# Shared genetic architecture between schizophrenia and subcortical brain volumes implicates early neurodevelopmental processes and brain development in childhood

Weiqiu Cheng (PhD)[1], Dennis van der Meer (PhD)[1,2], Nadine Parker (PhD)[1], Guy Hindley (MBBS MRes)[1,3], Kevin S. O'Connell (PhD)[1], Yunpeng Wang (PhD)[1,4], Alexey A. Shadrin (PhD)[1], Dag Alnæs (PhD)[1], Shahram Bahrami (PhD)[1], Aihua Lin (PhD)[1], Naz Karadag (MSc)[1], Børge Holen (MD)[1], Sara Fernandez-Cabello (PhD)[1], Chun-Chieh Fan (PhD)[5-7], Anders M. Dale (PhD)[5-8], Srdjan Djurovic (PhD)[9,10], Lars T. Westlye (PhD)[1,11], Oleksandr Frei (PhD)[1], Olav B. Smeland (MD, PhD)[1], Ole A. Andreassen (MD, PhD)[1]

[1] NORMENT, Division of Mental Health and Addiction, Oslo University Hospital & Institute of Clinical Medicine, University of Oslo, Oslo, Norway

[2] School of Mental Health and Neuroscience, Faculty of Health, Medicine and Life Sciences, Maastricht University, The Netherlands

[3] Psychosis Studies, Institute of Psychiatry, Psychology and Neurosciences, King's College London, United Kingdom

[4] Centre for Lifespan Changes in Brain and Cognition (LCBC), Department of Psychology, University of Oslo, Oslo, Norway

[5] Center for Multimodal Imaging and Genetics, University of California San Diego, La Jolla, CA 92093, United States of America

[6] Department of Psychiatry, University of California, San Diego, La Jolla, CA, United States of America

[7] Department of Neurosciences, University of California San Diego, La Jolla, CA 92093, United States of America

[8] Department of Radiology, University of California, San Diego, La Jolla, CA 92093, United States of America

[9] Department of Medical Genetics, Oslo University Hospital, Oslo, Norway

[10] NORMENT, Department of Clinical Science, University of Bergen, Bergen, Norway

[11] Department of Psychology, University of Oslo, Oslo, Norway

* Corresponding author: Weiqiu Cheng (PhD), and Ole A. Andreassen (MD, PhD)

E-mail: weiqiu.cheng@medisin.uio.no and ole.andreassen@medisin.uio.no;

Address: Kirkeveien 166, 0450 Oslo, Norway.

## Abstract

Patients with schizophrenia have consistently shown brain volumetric abnormalities, implicating both etiological and pathological processes. However, the genetic relationship between schizophrenia and brain volumetric abnormalities remains poorly understood. Here, we applied novel statistical genetic approaches (MiXeR and conjunctional false discovery rate analysis) to investigate genetic overlap with mixed effect directions using independent genome-wide association studies of schizophrenia (n=130,644) and brain volumetric phenotypes, including subcortical brain and intracranial volumes (n=33,735). We found brain volumetric phenotypes share substantial genetic variants (74%~96%) with schizophrenia, and observed 107 distinct shared loci with sign consistency in independent samples. Genes mapped by shared loci revealed (1) significant enrichment in neurodevelopmental biological processes, (2) three co-expression clusters with peak expression at the prenatal stage, and (3) genetically imputed thalamic expression of *CRHR1* and *ARL17A* was associated with the thalamic volume as early as in childhood. Together, our findings provide evidence of shared genetic architecture between schizophrenia and brain volumetric phenotypes and suggest altered early neurodevelopmental processes and brain development in childhood may be involved in schizophrenia development.

## Introduction

Schizophrenia is a common psychiatric disorder with a significant contribution to the global burden of disease[1], such as a 14.5-year reduction in life expectancy[2]. One of the most widely accepted pathological hypotheses of schizophrenia is the neurodevelopmental hypothesis[3], which postulates that the emergence of schizophrenia is the ultimate outcome of an abnormal brain development trajectory which begins many years before the occurrence of clinical symptoms, and is caused by genetic and early life environmental factors. Subcortical brain structures have been suggested to play an important role in schizophrenia etiopathology[4], supported by evidence of altered dopamine transmission within thalamus and amygdala[5], and predominant localization of antipsychotic-binding receptors, such as dopamine D2 receptors, in striatum[6]. Magnetic resonance imaging (MRI) has established brain volumetric alterations, including altered subcortical and intracranial volumes (ICV), as a feature of schizophrenia[7]. Compared to healthy controls, schizophrenia patients showed smaller hippocampus, amygdala, thalamus, accumbens, and ICVs, as well as a larger pallidum[7] on the group level. Brain volumetric alterations, such as reduced hippocampal volume, were also observed in unaffected individuals having first- or second-degree relatives with schizophrenia[8, 9], and in association with a higher polygenic risk score of schizophrenia[10]. However, the causative nature of the structural brain abnormalities in schizophrenia remains elusive.

As both schizophrenia and brain volumetric phenotypes, including subcortical volumes and ICV, are highly heritable[11, 12], the above findings may be explained by shared genetic architecture. Most genetic correlation analyses of schizophrenia and brain volumetric phenotypes have found non-significant correlations[13], except a marginal negative genetic correlation between schizophrenia and hippocampal volume[14]. Another study applying rank-rank hypergeometric overlap test did not find genetic overlap

between schizophrenia and brain volumetric phenotypes [13]. However, those methods are not able to capture genetic overlap with mixed effect directions[15], that is, when two phenotypes share a mixture of genetic variants with agonistic and antagonistic effects, warranting new statistical approaches. The conjunctional false discovery rate (conjFDR) analysis was developed to improve discovery of shared individual genetic variants by leveraging the combined power of two genome-wide association studies (GWASs)[16-18]. A previous study applying conjFDR analysis has reported 6 genetic loci shared between schizophrenia and hippocampal volume, putamen, and ICV[19]. Moreover, multivariate analytic methods have identified that chr6p22.1 was associated with schizophrenia-discriminating regional gray matter volume reduction[20]. Nonetheless, previous studies, based on moderate sample sizes, only examined polygenic overlap at the individual genetic variant level and characterized few shared common loci. Combining advanced new statistical models and recent large-cohort GWASs could investigate the polygenic overlap at the level of global genetic architecture and increase the power of discovering shared genomic loci, which may provide further evidence of overlapping genetic architecture between schizophrenia and brain volumetric phenotypes.

Here, we applied MiXeR[15] and conjFDR analysis[16-18] to the latest GWASs[11, 21] to assess the shared genetic basis between schizophrenia and brain volumetric phenotypes, including seven bilateral subcortical volumes (accumbens, amygdala, caudate, hippocampus, pallidum, putamen, and thalamus) and ICV. For each schizophrenia-volume pair, MiXeR[15] estimated the shared polygenic architecture between two traits irrespective of the direction of individual effects. The individual shared variants were identified using conjFDR analysis[16-18], and we further examined the effect directions of shared associations in independent samples. To gain an understanding of underlying molecular mechanisms, we applied the genes mapped by shared loci to functional enrichment analysis, co-expression enrichment analysis, and gene-based association analysis between brain expression and brain volumetric phenotypes in childhood. Together, our findings build a framework for understanding the

shared genetic basis and underlying molecular mechanisms between schizophrenia and brain morphology.

## Methods

*Genome-Wide Association Studies Datasets*

For MiXeR and conjFDR analysis, we obtained previously generated GWAS summary statistics of schizophrenia from Psychiatric Genomic Consortium (PGC) and of brain volumetric phenotypes from UK Biobank. The schizophrenia GWAS included 53,386 patients with schizophrenia and 77,258 healthy controls of European ancestry[11]. We acquired GWASs of eight brain volumetric phenotypes from UK Biobank (33,735 healthy participants, accession 27412)[21]. Details of data processing are described in Supplementary Methods and previous publications[11, 21, 22].

For validation, we included GWASs from a previously published meta-analysis of schizophrenia based on East Asian cohorts (22,778 cases and 35,362 controls)[23] and of brain volumetric phenotypes based on European participants (N=13,171)[24]. We further employed a height GWAS with 693,529 European participants[25] as a non-brain-related comparator. All GWASs were approved by the relevant ethics committees, and informed consent was obtained from all participants. The Regional Committee for Medical and Health Research Ethics of South-East Norway approved the inclusion of data from UK biobank. More details are provided in the corresponding publications[11, 21, 23-25].

*Adolescent Brain Cognitive Development Study Cohort*

The Adolescent Brain Cognitive Development (ABCD, data release 3.0) cohort is a deeply phenotyped imaging genetics study from the United States (N=11,878, ages: 9–10 year-old). Genetic data and neuroimaging data at baseline have been collected with each participant's written informed

consent and with approval from the respective local institutional review boards. After pre-processing, we used 3,757 European samples for gene-based association analysis. This sample included 1,771 females and the mean age was 9.942 years (standard deviation=0.619). Details of the cohort, data collection and data preprocessing are presented in Supplementary Methods and corresponding publications[26-28].

*Statistical Analysis of Genetic Overlap*

To examine the genetic overlap for each schizophrenia-volume pair, we first applied MiXeR[15] to estimate the overall shared polygenic architecture irrespective of the effect directions and coefficients. MiXeR goes around the intrinsically difficult problem of identifying the genomic locations of trait-influencing variants by estimating the number of shared and trait-specific trait-influencing variants. The 'trait-influencing variants' represent common variants with a non-zero additive effect on a trait. Conceptually, a variant with a non-zero additive effect indicates that it is likely to have an impact on the trait. However, please note that this definition is based on a statistical model and proof of true causality requires further in vitro and in vivo studies. To look into the shared genetic architecture and relevant mechanisms, we performed conjFDR[16-18] analysis to identify individual shared genomic loci. Descriptions of both tools are shown below, with more details provided in Supplementary Methods, and the original publications[15-18].

We applied MiXeR v1.3[15] to interpret the genetic characteristics of each trait (univariate analysis) and to evaluate the extent of polygenic overlap in each schizophrenia-volume pair (bivariate analysis). In univariate analysis[29], MiXeR modeled the additive genetic effect of each SNP on a trait and estimated heritability (a sum of effect across all trait-influencing variants), polygenicity (number of trait-influencing variants estimated to explain 90% heritability), and discoverability (the average magnitude of additive genetic effects among trait-influencing variants). Next, bivariate analysis[15]

modeled the estimated additive genetic effects on one schizophrenia-volume pair as a mixture of four components, representing null SNPs in both traits, SNPs with non-zero effect only on one trait, SNPs with non-zero effect only on the other trait, and SNP with non-zero effect on both traits. The number of estimated shared and trait-specific trait-influencing variants were presented in a Venn diagram. Meanwhile, MiXeR estimated the proportion of shared trait-influencing variants with concordant effects (i.e., the same effect directions) on both traits by modeling the effect direction of the shared component. Akaike information criterion (AIC) was employed for model fitting, with a positive AIC value indicating sufficient power of input GWASs.

We applied conjFDR analysis[16-18] to explore the individual shared genetic variants. The conjFDR analysis is built on an empirical Bayesian statistical framework and it leverages cross-trait SNP enrichment to improve the power to discover shared variants. For each schizophrenia-volume pair, we first used SNP associations of the brain volumetric phenotype to re-rank test statistics and re-calculate the significance of associations between these SNPs and schizophrenia. Then, we inverted the roles of two phenotypes to re-calculate the significance of SNP associations to the brain volumetric phenotypes. The conjFDR analysis conservatively estimates the posterior probability that a SNP has no association with either trait, represented as a conjFDR value. A SNP with a conjFDR value $< 0.05$ was considered as a shared SNP[16, 18, 22].

*Distinct Genomic Locus Definition, Functional Annotation, and Gene Mapping*

For each schizophrenia-volume pair, we first applied FUMA protocol[30] to define its shared loci (Supplementary Methods). We defined genomic borders (i.e., start basepair to end basepair) for each locus based candidate SNPs for the lead SNP, which is the SNP with the most significant conjFDR value. Next, we compared shared loci across schizophrenia-volume pairs using Bedtools[31], and defined distinct shared loci across pairs. If genetic borders of two or more loci from different

schizophrenia-volume pairs overlapped, we merged them into one distinct shared locus with the union of their genomic borders.

We annotated variant function, the predicted deleteriousness, and regulatory effect for candidate variants using FUMA[30] and significant brain *cis*-expression quantitative trait loci (eQTLs) for lead SNPs using GTEx v8[32] (Supplementary Methods). Fisher exact analysis examined the enrichment of shared variants in functional regions.

We mapped the shared loci to genes using positional mapping, and brain tissue-based eQTL mapping and chromatin interaction mapping (Supplementary Methods). A gene that meets at least one of the following criteria was defined as a shared gene: (1) mapped by at least two strategies either by the same variant or by different variants; (2) mapped by a putatively deleterious exonic variant using a Combined Annotation Dependent Depletion threshold of 12.37[33].

*SNP Sign Test*

In line with previous studies[13, 34], we performed SNP sign test (Supplementary Methods) to examine whether the shared genomic loci replicated *en masse* in independent samples. Using lead SNPs, SNP sign test verified directionality of allelic association between the discovery and independent cohorts, with the null hypothesis of randomly oriented associations. As the independent GWAS for schizophrenia and for brain volumetric phenotypes were based on different ancestries, we applied the SNP sign test for each trait separately. FDR correction was performed across all SNP sign tests.

*Functional Enrichment Analysis*

Functional enrichment analysis of four categories, including Gene Ontology, KEGG pathways, cell type signatures, and immunological signature, was performed for shared genes using Molecular

Signatures Database (v7.4)[35].The significance of enrichment was calculated from the hypergeometric distribution and FDR correction was performed cross categories. For all functional enrichment analysis within the present study, genes located within complex LD regions were excluded.

*Lifespan Spatiotemporal Gene Expression Trajectory*

We applied a hypergeometric test[36] to calculate the enrichment of shared genes on 73 previously reported co-expression modules from PsychENCODE human brain development study (http://development.psychencode.org/)[37]. We applied the hypergeometric test for each schizophrenia-volume pair separately, and performed FDR correction across all tests. We named the co-expressed shared genes, which were enriched in one co-expression module, as a co-expression cluster.

The spatiotemporal brain expression trajectory was evaluated using the processed human mRNA-seq data from the PsychENCODE study[37]. For each gene, we estimated its brain-tissue expression (i.e., expression in each anatomic tissue) and its whole-brain expression level (i.e., the average expression of this gene across all available brain tissues). For a co-expression cluster, its expression profile was calculated by averaging the expression levels of clustered genes. Gene expression levels were represented using reads per kilobase of transcript per million reads mapped (RPKM). Details are provided in Supplementary Methods and the original publication[37].

*Gene-Based Expression-Trait Associations Analysis*

For a schizophrenia-volume pair, we investigated whether the expression of its shared genes is associated the corresponding volumetric phenotype in late childhood. To this end, we applied PrediXcan[38] analysis using 3,757 European participants in late childhood from ABCD cohort. For a schizophrenia-volume pair (e.g., schizophrenia-thalamus pair), PrediXcan first imputed the expression of its shared genes in the corresponding brain tissue (e.g., thalamic expression) based on individual

genotype and pre-computed eQTL prediction models. Next, PrediXcan examined the association between the imputed expressions and the corresponding phenotype (e.g., thalamic volume) in late childhood, controlling age, sex, MRI scanners, ICV (for non-ICV phenotype) and the top 20 principal components. The eQTL prediction models are elastic net prediction models trained with GTEx v8 database[32]. Among those eight brain volumetric phenotypes, PredictDB provides eQTL prediction models for amygdala, caudate, hippocampus, hypothalamus, accumbens, and putamen. As PrediXcan imputed gene expression using eQTL prediction models, shared genes without any significant brain eQTL in GTEx v8 were excluded. FDR correction was applied across all expression-trait associations. An expression-trait association with an FDR<0.05 was considered significant and the gene was defined as a candidate gene for the corresponding volumetric phenotype.

For a candidate gene, we examined whether its brain gene expression was associated with schizophrenia using two methods (Supplementary Method). The first method is S-PrediXcan[39], which is a summary-version of PrediXcan and also aims to compute expression-trait associations. Although PrediXcan uses individual level genotype data while S-PrediXcan uses GWAS summary statistics[39], PrediXcan and S-PrediXcan have a high concordance ($r^2$>0.99)[39]. S-PrediXcan evaluated imputed expression-trait associations based on the schizophrenia GWAS[11] and brain-tissue eQTL prediction models. Meanwhile, we assessed whether candidate genes are significantly differently expressed between schizophrenia cases and controls using the differential expression profiles in DLPFC and hippocampus from Brainseq consortium (http://eqtl.brainseq.org/)[40]. Those differential expression profiles are based on CommonMind Consortium (CMC) samples[41] and BrainSeq Phase 2 samples[42].

*Code availability*

The present study applied previously published approaches, of which codes are shared on public repositories: MiXeR v1.3 (https://github.com/precimed/mixer), conjFDR analysis

(https://github.com/precimed/pleiofdr), PrediXcan (https://github.com/hakyimlab/PrediXcan), and S-PrediXcan (https://github.com/hakyimlab/MetaXcan).

## Results

*Schizophrenia and Brain Volumetric Phenotypes Share Variants with Mixed Effect Directions.*

MiXeR univariate analysis estimated a higher SNP heritability (0.38), a higher polygenicity (9,574 trait-influencing variants), and a lower discoverability (6.13E-05) for schizophrenia than all brain volumetric phenotypes (SNP heritability ranges from 0.19 to 0.30, polygenicity ranges from 994 to 2,861 trait-influencing variants, and discoverability ranges from 1.39E-04 to 3.47E-04, **Fig. 1a**). MiXeR bivariate analysis showed that a substantial proportion of SNPs associated with a brain volumetric phenotype was also related to the risk of schizophrenia (**Fig. 1b**). The thalamus shared the largest proportion of trait-influencing variants with schizophrenia (96%, 1,888 out of 1,961 variants) and accumbens shared the smallest proportion but still 74% (1,814 out of 2,455 variants). The proportion of shared variants with concordant effects on schizophrenia and brain volumetric phenotypes ranges from 42% for schizophrenia and pallidum to 54% for schizophrenia and ICV (average proportion: 48%, **Fig. 1c**). Corroborating previous reports[13], LD score regression (LDSC)[43] demonstrated non-significant genetic correlations between schizophrenia and brain volumetric phenotypes, except for pallidum ($r_g$= -0.10, FDR-corrected P, FDR=0.014, **Fig. 1c**). We also included height as a non-brain comparator[25]. In line with prior studies[22], height showed a smaller proportion of polygenic overlap with schizophrenia (28%, 1,117 out of 4,026 variants) and a non-significant genetic correlation ($r_g$= -0.02, FDR = 0.228). Details of MiXeR and LDSC results are provided in Supplementary Table 1-3, and Supplementary Fig. 1.

*Shared Genetic Loci are enriched in Neurodevelopmental Processes*

To explore the MiXeR estimated shared basis at individual variant level, we applied conjFDR analysis for each schizophrenia-volume pair, separately. As shown in **Fig. 2a**, conjFDR analysis showed that schizophrenia shares a group of loci with the volume of accumbens (N=10), amygdala (N=8), caudate (N=15), hippocampus (N=13), pallidum (N=12), putamen (N=21), and thalamus (N=25), as well as ICV (N=36, Supplementary Table 4). Among the shared associations, a total of 21 genomic loci were jointly associated with schizophrenia and at least two brain volumetric phenotypes (Supplementary Table 5), resulting in 107 distinct shared loci in total.

We next examined the validity of shared loci using the SNP sign test[34]. For schizophrenia, 95 out of 107 lead SNPs were available in independent samples and 73% of associated alleles showed sign consistency (FDR=5.32E-05, **Fig. 2b**, and Supplementary Tables 4 and 5). SNP associations of five volumetric phenotypes, including volume of caudate (FDR=0.003), hippocampus (FDR=0.049), putamen (FDR=0.001), and thalamus (FDR=0.001), and ICV (FDR=7.29E-04), showed 82% to 92% significant sign consistency in the independent cohort. We also evaluated previously reported shared genetic loci in our findings. The present study replicated 67% (4 out of 6) shared loci reported in a previous conjFDR analysis[19], and validated previously implicated associations of rs13107325 and chr6p22.1 with schizophrenia and brain volumetric phenotypes (Supplementary Table 6).

To investigate underlying biological processes, we mapped the 107 distinct shared loci to genes, resulting in 795 shared genes in total (Supplementary Table 7). Details for further gene-based analysis are shown in **Fig. 3**. The 459 shared genes outside of LD regions showed enrichment in the axon guidance pathway and neurodevelopmental biological processes, such as neurogenesis, neuron differentiation, neuron development, negative regulation of axonogenesis, and central nervous system development (Supplementary Table 8), indicating brain developmental processes underlying this

shared genetic architecture. We also found significant enrichment in human brain cells, including oligodendrocytes in embryonic cerebral cortex, medial neuroblasts, and GABAergic neurons (Supplementary Table 8). As our previous study suggested that genetic overlap between schizophrenia and cortical structure are enriched in immunologic processes[22], we evaluated gene-set enrichment regarding immunologic signatures and identified 27 enriched cell states and perturbations within the immune system (Supplementary Table 8).

*Distinct Spatiotemporal Brain Expression Trajectories throughout Lifespan.*

Among the lead SNPs of those shared loci, the majority were located within non-coding regions, with significant enrichment in intronic (FDR=0.034, OR=1.55, 95% confidence interval, CI: 1.08, 2.21, Supplementary Table 9) and upstream regions (FDR=0.034, OR=3.66, 95% CI: 1.17, 8.78), suggesting the shared genetic variants may be involved in gene regulation[44]. We next examined whether those shared lead SNPs have genetic regulatory effects on gene expression using the significant *cis*-eQTLs. We found 389 brain *cis*-eQTLs for shared lead SNPs (**Fig. 2c**, Supplementary Table 10), suggesting that those shared variants are associated with gene expression variations in brain tissues.

As coordinated expression of genes within brain tissues plays a vital role in brain development and neuropsychiatric disorders[37, 45], the transcriptional landscape of the shared genes may provide insights into the neuropathology of schizophrenia. Therefore, we first investigated the co-expression patterns among shared genes (**Fig. 3**). Five schizophrenia-volume pairs (accumbens, hippocampus, pallidum, putamen, and ICV) showed that shared genes were significantly enriched in five previously reported co-expression modules (Supplementary Table 11)[37]. Those five co-expression modules were labeled as modules 2, 3, 5, 7, and 9 in the original publication[37], and contained 47, 23, 11, 18, and 12 shared genes, respectively. As the enriched shared genes were only a subset of co-expressed genes within a

module, we specified the shared genes enriched in modules 2, 3, 5, 7, and 9 as clusters red, pink, azure, blue, and black, respectively. Additionally, amygdala, caudate, and thalamus showed nominally significant co-expression module enrichment (Supplementary Table 11). Next, we assessed the whole-brain and tissue-specific lifespan expression patterns of those five co-expression clusters. Both whole-brain (**Fig. 2d**) and tissue-specific (Supplementary Fig. 2) expression showed two main developmental expression features —cluster red, azure, and black showed peak expression in fetal period while cluster pink and blue exhibited peak expression in postnatal period. As genes with similar expression pattern are likely to participate in the same functional processes[46], we explored functional enrichment for each cluster. Those five clusters pointed to different biological mechanisms (Supplementary Table 12). Cluster red showed significant enrichment in calcium ion binding and cell-cell adhesion-related gene sets. Cluster azure demonstrated shared genes are enriched in two types of brain cells, including neural stem cells in embryonic cortex and oligodendrocyte progenitor cells in the prefrontal cortex. Cluster black revealed shared genes are overrepresented in nuclear-related and spliceosomal complex gene sets.

*Shared Genes Highlight Genetic Risk of Schizophrenia in Brain Development in Childhood*

In light of the above findings and the neurodevelopmental theory of schizophrenia[3], we examined whether the brain expression of shared genes is relevant to brain volumetric phenotype before the peak age of onset for schizophrenia. To this end, we used PrediXcan to evaluate associations between gene expression and brain volumetric phenotypes in childhood using 3,757 ABCD participants (**Fig. 3**). In total, PrediXcan examined 195 imputed expression-trait associations for volumetric phenotypes, corresponding to 161 shared genes. The analysis highlighted two candidate genes for the development of thalamus —*CRHR1* (ENSG00000120088, **Fig. 4a**) and *ARL17A* (ENSG00000185829) — by showing the imputed hypothalamic expression levels of those two genes are associated with thalamic volume in late childhood (*CRHR1*: FDR=0.024, Z= 3.67, and *ARL17A*: FDR=0.024, Z= -3.68,

Supplementary Table 13). Both genes were mapped to the same shared locus on chromosome 17 (chr17:43859929-44865603, lead SNP: rs70602, **Fig. 4b**). The brain *cis*-eQTL pairs showed that rs70602 was associated with the expression of *CRHR1* and *ARL17A* in brain tissues (**Fig. 4c**).

We did not find any available dataset to evaluate the association between hypothalamic expression of those two genes and schizophrenia. However, it has been suggested that the expression of *CRHR1* in thalamus is not significantly different from most brain tissues (Supplementary Figure 3)[37]. The expression trajectories of *CRHR1* and *ARL17A* in nine brain tissues were further shown in **Fig. 4d**. Thus, we investigated the association between schizophrenia and expression of *CRHR1* and *ARL17A* in other brain tissues. S-PrediXcan showed a significant association between schizophrenia and *ARL17A* in cerebellum: FDR=2.81E-04, Z=3.63) and in putamen (FDR=2.81E-04, Z=3.63). Meanwhile, CMC samples[47] showed that schizophrenia cases have significantly reduced expression of *CRHR1* (P=0.032, $\log_2$ fold change, $\log_2$FC=-0.04) and significantly increased expression of *ARL17A* (P=0.033, $\log_2$FC=0.06) in DLPFC than healthy controls. This up-regulated DLPFC expression of *ARL17A* in schizophrenia cases were also found in the BrainSeq Phase 2 samples with nominal significance (P=0.089, $\log_2$FC=0.11, **Fig. 4a**)[42]. We did not observe any schizophrenia-associated expression in hippocampus. More details are provided in Supplementary Methods and Supplementary Table 14.

## Discussion

This study provided evidence of shared genetic architecture between schizophrenia and brain volumetric phenotypes and indicated 107 shared genetic loci. The SNP sign test for those shared genetic associations suggested sign consistency between the original GWAS and independent datasets. The shared genes showed two distinct expression trajectories, with one peaking prenatally and the other peaking postnatally. We next remarked two candidate genes (*ARL17A* and *CRHR1*) for the

development of the thalamus by showing the imputed hypothalamic expression levels of those genes are associated with thalamic volume in late childhood. Along with the enrichment of shared genes in neurodevelopment, these findings suggest that the shared genetic component between schizophrenia and brain volumetric phenotypes may be involved in brain development in as early as childhood, adding support to the neurodevelopmental hypothesis[3].

It has been well-established that patients with schizophrenia have altered brain morphology when compared to healthy controls on the group level, but the genetic findings have been inconsistent. By applying tools that take mixed effects into account, the present study suggested shared genetic architecture with mixed effect directions underlying the phenotypic associations[22]. Here, we provided evidence of an overlapping genetic basis from two aspects. First, MiXeR revealed that all brain volumetric phenotypes share a large proportion of variants (74% to 96%) with schizophrenia, which was further supported by the positive AIC values in bivariate analysis. The shared genetic associations were found by conjFDR at individual variant and locus levels, with consistent genetic effects in independent samples and previous studies[19, 20, 48]. Secondly, given the non-significant genetic correlation might be due to mixed directions of allelic effects among shared variants[18, 22], we examined the effect directions of those shared variants, revealing approximately 50% shared variants with concordant effects on schizophrenia and the brain volumetric phenotype while the rest have discordant effects. ConjFDR analysis also showed mixed effect directions with both concordant and discordant effects found in lead shared SNPs. This indicates a scenario of genetic overlap without correlation[18] for schizophrenia and brain volumetric phenotypes. The above findings are also relevant to the "polygenic-pleiotropy" model of the brain[18, 22], which suggests that brain-related polygenic phenotypes may share a genetic basis along with their own specific genetic characteristics.

Brain morphological measures have been hypothesized to be less polygenic than schizophrenia, forming the basis for the endophenotype approach[49] in schizophrenia. Our findings further support

this hypothesis by revealing subcortical volumes and ICV have a lower polygenicity and a higher discoverability than schizophrenia, consistent with previous reports[49]. The positive AIC values indicate sufficient power for univariate analysis, suggesting the difference in polygenicity and discoverability is not mainly driven by different GWAS sample sizes. A possible explanation is that schizophrenia is a complex and heterogeneous disorder with a broad range of symptoms and functional deficits[50], so its high polygenicity is likely the result of a sum of SNPs associated with those clinical and course features. In contrast, brain volumetric phenotypes are less complex and have been proposed as "closer to the underlying biology," which may lead to a smaller polygenicity[49]. Moreover, polygenicity has been hypothesized to be related to etiological heterogeneity, i.e., how many etiological mechanisms are underlying the development of the phenotype[49, 50]. A further investigation on the underlying etiological mechanisms is therefore suggested.

The enrichment of Gene Ontology, pathways, and cell types suggested the shared genes between schizophrenia and brain volumetric phenotypes are involved in neural developmental processes, consistent with previous reports[51]. This is also supported by the lifespan expression trajectories that revealed two expression patterns for shared genes, with one group of genes showing high expression levels in the prenatal period. Of particular interest are the co-expression clusters red and azure, corresponding to shared genes enriched in modules 2 and 5[37]. Both clusters revealed peak expression in prenatal period, consistent with the original publication[37] that modules 2 and 5 have a significant higher expression in prenatal period than postnatal period. As modules 2 and 5 has shown a significant coincident occurrence of genes whose expression was enriched in neurons[37], we also examined the functional enrichment for each co-expression cluster. Although those five co-expression clusters only revealed few enriched gene sets, which may be due to the limited number of shared genes, we still found that cluster azure, corresponding to module 5, was enriched in stem and

progenitor cells in the brain, both important in cell proliferation and differentiation in early-stage brain development[52].

Additionally, Luo et al[48] have suggested an interplay between the development of schizophrenia and the brain morphology in adolescence. We therefore explored whether those shared genes are associated with the brain development in healthy participants before the peak age of schizophrenia onset. We found that the genetically imputed hypothalamic expression of *CRHR1* and *ARL17A* is associated with the thalamic volume in childhood, and the expression of those two genes in putamen, cerebellum, and DLPFC were implicated in schizophrenia. Specifically, *CRHR1* encodes the main receptor of corticotrophin-releasing hormone, and it is well known for its role in the hypothalamic-pituitary-adrenal axis-mediated response to stress, a well-established modulator of brain morphology[53] and schizophrenia[54]. *ARL17A* belongs to the ADP-ribosylation factor family which regulates membrane trafficking and vesicular transport[55], relevant for neuronal development and communication[56]. The expression-trait associations for *CRHR1* and *ARL17A* in childhood structural brain development and schizophrenia provides insights into recent observations that early brain volumetric alterations, which are associated with the later emergence of schizophrenia, predate the onset of clinical syndromes[57]. Combining the above functional enrichment and gene expression findings, the present study adds support to the neurodevelopmental theory of schizophrenia by suggesting that shared genetic underpinning between schizophrenia and brain volumetric phenotypes might be involved in early neurodevelopmental processes and those shared genes may modulate the brain morphology years before the peak age of schizophrenia onset.

It is important to note that our results are limited to common variants on autosomal chromosomes. Rare variants, such as copy number variations, individually shows high penetrance for schizophrenia[58] and schizophrenia-associated copy number variations have shown associations with reduced subcortical volumes[59]. The genetic factors on sex chromosomes have also been implicated in both

schizophrenia[11] and brain development[60]. Hence, future studies on rare variants and sex chromosomes are needed. The second limitation is that the majority of our analysis was based on the European population. To address this, we included the East Asian population-based schizophrenia GWAS and found a significant trans-ancestry sign consistency for the associations between shared lead SNPs and schizophrenia. Thirdly, the current study indicated the need to maximize sample sizes in GWAS. For example, in the MiXeR bivariate analysis for schizophrenia and thalamus, AIC values suggest extensive polygenic overlap and the requirement to apply larger GWASs for a precise estimate of polygenic overlap. The power of GWASs also affects the conjFDR analysis and the follow-up analysis, such as SNP sign test and co-expression enrichment analysis. SNP sign test for accumbens, amygdala, and pallidum returned 67~75% sign consistency in independent samples but non-significant. This might be due to a lack of power as the number of input SNPs was relatively small. Also, we only observed significant co-expression pattern for accumbens, hippocampus, pallidum, putamen, and ICV but not for amygdala, caudate, and thalamus. This may suggest differences in the genetic architecture or neurobiological mechanisms across brain volumetric phenotypes. However, the co-expression enrichment is limited by the SNP discovery in conjFDR analysis and the selection of shared genes. Further work with larger GWASs and advanced statistical approaches may help valid those findings. Finally, although we found two shared genes of which predicted thalamic expression is associated with thalamic volume in childhood, the current findings cannot distinguish causality (where the association between rs70602 and thalamic volume is mediated by hypothalamic expression of *CRHR1* and *ARL17A*) or pleiotropy (where rs70602 is associated with both thalamic volume and hypothalamic expressions separately). Another note of caution is due here since both genes were mapped at the same locus and their thalamic expression was predicted via eQTL. However, whether both of these genes are indeed involved in SCZ pathogenesis remains to be determined. The relatively low postnatal brain expression of *ARL17A* may also limit the expression-trait association analysis in childhood. Further research, such as in vitro and in vivo experiments, should be undertaken to

investigate the causality and pleiotropy assumptions, and to retrace the impact of each gene on brain development trajectory from prenatal stage to childhood.

In summary, we quantified the overlapping genetic architecture and suggested 107 shared genetic loci between schizophrenia and subcortical volumes and ICV. The genes mapped by those shared loci implicated neurodevelopmental processes prior to schizophrenia development and highlighted *CRHR1* and *ARL17A* of which thalamic expression was predicted as associated with the development of thalamus in childhood. These findings provide genetic insights into the brain volumetric alternations in schizophrenia patients, and a further study examining the role of the altered neurodevelopment in schizophrenia is therefore suggested.

## Acknowledgements

The authors thank the Psychiatric Genetics Consortium (PGC, https://www.med.unc.edu/pgc/), the UK biobank (https://www.ukbiobank.ac.uk/), the Enhancing NeuroImaging Genetics through Meta-Analysis (ENIGMA, http://enigma.ini.usc.edu/) consortium for access to GWAS data, as well as Adolescent Brain Cognitive Development (ABCD) Study (https://abcdstudy.org) for access to genotype and brain imaging data. We thank all participants and employees of PGC, UK biobank (Application ID 27412), and ENIGMA, ABCD for their contribution to this study. The ABCD Study is supported by the National Institutes of Health and additional federal partners under award numbers: U01DA041022, U01DA041028, U01DA041048, U01DA041089, U01DA041106, U01DA041117, U01DA041120, U01DA041134, U01DA041148, U01DA041156, U01DA041174, U24DA041123, and U24DA041147. A full list of supporters is available at https://abcdstudy.org/federal-partners/.

The authors were funded by the National Institutes of Health (grants NS057198 and EB00790), the Research Council of Norway (grants 229129, 276082, 213837, 223273, and 248980), the South-East

Regional Health Authority (grant 2017-112), Stiftelsen Kristian Gerhard Jebsen (grants SKGJ-MED-008 and SKGJ-MED-021), the Psychiatric Genomics Consortium US Norway Collaboration (RCN 248980), part of convergence environment (MultiModal Mental Models [4MENT]) funded by the University of Oslo Life Science and Scientia Fellows, European Union's Horizon 2020 Research and Innovation Active Grant (847776 CoMorMent), and European Union's Horizon 2020 research and innovation programme (801133 Marie Skłodowska-Curie grant agreement), and have received internationalization support from UiO:Life Science. This work was performed on resources provided by Sigma2 (the National Infrastructure for High Performance Computing and Data Storage in Norway) and the TSD (Tjeneste for Sensitive Data) facilities.

## Conflict of Interest

Dr. Andreassen reported grants from Stiftelsen Kristian Gerhard Jebsen, South-East Regional Health Authority, Research Council of Norway, and European Union's Horizon 2020 during the conduct of the study; personal fees from HealthLytix (stock options), Lundbeck (speaker's honorarium), and Sunovion (speaker's honorarium) outside the submitted work; and had a pending patent for systems and methods for identifying polymorphisms. Dr. Dale reported grants from the National Institutes of Health outside the submitted work; had a patent for US7324842 licensed to Siemens Healthineers; is a founder of and holds equity in Cortechs Labs and serves on its scientific advisory board; is a member of the scientific advisory board of Human Longevity; a member of the scientific advisory board of Healthlytix; and receives funding through a research agreement with GE Healthcare. No other disclosures were reported.

## Author contributions

W.Q. and O.A. designed the study; W.Q., and D.v.d.M. pre-processed the data. W.Q. performed all analyses, with conceptual input from O.A., D.v.d.M., Y.W., N.P., G.H., and O.S.. All authors helped shape the research, and contributed to interpretation of results. W.Q. drafted the manuscript with input from O.A., D.v.d.M., N.P., G.H., and O.S.. All authors contributed to and approved the final manuscript.

## References


1. Anderson KK. Towards a public health approach to psychotic disorders. *The Lancet Public Health* 2019; **4**(5)**:** e212-e213.

2. Hjorthøj C, Stürup AE, McGrath JJ, Nordentoft M. Years of potential life lost and life expectancy in schizophrenia: a systematic review and meta-analysis. *The Lancet Psychiatry* 2017; **4**(4)**:** 295-301.

3. Weinberger DR. Implications of normal brain development for the pathogenesis of schizophrenia. *Archives of general psychiatry* 1987; **44**(7)**:** 660-669.

4. Carlsson M, Carlsson A. Schizophrenia: a subcortical neurotransmitter imbalance syndrome? *Schizophr Bull* 1990; **16**(3)**:** 425-432.

5. Laviolette SR. Dopamine Modulation of Emotional Processing in Cortical and Subcortical Neural Circuits: Evidence for a Final Common Pathway in Schizophrenia? *Schizophrenia Bulletin* 2007; **33**(4)**:** 971-981.

6. Ghahremani DG, Lee B, Robertson CL, Tabibnia G, Morgan AT, De Shetler N *et al.* Striatal dopamine $D_2/D_3$ receptors mediate response inhibition and related activity in frontostriatal neural circuitry in humans. *J Neurosci* 2012; **32**(21)**:** 7316-7324.

7. van Erp TGM, Hibar DP, Rasmussen JM, Glahn DC, Pearlson GD, Andreassen OA *et al.* Subcortical brain volume abnormalities in 2028 individuals with schizophrenia and 2540 healthy controls via the ENIGMA consortium. *Molecular psychiatry* 2016; **21**(4)**:** 547-553.

8. Boos HB, Aleman A, Cahn W, Hulshoff Pol H, Kahn RS. Brain volumes in relatives of patients with schizophrenia: a meta-analysis. *Archives of general psychiatry* 2007; **64**(3)**:** 297-304.

9. McIntosh AM, Owens DC, Moorhead WJ, Whalley HC, Stanfield AC, Hall J *et al.* Longitudinal volume reductions in people at high genetic risk of schizophrenia as they develop psychosis. *Biological psychiatry* 2011; **69**(10)**:** 953-958.

10. Alnæs D, Kaufmann T, van der Meer D, Córdova-Palomera A, Rokicki J, Moberget T *et al.* Brain Heterogeneity in Schizophrenia and Its Association With Polygenic Risk. *JAMA psychiatry* 2019; **76**(7)**:** 739-748.

11. Trubetskoy V, Pardiñas AF, Qi T, Panagiotaropoulou G, Awasthi S, Bigdeli TB *et al.* Mapping genomic loci implicates genes and synaptic biology in schizophrenia. *Nature* 2022; **604**(7906)**:** 502-508.

12. Zhao B, Luo T, Li T, Li Y, Zhang J, Shan Y *et al.* Genome-wide association analysis of 19,629 individuals identifies variants influencing regional brain volumes and refines their genetic co-architecture with cognitive and mental health traits. *Nature Genetics* 2019; **51**(11)**:** 1637-1644.

13. Franke B, Stein JL, Ripke S, Anttila V, Hibar DP, van Hulzen KJE *et al.* Genetic influences on schizophrenia and subcortical brain volumes: large-scale proof of concept. *Nature Neuroscience* 2016; **19**(3)**:** 420-431.

14. Ohi K, Shimada T, Kataoka Y, Yasuyama T, Kawasaki Y, Shioiri T *et al.* Genetic correlations between subcortical brain volumes and psychiatric disorders. *The British journal of psychiatry : the journal of mental science* 2020; **216**(5)**:** 280-283.

15. Frei O, Holland D, Smeland OB, Shadrin AA, Fan CC, Maeland S *et al.* Bivariate causal mixture model quantifies polygenic overlap between complex traits beyond genetic correlation. *Nature Communications* 2019; **10**(1)**:** 2417.

16. Andreassen OA, Djurovic S, Thompson WK, Schork AJ, Kendler KS, O'Donovan MC *et al.* Improved detection of common variants associated with schizophrenia by leveraging pleiotropy with cardiovascular-disease risk factors. *American journal of human genetics* 2013; **92**(2)**:** 197-209.

17. Andreassen OA, Thompson WK, Dale AM. Boosting the power of schizophrenia genetics by leveraging new statistical tools. *Schizophr Bull* 2014; **40**(1)**:** 13-17.

18. Smeland OB, Frei O, Dale AM, Andreassen OA. The polygenic architecture of schizophrenia — rethinking pathogenesis and nosology. *Nature Reviews Neurology* 2020; **16**(7)**:** 366-379.

19. Smeland OB, Wang Y, Frei O, Li W, Hibar DP, Franke B *et al.* Genetic Overlap Between Schizophrenia and Volumes of Hippocampus, Putamen, and Intracranial Volume Indicates Shared Molecular Genetic Mechanisms. *Schizophr Bull* 2018; **44**(4)**:** 854-864.

20. Chen J, Calhoun VD, Lin D, Perrone-Bizzozero NI, Bustillo JR, Pearlson GD *et al.* Shared Genetic Risk of Schizophrenia and Gray Matter Reduction in 6p22.1. *Schizophr Bull* 2019; **45**(1)**:** 222-232.

21. van der Meer D, Shadrin AA, O'Connell K, Bettella F, Djurovic S, Wolfers T *et al.* Boosting Schizophrenia Genetics by Utilizing Genetic Overlap With Brain Morphology. *Biological psychiatry* 2022.

22. Cheng W, Frei O, van der Meer D, Wang Y, O'Connell KS, Chu Y *et al.* Genetic Association Between Schizophrenia and Cortical Brain Surface Area and Thickness. *JAMA psychiatry* 2021; **78**(9)**:** 1020-1030.

23. Lam M, Chen C-Y, Li Z, Martin AR, Bryois J, Ma X *et al.* Comparative genetic architectures of schizophrenia in East Asian and European populations. *Nature Genetics* 2019; **51**(12)**:** 1670-1678.

24. Hibar DP, Stein JL, Renteria ME, Arias-Vasquez A, Desrivières S, Jahanshad N *et al.* Common genetic variants influence human subcortical brain structures. *Nature* 2015; **520**(7546)**:** 224-229.

25. Yengo L, Sidorenko J, Kemper KE, Zheng Z, Wood AR, Weedon MN *et al.* Meta-analysis of genome-wide association studies for height and body mass index in ~700000 individuals of European ancestry. *Human molecular genetics* 2018; **27**(20)**:** 3641-3649.

26. Casey BJ, Cannonier T, Conley MI, Cohen AO, Barch DM, Heitzeg MM *et al.* The Adolescent Brain Cognitive Development (ABCD) study: Imaging acquisition across 21 sites. *Developmental cognitive neuroscience* 2018; **32:** 43-54.

27. Kowalski MH, Qian H, Hou Z, Rosen JD, Tapia AL, Shan Y *et al.* Use of >100,000 NHLBI Trans-Omics for Precision Medicine (TOPMed) Consortium whole genome sequences improves imputation quality and detection of rare variant associations in admixed African and Hispanic/Latino populations. *PLOS Genetics* 2019; **15**(12)**:** e1008500.

28. Volkow ND, Koob GF, Croyle RT, Bianchi DW, Gordon JA, Koroshetz WJ *et al.* The conception of the ABCD study: From substance use to a broad NIH collaboration. *Developmental cognitive neuroscience* 2018; **32:** 4-7.

29. Holland D, Frei O, Desikan R, Fan C-C, Shadrin AA, Smeland OB *et al.* Beyond SNP heritability: Polygenicity and discoverability of phenotypes estimated with a univariate Gaussian mixture model. *PLOS Genetics* 2020; **16**(5)**:** e1008612.

30. Watanabe K, Taskesen E, van Bochoven A, Posthuma D. Functional mapping and annotation of genetic associations with FUMA. *Nature Communications* 2017; **8**(1)**:** 1826.

31. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 2010; **26**(6)**:** 841-842.

32. The Genotype-Tissue Expression (GTEx) pilot analysis: Multitissue gene regulation in humans. *Science (New York, NY)* 2015; **348**(6235)**:** 648.

33. Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants. *Nat Genet* 2014; **46**(3)**:** 310-315.

34. Ripke S, Neale BM, Corvin A, Walters JTR, Farh K-H, Holmans PA *et al.* Biological insights from 108 schizophrenia-associated genetic loci. *Nature* 2014; **511**(7510)**:** 421-427.

35. Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst* 2015; **1**(6)**:** 417-425.

36. Johnson NL, Kemp AW, Kotz S. *Univariate discrete distributions*, vol. 444. John Wiley & Sons2005.

37. Li M, Santpere G, Imamura Kawasawa Y, Evgrafov OV, Gulden FO, Pochareddy S *et al.* Integrative functional genomic analysis of human brain development and neuropsychiatric risks. *Science (New York, NY)* 2018; **362**(6420).

38. Gamazon ER, Wheeler HE, Shah KP, Mozaffari SV, Aquino-Michaels K, Carroll RJ *et al.* A gene-based association method for mapping traits using reference transcriptome data. *Nature Genetics* 2015; **47**(9)**:** 1091-1098.

39. Barbeira AN, Dickinson SP, Bonazzola R, Zheng J, Wheeler HE, Torres JM *et al.* Exploring the phenotypic consequences of tissue specific gene expression variation inferred from GWAS summary statistics. *Nature Communications* 2018; **9**(1)**:** 1825.

40. BrainSeq: Neurogenomics to Drive Novel Target Discovery for Neuropsychiatric Disorders. *Neuron* 2015; **88**(6)**:** 1078-1083.

41. Fromer M, Roussos P, Sieberts SK, Johnson JS, Kavanagh DH, Perumal TM *et al.* Gene expression elucidates functional impact of polygenic risk for schizophrenia. *Nature Neuroscience* 2016; **19**(11)**:** 1442-1453.

42. Collado-Torres L, Burke EE, Peterson A, Shin J, Straub RE, Rajpurohit A *et al.* Regional Heterogeneity in Gene Expression, Regulation, and Coherence in the Frontal Cortex and Hippocampus across Development and Schizophrenia. *Neuron* 2019; **103**(2)**:** 203-216.e208.

43. Bulik-Sullivan B, Finucane HK, Anttila V, Gusev A, Day FR, Loh P-R *et al.* An atlas of genetic correlations across human diseases and traits. *Nature Genetics* 2015; **47**(11)**:** 1236-1241.

44. Maurano MT, Humbert R, Rynes E, Thurman RE, Haugen E, Wang H *et al.* Systematic localization of common disease-associated variation in regulatory DNA. *Science (New York, NY)* 2012; **337**(6099)**:** 1190-1195.

45. Kim M, Haney JR, Zhang P, Hernandez LM, Wang L-k, Perez-Cano L *et al.* Brain gene co-expression networks link complement signaling with convergent synaptic pathology in schizophrenia. *Nature Neuroscience* 2021; **24**(6)**:** 799-809.

46. Carpenter AE, Sabatini DM. Systematic genome-wide screens of gene function. *Nature Reviews Genetics* 2004; **5**(1)**:** 11-22.

47. Jaffe AE, Straub RE, Shin JH, Tao R, Gao Y, Collado-Torres L *et al.* Developmental and genetic regulation of the human cortex transcriptome illuminate schizophrenia pathogenesis. *Nature Neuroscience* 2018; **21**(8)**:** 1117-1125.

48. Luo Q, Chen Q, Wang W, Desrivières S, Quinlan EB, Jia T *et al.* Association of a Schizophrenia-Risk Nonsynonymous Variant With Putamen Volume in Adolescents: A Voxelwise and Genome-Wide Association Study. *JAMA psychiatry* 2019; **76**(4)**:** 435-445.

49. Matoba N, Love MI, Stein JL. Evaluating brain structure traits as endophenotypes using polygenicity and discoverability. *Hum Brain Mapp* 2020.

50. Fanous AH, Kendler KS. Genetic heterogeneity, modifier genes, and quantitative phenotypes in psychiatric illness: searching for a framework. *Molecular psychiatry* 2005; **10**(1)**:** 6-13.

51. Meer Dvd, Kaufmann T, Shadrin AA, Makowski C, Frei O, Roelfs D *et al.* The genetic architecture of human cortical folding. *Science Advances* 2021; **7**(51)**:** eabj9446.

52. Farkas LM, Huttner WB. The cell biology of neural stem and progenitor cells and its significance for their proliferation versus differentiation during mammalian brain development. *Curr Opin Cell Biol* 2008; **20**(6)**:** 707-715.

53. Oitzl MS, Champagne DL, van der Veen R, de Kloet ER. Brain development under stress: hypotheses of glucocorticoid actions revisited. *Neurosci Biobehav Rev* 2010; **34**(6)**:** 853-866.

54. Howes OD, Murray RM. Schizophrenia: an integrated sociodevelopmental-cognitive model. *Lancet (London, England)* 2014; **383**(9929)**:** 1677-1687.

55. Seixas E, Barros M, Seabra MC, Barral DC. Rab and Arf proteins in genetic diseases. *Traffic* 2013; **14**(8)**:** 871-885.

56. Pfenninger KH. Plasma membrane expansion: a neuron's Herculean task. *Nat Rev Neurosci* 2009; **10**(4)**:** 251-261.

57. Pantelis C, Velakoulis D, McGorry PD, Wood SJ, Suckling J, Phillips LJ *et al.* Neuroanatomical abnormalities before and after onset of psychosis: a cross-sectional and longitudinal MRI comparison. *Lancet (London, England)* 2003; **361**(9354)**:** 281-288.

58. Kirov G, Rees E, Walters JTR, Escott-Price V, Georgieva L, Richards AL *et al.* The Penetrance of Copy Number Variations for Schizophrenia and Developmental Delay. *Biological psychiatry* 2014; **75**(5)**:** 378-385.

59. Warland A, Kendall KM, Rees E, Kirov G, Caseras X. Schizophrenia-associated genomic copy number variants and subcortical brain volumes in the UK Biobank. *Molecular psychiatry* 2020; **25**(4)**:** 854-862.

60. Mallard TT, Liu S, Seidlitz J, Ma Z, Moraczewski D, Thomas A *et al.* X-chromosome influences on neuroanatomical variation in humans. *Nature Neuroscience* 2021; **24**(9)**:** 1216-1224.

## Figure Legends

**Figure 1. Shared genetic basis between schizophrenia and eight brain volumetric phenotypes.**

**a**., The MiXeR-estimated heritability, polygenicity, and discoverability for each phenotype. The SNP heritability is a sum of effect across all trait-influencing variants. The polygenicity of each trait is represented by the number of trait-influencing variants to explain 90% heritability. The discoverability of each trait is defined as effect size per trait-influencing variant. A trait-influencing variant is defined as a common variant associated with the trait of interest after controlling for linkage disequilibrium (LD). **b**., The MiXeR-modeled the number of shared and trait-specific trait-influencing variants in thousands. The colors reflect the shared (gray) and trait-specific (in color) trait-influencing variants. The standard deviations are shown in bracket. **c**., The LD Score-estimated genetic correlation and MiXeR-modeled proportion of shared variants with concordant effect direction on both traits. Only a significant false discovery rate-corrected P value (FDR) of genetic correlation is labeled within the figure. The error bar of genetic correlation presents standard error. In the proportion of concordant shared variants, the error bar reflects standard deviation. All details are provided in supplementary Table 1. SCZ: schizophrenia; ICV: intracranial volume.

**Figure 2. Individual shared associations between schizophrenia and brain volumetric phenotypes and relevant gene expression patterns.**

**a**., The Manhattan plot of shared associations between schizophrenia and brain volumetric phenotypes. The x-axis stands for the chromosomal number and position and the y-axis represents -log10 transformed conjunctional false discovery rate (FDR) values with a dotted horizontal line reflecting significance. Each dot represents a SNP and the border indicates a lead SNP. **b**., The SNP sign test for shared lead SNPs. $N_{SNP}$ counts the number of lead SNPs available in the independent

datasets and $N_{CON}$ calculates the number of lead SNPs showing consistent effect between discovery and independent cohorts. The x-axis reflects the proportion of consistency. The solid points indicate significance of SNP sign test after FDR correction with the P value showed in FDR. **c**., Co-expression patterns among shared genes. Shared genes between schizophrenia and four brain volumetric phenotypes, including accumbens, hippocampus, pallidum, putamen, and ICV, are enriched in five co-expression modules reported by PsychoENCODE human brain development study. Those five co-expression modules were labeled as modules (MEs) 2, 3, 5, 7, and 7 to be consistent with the PsychoENCODE study. **d.**, The five sets of co-expressed shared genes showed distinct temporal expression trajectory. The cluster red, pink, azure, blue, and black represented the shared genes enriched in MEs 2, 3, 5, 7, and 9 as respectively. The whole-brain expression level (y-axis) represents log2-median transformed mean expression value across brain tissues. The whole-brain expression trajectory was shown by a fitted non-linear LOESS regression line with the 95% confidence interval presented by shaded areas. The age (x-axis) was presented using post-conception days, which were further divided into five stages—fetal development (8 PCW <= age < 28 PCW), infancy (40 PCW <= age < 1 PY), childhood (1 PY<= age < 12 PY), adolescence (12 PY <= age < 20 PY), and adulthood (20 PY <= age <= 40 PY). Details of those five stages and nine anatomic tissues are provided in Supplementary Method. ICV: intracranial volume.

**Figure 3. The flowchart of gene-based functional analysis.**

For all functional enrichment analysis in the present study, we excluded shared genes located in a complex linkage disequilibrium (LD) region, including major histocompatibility complex (MHC), Chromosome 8p23.1 deletion (8p23.1), microtubule-associated protein tau (MAPT), and apolipoprotein E (APOE) regions. The gene-based association analysis by PrediXcan was illustrated using the pair of schizophrenia and thalamus as an example. A green rectangle represents a process, with a decision was indicated by a diamond shape. The input and output are reflected in purple

parallelograms. The data from PsychoENCODE, GTEx, and the Adolescent Brain Cognitive Development (ABCD) cohort are shown in database symbols in light blue. More details are provided in Methods. MSigDB: Molecular Signatures Database; eQTLs: expression quantitative trait loci.

**Figure 4. Two candidate genes for schizophrenia and thalamus volume in late childhood.**

**a.,** The expression-trait associations between two candidate genes (*ARL17A* and *CRHR1*) and traits of interest. The x-axis indicates data resources and corresponding tissue, as well as the traits of interest labeled below. The expression-trait associations were genetically imputed from PrediXcan and S-PrediXcan, and differential gene expression profiles were based on CommonMind Consortium (CMC) and BrainSeq Phase 2 studies. The asterisk implies the significance. DLPFC: dorsolateral prefrontal cortex. **b.,** Two candidate genes were mapped by a shared locus (chr17:43859929-44865603) between schizophrenia and thalamus. The violin plot represents the original GWAS Z scores of candidate SNPs within this shared locus. **c.,** The brain *cis*-expression quantitative trait locus (*cis*-eQTL) between the lead SNP and each candidate gene. The significant *cis*-eQTLs are shown in colors with indication of the direction of effect. The significant brain *cis*-eQTLs were annotated from GTEx portal (v8). **d.,** The lifespan temporal expression trajectory of two candidate genes in nine brain tissues. The brain tissue expression (y-axis) was log2-median transformed. The expression trajectory for each brain tissue was revealed by a fitted non-linear LOESS regression line with the 95% confidence interval (shaded areas). The time span was shown using five stages, including fetal development (8 PCW <= age < 28 PCW), infancy (40 PCW <= age < 1 PY), childhood (1 PY<= age < 12 PY), adolescence (12 PY <= age < 20 PY), and adulthood (20 PY <= age <= 40 PY). ICV: intracranial volume.

**Figure 1. Shared genetic basis between schizophrenia and eight brain volumetric phenotypes.**

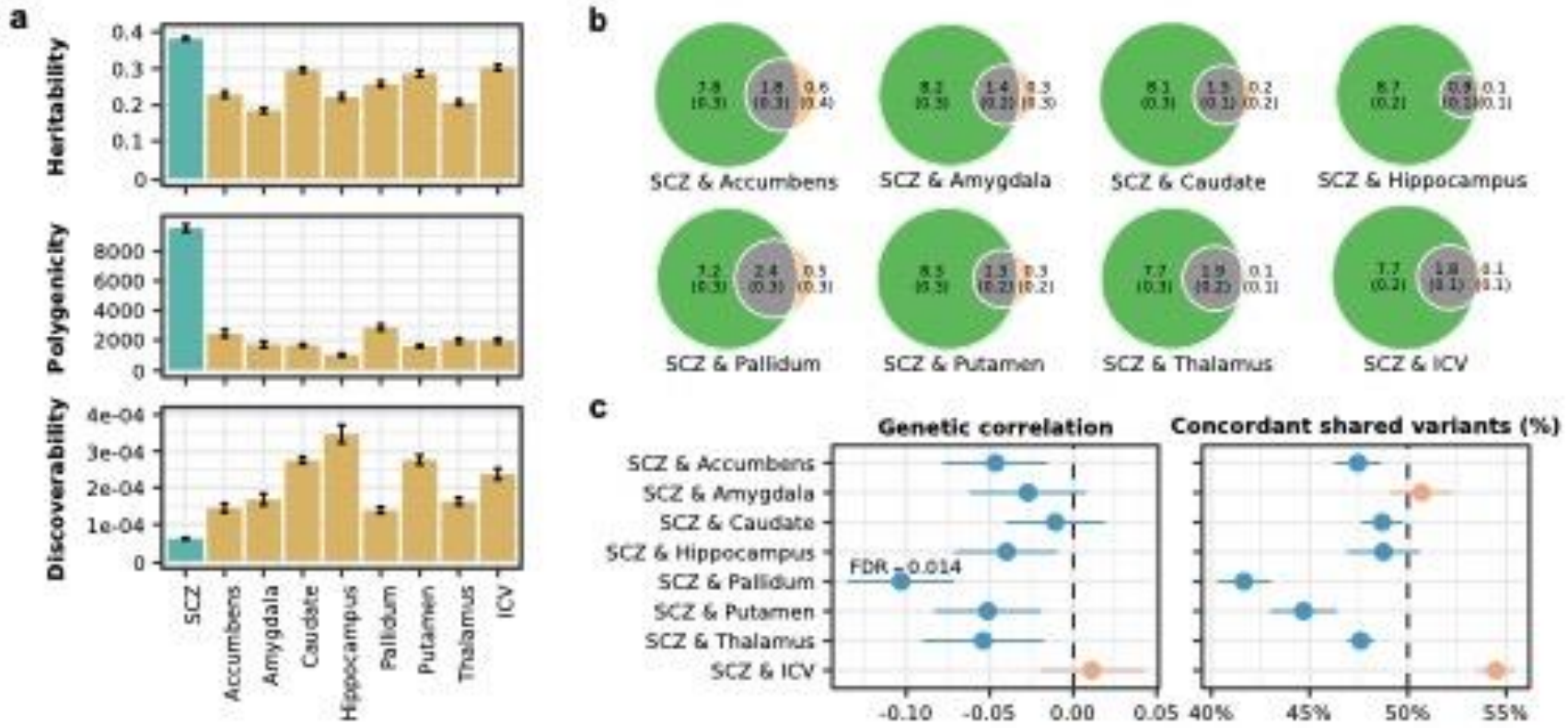

**Figure 2. Individual shared associations between schizophrenia and brain volumetric phenotypes and relevant gene expression patterns.**

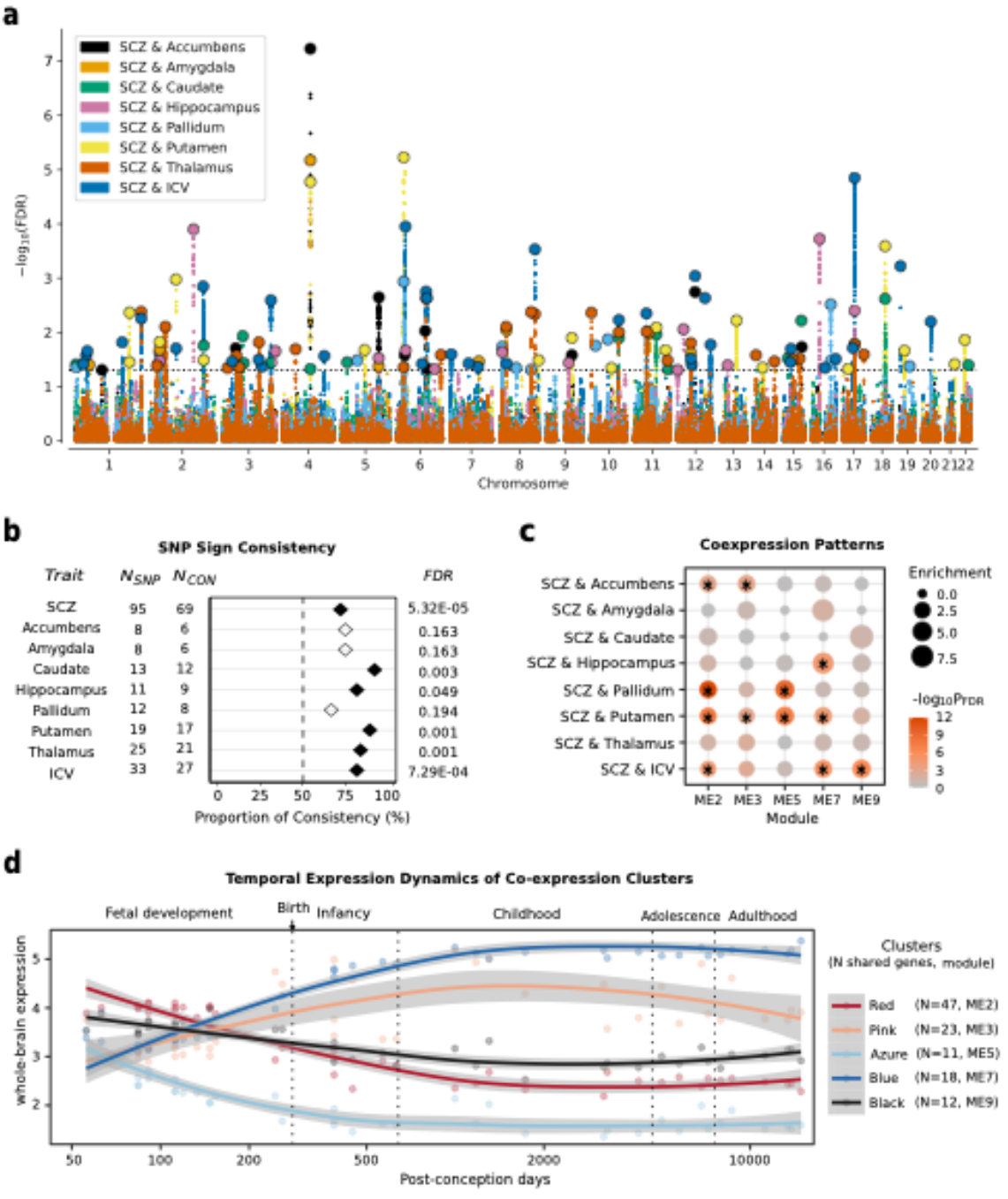

**Figure 3. The flowchart of gene-based functional analysis.**

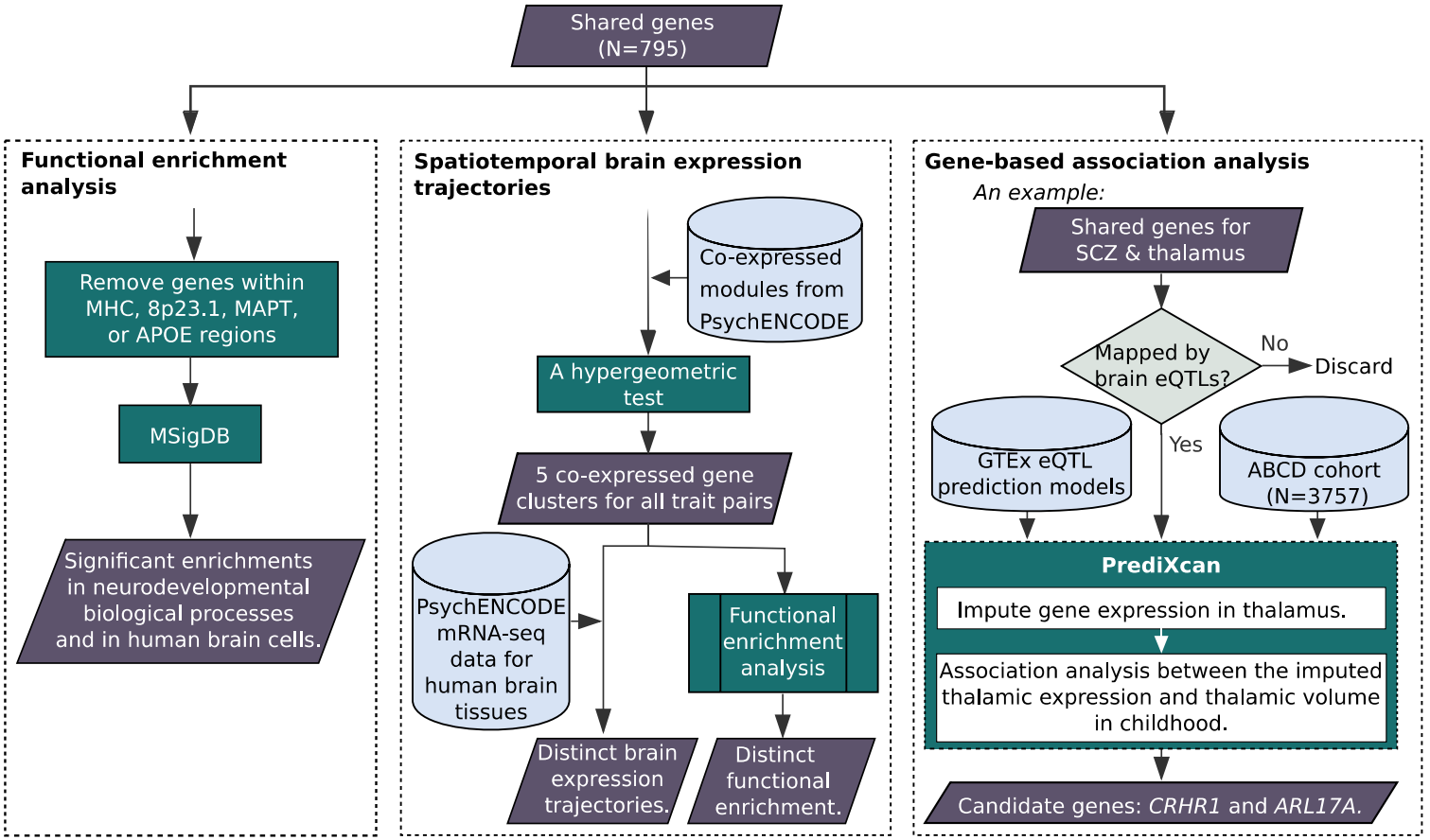

**Figure 4. Two candidate genes for schizophrenia and thalamus volume in late childhood.**