been made other than those potentially (hiddenly) associated with our sub-groupings for our main phenotype related ascertainment (i.e., self-report, Biobank, comorbid, clinical groups).

Population characteristics

This is a meta-analysis of existing (in part unpublished) summary statistics. Details for the respective individual studies are provided in the Supplemental material and vary. The most common grouping exists for genotyping batches (especially for the large scale cohort and biobanking studies (including 23andMe)).

Recruitment

Details are provided in the Supplemental material and vary for the different studies that are part of the meta-analysis. In brief, samples were either ascertained in clinical setting, as part of larger biobanking efforts, via 23andMe, or as part of studies looking into comorbid phenotypes (such as MDD, ASD, and ADHD).

Ethics oversight

The details of the IRB/oversight body that provided approval or exemption for the research described are given below: were obtained from the Hopkins Medicine Institutional Review Boards. OCGAS-ab: The study was approved by REB at Hospital for Sick Children. Ethics approvals for the OCGAS study were obtained from the Hopkins Medicine Institutional Review Boards. OCGAS-gh: Informed written consent was obtained in all cases by the participants or their parents. The study was approved by the ethical commissions of all involved universities in accordance with the latest version of the Declaration of Helsinki, including an ethical permission granted by the Ethic Committees from Aachen, Wuerzburg, Marburg, Freiburg, and the Cantonal Ethic Commission of Zurich (Ref. Nr. 39/97, 140/3 and EK: KEK-ZHNR. 2010-0340/3). OCD-WWF: The study was approved by the ethics committee of the University of Wuerzburg, Germany and was conducted according to the ethical principles of the Helsinki Declaration. All patients gave written informed consent prior to participation. Psych_Broad: Both studies have been approved by the Clinical Research Ethics Committee (CREC) of Hospital Universitari Vall d'Hebron. All methods were performed in accordance with the relevant guidelines and regulations and written informed consent was obtained from participant parents before inclusion into the study. UKBB: Research on the UK Biobank is conducted under a generic Research Tissue Bank approval from the UK North West Multi-centre Research Ethics Committee (MREC). This research was approved to be conducted under that approval by the governing Research Ethics Committee of the UK Bio bank. The analyses in this paper were performed under an approved extension to project 16577. Yale-Penn: Participants were recruited from eastern U.S sites and provided written informed consent as approved by the institutional review board at each site. 23andMe: Participants provided informed consent and volunteered to participate in the research on line, under a protocol approved by the external AAHRPP-accredited IRB, Ethical & Independent (E&I) Review Services. As of 2022, E&I Review Services is part of Salus IRB (<https://www.versiticlinicaltrials.org/salusirb>). AGDS: All study protocols were approved by the QIMR Berghofer Medical Research Institute Human Research Ethics Committee. The protocol for approaching participants through the DHS, enrolling them in the study, and consenting for all phases of the study (including invitation to future related studies) and accessing MBS and PBS records was approved by the Ethics Department of the Department of Human Services. BioVU: The Vanderbilt University Medical Center Institutional Review Board oversees BioVU and approved this project (IRB201609). COGA: Institutional review boards at all sites approved the study and all participants provided informed consent. EGOS: Ethical approvals were obtained from the Institutional Review Board (IRB) at the Icahn School of Medicine at Mount Sinai, New York, NY, and the Regional Ethical Review Board in Stockholm. EPOC: The study was in accordance with the revised Declaration of Helsinki and approved by the local ethics committees of the Charité University Medicine Berlin and the University Hospital Bonn. Est BB: At recruitment, participants signed a consent allowing follow-up linkage of their electronic health records (EHRs), thereby providing a longitudinal collection of their phenotypic information. FinnGen: The Ethical Review Board of the Hospital District of Helsinki and Uusimaa approved the FinnGen study protocol Nr. HUS/990/2017. The FinnGen project was approved by Finnish Institute for Health and Welfare (THL), approval numbers THL/2031/6.02.00/2017, amendments THL/341/6.02.00/2018, THL/2222/6.02.00/2018, and THL/283/6.02.00/2019. HUNT:

The HUNT study was approved by the Regional Committee for Medical and Health Research Ethics, Norway (2015/575). IOCDF/OCGAS: This work was approved by the relevant IRBs at all participating sites, and all participants provided written informed consent. iPSYCH: The study was approved by the Regional Scientific Ethics Committee in Denmark and the Danish Data Protection Agency. MVP: The U.S. Department of Veterans Affairs (VA) Million Veteran Program (MVP) is collecting genetic and electronic health record (EHR) data in the U.S with ethical approval given by the Central VA Institutional Review Board (IRB) and site-specific IRBs. All relevant ethical regulations for work with human subjects were followed in the conduct of the study, and informed consent was obtained from all participants. MoBa: The establishment of MoBa and initial data collection was based on a license from the Norwegian Data Protection Agency and approval from The Regional Committees for Medical and Health Research Ethics. The MoBa cohort is now based on regulations related to the Norwegian Health Registry Act. NORDIC-SWE: This study was approved by the Regional Ethics Committee, Stockholm (EPN Stockholm) and the Institutional Review Board (IRB) at the University of North Carolina at Chapel Hill and all subjects provided informed consent. NORDIC-NOR: The NORDIC-NOR study was approved by the Norwegian Regional Committee for Medical and Health Research Ethics (IRB00001872 REK West) under project number 2018/52 REK Vest (PI: Bjarne Hansen) and project number: 2014/75 REK Vest (PI: Jan Haavik) and all subjects provided informed consent. OCGAS-all: Ethics approvals for the OCGAS study were obtained from the Hopkins Medicine Institutional Review Boards, the Butler Institutional Review Board, the UCLA Institutional Review Boards, the Mass General Brigham Human Research Committee, the Columbia University Institutional Review Boards, and the National Institutes of Health Institutional Review Board (NIH IRB). OCGAS-nestadt: Ethics approvals for the OCGAS study were obtained from the Hopkins Medicine Institutional Review Boards. OCGAS-ab: The study was approved by REB at Hospital for Sick Children. Ethics approvals for the OCGAS study were obtained from the Hopkins Medicine Institutional Review Boards. OCGAS-gh: Informed written consent was obtained in all cases by the participants or their parents. The study was approved by the ethical commissions of all involved universities in accordance with the latest version of the Declaration of Helsinki, including an ethical permission granted by the Ethic Committees from Aachen, Wuerzburg, Marburg, Freiburg, and the Cantonal Ethic Commission of Zuerich (Ref. Nr. 39/97, 140/3 and EK: KEK-ZHNr. 2010-0340/3). OCD-WWF: The study was approved by the ethics committee of the University of Wuerzburg, Germany and was conducted according to the ethical principles of the Helsinki Declaration. All patients gave written informed consent prior to participation. Psych_Broad: Both studies have been approved by the Clinical Research Ethics Committee (CREC) of Hospital Universitari Vall d'Hebron. All methods were performed in accordance with the relevant guidelines and regulations and written informed consent was obtained from participant parents before inclusion into the study. UKBB: Research on the UK Biobank is conducted under a generic Research Tissue Bank approval from the UK North West Multi-centre Research Ethics Committee (MREC). This research was approved to be conducted under that approval by the governing Research Ethics Committee of the UK Biobank. The analyses in this paper were performed under an approved extension to project 16577. Yale-Penn: Participants were recruited from eastern U.S sites and provided written informed consent as approved by the institutional review board at each site.

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

Field-specific reporting

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

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

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Life sciences study design

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

Sample size	No sample size calculation was made. Individual studies were selected for inclusion if they had at least an effective samples size of 100. Sample size of the meta-analysis is a result of inclusion of all world-wide available samples of European ancestry at the time of study,
Data exclusions	Within each analyzed cohort we aimed at analyzing genetically homogeneous samples of unrelated individuals. Related individuals were excluded based on Identity by State analyses (pseudo controls were used for trios) and genetic outliers were excluded based on principal component analyses.
Replication	Replication in a narrow sense could not be provided given that the meta-analysis comprised of most if not all samples currently available for research into the genetics of OCD (at the time of analysis). Consistency was checked via different sensitivity analyses using sub-groupings, including genomic SEM analyses, MTAG analyses, internal genetic correlation and other approaches. Our analyses did not reveal a concerning amount of heterogeneity in the samples.
Randomization	This study utilized naturally occurring genetic variation as the experimental condition. Because an individual's genotype is fixed at conception, we could not randomly assign participants to experimental groups. Instead, we examined the association between pre-determined genetic profiles and the observed phenotype of OCD.
Blinding	Blinding is not relevant to the current study. Samples were not allocated to different conditions by the researchers, and the phenotype ascertainment process is fully separate from the genotyping process.

Reporting for specific materials, systems and methods

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

Materials & experimental systems

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern
<input checked="" type="checkbox"/>	<input type="checkbox"/> Plants

Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

Plants

Seed stocks	Report on the source of all seed stocks or other plant material used. If applicable, state the seed stock centre and catalogue number. If plant specimens were collected from the field, describe the collection location, date and sampling procedures.
Novel plant genotypes	Describe the methods by which all novel plant genotypes were produced. This includes those generated by transgenic approaches, gene editing, chemical/radiation-based mutagenesis and hybridization. For transgenic lines, describe the transformation method, the number of independent lines analyzed and the generation upon which experiments were performed. For gene-edited lines, describe the editor used, the endogenous sequence targeted for editing, the targeting guide RNA sequence (if applicable) and how the editor was applied.
Authentication	Describe any authentication procedures for each seed stock used or novel genotype generated. Describe any experiments used to assess the effect of a mutation and, where applicable, how potential secondary effects (e.g. second site T-DNA insertions, mosaicism, off-target gene editing) were examined.