GWAS SNPs Impact Shared Regulatory Pathways Amongst Multimorbid Psychiatric Disorders and Cognitive Functioning

Evgeniia Golovina[1,2], Mark H. Vickers[1], Christopher D. Erb[3] and Justin M. O'Sullivan[1,2]*

[1] Liggins Institute, University of Auckland, Auckland, New Zealand, [2] A Better Start National Science Challenge, Auckland, New Zealand, [3] School of Psychology, University of Auckland, Auckland, New Zealand

Background: Epidemiological research has reported that attention-deficit hyperactivity disorder (ADHD), anxiety, bipolar disorder (BD), schizophrenia (SCZ), and unipolar depression (UD) are multimorbid conditions that are typically accompanied by cognitive advantages or deficits, suggesting that common biological mechanisms may underlie these phenotypes. Genome-wide association studies (GWAS) have identified single-nucleotide polymorphisms (SNPs) associated with psychiatric disorders and cognitive functioning. However, the mechanisms by which these SNPs contribute to multimorbidities amongst psychiatric and cognitive phenotypes remains largely unknown.

Objective: To identify shared regulatory mechanisms amongst multimorbid psychiatric disorders and cognitive functioning.

Methods: We integrated data on 3D genome organization, expression quantitative trait loci (eQTLs), and pathway analyses to identify shared and specific regulatory impacts of 2,893 GWAS SNPs ($p < 1 \times 10^{-6}$) associated with ADHD, anxiety, BD, SCZ, UD, and cognitive functioning on genes and biological pathways. Drug-gene interaction analysis was performed to identify potential pharmacological impacts on these genes and pathways.

Results: The analysis revealed 33 genes and 62 pathways that were commonly affected by tissue-specific gene regulatory interactions associated with all six phenotypes despite there being no common SNPs in our original dataset. The analysis of brain-specific regulatory connections revealed similar patterns at eQTL and eGene levels, but no pathways shared by all six phenotypes. Instead, pairwise overlaps and individualized pathways were identified for psychiatric and cognitive phenotypes in brain tissues.

Conclusions: This study offers insight into the shared genes and biological pathways that are affected by tissue-specific regulatory impacts resulting from psychiatric- and cognition-associated genetic variants. These results provide limited support for the "p-factor" hypothesis for psychiatric disorders and potential mechanisms that explain drug

side-effects. Our results highlight key biological pathways for development of therapies that target single or multiple psychiatric and cognitive phenotypes.

INTRODUCTION

Attention-deficit hyperactivity disorder (ADHD), anxiety, bipolar disorder (BD), schizophrenia (SCZ), and unipolar depression (UD) are highly prevalent psychiatric disorders (1). Epidemiological research has reported that ADHD (2), anxiety (2, 3), BD (2), SCZ (4), and UD (5) are highly heritable (6) and represent multimorbid conditions that can be accompanied by cognitive advantages or deficits (7). Consistent with this, multiple genetic studies of psychiatric disorders and cognitive functioning have revealed a surprising degree of genetic overlap amongst these phenotypes, with the highest genetic correlations observed between SCZ and BP ($r_g = 0.70$), ADHD and UD ($r_g = 0.44$), BP and UD ($r_g = 0.36$), SCZ and UD ($r_g = 0.34$) (8–10). Collectively, this suggests the possible existence of common biological mechanisms underlying these phenotypes. Recent studies have proposed the term p-factor to represent common features of psychiatric disorders and an individual's overall liability to psychopathology (11, 12). Epidemiological evidence on the persistence, co-occurrence, sequential psychiatric multimorbidities, and genomic research tends to support the p-factor hypothesis. However, how the genetic architectures of psychiatric phenotypes contribute to the psychiatric multimorbidity (i.e., the "p-factor") and the relationship with cognitive functioning is incompletely understood (11, 12). Furthermore, assuming that common drug targets may trigger similar therapeutic effects that induce similar signaling pathways and, thus, similar side effects, understanding the p-factor will be useful for prognosing therapeutic approaches and predicting drug side-effects.

Genome-wide association studies (GWAS) have identified thousands of single nucleotide polymorphisms (SNPs) that are associated with psychiatric and cognitive phenotypes (6). The majority of these genetic variants are non-coding and are likely to be expression quantitative trait loci (eQTLs) associated with the regulation of gene expression (13, 14). However, identifying the functional impacts of these eQTL SNPs is a significant hurdle (15).

Tissue-specific regulatory regions harbor substantial genetic risk for complex phenotypes (16–18). As such, decoding tissue-specific regulatory networks is an important step toward understanding the molecular mechanisms underlying psychiatric and cognitive phenotypes and identifying novel pathway-based therapeutic strategies.

The three-dimensional (3D) organization of the genome includes cell-type and tissue-specific spatial interactions between regulatory regions (marked by eQTL SNPs) and the genes that they control (hereafter eGenes) (19). These spatial interactions include cis- (<1 Mb between regulatory region and target eGene) and trans-acting (>1 Mb) intra- and interchromosomal connections. Recent studies have demonstrated the potential of integrating phenotype-associated SNPs, genome structure, and eQTLs for the identification of functional regulatory interactions in SCZ (20, 21). However, if or how genetic variation impacts on the regulation of shared or specific biological pathways to contribute to the development of psychiatric and cognitive multimorbidities remains unknown.

In this study, we integrated data on 3D genome organization and eQTLs to identify the tissue-specific spatial regulatory impacts of SNPs associated with cognitive functioning and five psychiatric disorders (i.e., ADHD, anxiety, BD, UD, SCZ). We hypothesized that this approach would enhance the discovery of shared characteristics (i.e., eGenes and biological pathways) among psychiatric and cognitive phenotypes, thereby potentially disclosing the mechanism for the p-factor that contributes to the observed multimorbidities.

MATERIALS AND METHODS

Ethical Approvals for Data Access

Data access approval was obtained from dbGaP (*https://www.ncbi.nlm.nih.gov/gap/*) for Hi-C data sets for HeLa (project #18446: "Finessing predictors of cognitive development (part 2)"), cortical plate and germinal zone neurons (project #16489: "Finessing predictors of cognitive development").

Identification of Psychiatric and Cognition-Associated SNPs

Single-nucleotide polymorphisms (SNPs) associated (suggestive $p < 1 \times 10^{-6}$ cut-off was used) with ADHD, anxiety, BD, UD, SCZ, and cognitive functioning were downloaded from the GWAS Catalog (*www.ebi.ac.uk/gwas/*; 07/12/2018 and 14/07/2018; **Supplementary Spreadsheet 1**). Nine cognitive traits (i.e., intelligence, information processing speed, cognition, reading ability, reasoning, mathematical ability, infant expressive language ability, language performance, and speech perception) were combined to create the "cognitive functioning" category. Genomic positions and annotations of SNPs were obtained for the human genome build hg19 release 75 (GRCh37). To obtain genomic region information for SNPs, functional SNP annotation was performed using wANNOVAR (22, 23) (*http://wannovar.wglab.org/*).

Hi-C Data Processing

Three-dimensional (3D) chromatin interactions in the nucleus can bring genes and distant regulatory elements into close spatial proximity and, thus, significantly influence gene expression regulation. Chromosome Conformation Capture (3C)-based technologies, such as Hi-C, capture these chromatin interactions in a genome-wide fashion and enable the investigation of the

relationship between genome organization and genome activity. In order to study spatial tissue-specific regulatory interactions, we analyzed 28 cell type and tissue-specific Hi-C chromatin interaction libraries, including high-resolution data (24) and brain specific datasets (19, 25, 26) (**Supplementary Table 1**). Hi-C interaction data were downloaded from GEO (*https://www.ncbi.nlm.nih.gov/geo/*). Raw data were analyzed according to Rao et al. (24) [Juicer (27), version 1.5, *https://github.com/aidenlab/juicer*] to generate Hi-C libraries. This analysis pipeline included BWA alignment of paired-end reads onto the hg19 reference genome (BWA, version 0.7.15), merging paired-end read alignments, and removing chimeric, unmapped and duplicated reads. The remaining read pairs we refer to as "contacts." Only Hi-C libraries that contain >90% alignable unique read pairs, and >50% unique contacts (<40% duplication rate) within the total sequenced read pairs were included in the analysis. Files containing cleaned Hi-C contacts locations (i.e., *_merged_nodups.txt files) were processed to obtain Hi-C chromatin interaction libraries in the following format: read name, str1, chr1, pos1, frag1 mapq1, str2, chr2, pos2, frag2, mapq2 (str = strand, chr = chromosome, pos = position, frag = restriction site fragment, mapq = mapping quality score, 1 and 2 correspond to read ends in a pair). Reads where both ends had a mapq $\geq$ 30 were included in the final library. Hi-C chromatin interactions represent all captured pairs of interacting restriction fragments in the genome (**Supplementary Figure 1**). As such, restriction fragments were used to identify regulatory interactions between SNPs and genes.

Identification of Spatial Regulatory Interactions Using CoDeS3D

CoDeS3D (28) (*https://github.com/Genome3d/codes3d-v1*) was used to identify spatial regulatory interactions between genes and phenotype-associated SNPs (**Supplementary Figure 1**). Briefly, the reference genome (hg19) was digested using the restriction enzyme that was used in the Hi-C library preparation (i.e., MboI or HindIII). The restriction fragments containing phenotype-associated SNPs were identified. Next, the algorithm captured the restriction fragments interacting with the SNP-containing fragments in each of 28 Hi-C chromatin interaction libraries. Only interactions between SNP-containing fragments and restriction fragments overlapping a gene (defined using GENCODE transcript model version 19) were further analyzed (hereafter SNP-gene pairs). Next, the GTEx database (*https://gtexportal.org/*, GTEx multi-tissue dataset v7, **Supplementary Table 2**) was queried to identify only those spatial SNP-gene pairs, where expression levels of the gene (i.e., eGene) were associated with a SNP (i.e., eQTL SNP) in one or more of 48 tissues. Lastly, significant eQTL SNP-eGene-tissue interactions were identified using the Benjamini-Hochberg FDR control algorithm (29) to adjust eQTL p-values (FDR < 0.05). The FDR correction was performed across all tissue types. Significant eQTL-eGene interactions were defined as brain-specific if: (1) the spatial interaction was present in one or more brain-specific Hi-C dataset; and (2) the eQTL occurred in brain and spinal cord tissues.

Bootstrapping Analysis

We performed bootstrapping ($N = 10,000$ iterations) with two reference eGene sets to test if the observed spatial eGene associations and overlaps were non-random. Reference set 1 contained the list of all genes in the genome (GENCODE transcript model version 19). Reference set 2 included the genes that were identified as interacting with the phenotype associated SNP containing fragments within the Hi-C libraries, to account for any potential bias in Hi-C gene coverage (i.e., the output of the find_genes module). Each bootstrapping iteration generated samples containing an equivalent number of randomly selected eGenes, as identified for each of the phenotypes being tested. The number of shared eGenes amongst phenotypes was counted for each bootstrap iteration. After 10,000 iterations we counted those instances where the number of shared eGenes in the bootstrapped overlap (*eGenes_bootstrapped*) is greater than or equal to the number of shared eGenes in the observed overlap (*eGenes_observed*). The p-value was calculated as the sum of these instances divided by the total number of iterations N,

$$p\ value = \frac{\sum(eGenes_{bootstrapped} \geq eGenes_{observed})}{N} \quad (1)$$

For bootstrapping, if the $p < 0.01$ we reject the null hypothesis (i.e., the assumption that the observed eGene overlap is due to chance) and accept the alternative hypothesis that the observed relationship is non-random.

Functional Enrichment Analysis

Functional gene enrichment analysis was performed using the g:GOSt tool from g:Profiler (30) (*https://biit.cs.ut.ee/gprofiler/*). The g:GOSt tool maps genes to known functional annotations and identifies the most impacted and statistically significant enriched annotations. eGene enrichment was tested within the biological process, molecular function, and cellular component gene ontology terms. All known human genes were chosen as the statistical domain scope. The significance of the overrepresented GO terms was corrected using the SCS algorithm (30) (adjusted $p < 0.05$).

Pathway Analysis

eGenes were analyzed using iPathwayGuide (*https://www.advaitabio.com/ipathwayguide*, 09/13/2019) to identify enriched biological pathways. iPathwaysGuide uses the KEGG pathways database (31) and considers the role, position and relationships of each gene within a pathway to significantly reduce the number of false positives and identify truly impacted pathways. The FDR algorithm (29) was applied to correct p-values for multiple testing and determine significance at the pathway level (FDR < 0.05).

As for the gene analysis, bootstrapping ($N = 10,000$ iterations) was performed to determine the significance of the observed pathway overlaps. Here the KEGG pathways database (31) ($n = 536$; iPathwaysGuide, 22/01/2020) was used as the reference set.

Correlation Analysis

A Pearson's correlation analysis was performed to measure the association between GTEx tissue sample size and the number of cis- and trans-acting eQTL-eGenes interactions.

Drug-eGene Interaction Analysis

The Drug Gene Interaction database (32) (DGIdb, v3.0.2—sha1 ec916b2, *http://www.dgidb.org/*) consolidates and organizes known drug-gene interactions and gene druggability information from 30 databases and clinical trials. The collapsed resources include: DrugBank, PharmGKB, ChEMBL, FDA Biomarkers, Guide To Pharmacology, Jax-Clinical Knowledgebase, TDG Clinical Trials, CIViC, and CancerCommons. DGIdb supports over 40,000 genes and 10,000 drugs that are involved in over 15,000 drug-gene interactions. We queried DGIdb to identify drugs (their effects and mechanisms of action) that target the eGene products.

RESULTS

SNPs Associated With Psychiatric and Cognitive Phenotypes Regulate Distant Genes

We hypothesized that multimorbidity is driven by genetic variants (e.g., SNPs, structural variants, indels) that regulate tissue-specific expression of genes that co-occur within specific biological pathways and thus affect the phenotype (**Figure 1**). CoDeS3D (28) was used to integrate genome structure and eQTL data to identify tissue-specific spatial eQTLs for SNPs ($n = 2,893$) associated ($p < 1 \times 10^{-6}$) with cognitive functioning, ADHD, anxiety, BD, UD, and SCZ. We identified a total of 45,269 significant tissue-specific eQTL-eGene pairs (FDR < 0.05) from 48 different human tissues (**Figure 2**, **Supplementary Spreadsheet 2**). In total, 2,088 SNPs were identified as eQTLs (**Figure 2**). Approximately 60% of the SNPs associated with each phenotype were eQTLs (**Supplementary Figure 2A**). The majority of the 2,088 eQTLs we identified were located within introns and intergenic regions (**Supplementary Figure 2B**). The patterns of cis- and trans-acting regulatory interactions formed by eQTLs within introns and intergenic regions were similar across all phenotypes (**Supplementary Figures 3A,B**).

SNPs and eQTLs Are Mostly Unique to Individual Phenotypes

Previous studies have described a limited number of SNPs associated with combined phenotypes [i.e., SCZ+cognition (7), BD+cognition (7) and BD+SCZ (33)]. Therefore, we intersected the SNP and spatial eQTL sets associated with ADHD, anxiety, BD, UD, SCZ, and cognitive functioning to identify shared genetic variation between these phenotypes. We found that the total numbers of shared SNPs across phenotype combinations are similar to those for shared eQTLs (**Figures 3A,B**). No SNPs, or eQTLs were common to all phenotypes (**Figures 3A,B**, respectively). However, there were limited eQTL overlaps amongst the psychiatric disorders and between psychiatric disorders and cognition (**Figure 3B**). We did not observe any overlapping eQTLs for ADHD and cognitive phenotypes. The limited SNP and eQTL overlaps we observed supports the hypothesis that regulatory impacts on genes and pathways can make a contribution to multimorbidity.

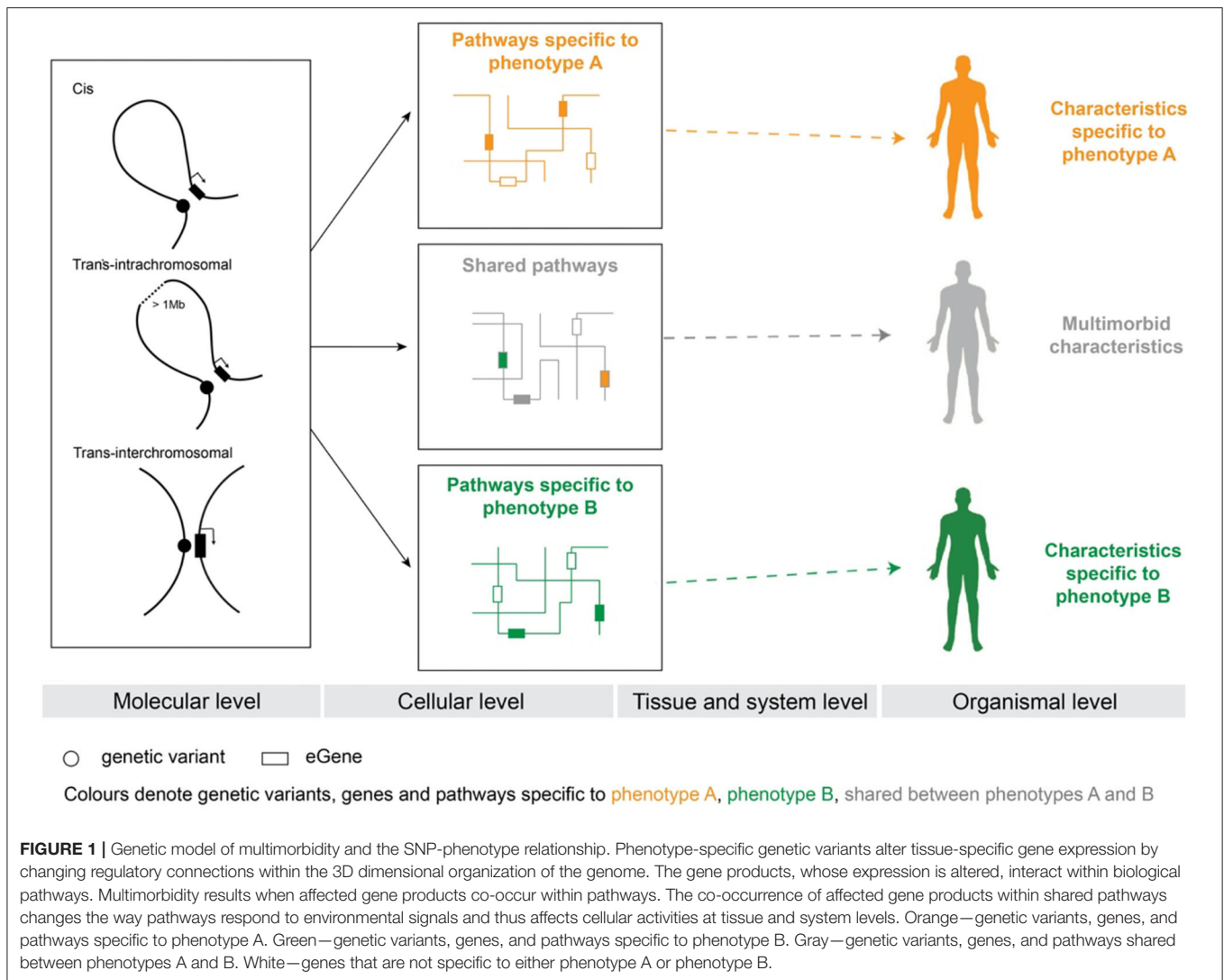## Psychiatric and Cognitive Phenotypes Share Common eGenes

Multimorbidities among psychiatric disorders and cognition could result from the regulatory effects of shared or phenotype-specific eQTLs on shared gene targets. Shared eGenes ($n = 33$) were identified for eQTLs associated with all six phenotypes (**Figure 3C**). SCZ and UD shared the greatest number of eGenes ($n = 374$). ADHD and cognition have 58 shared eGenes despite having no shared eQTLs (**Figures 3B,C**). Bootstrapping simulations ($n = 10,000$) confirmed that the overlaps were significant ($p < 0.001$; **Supplementary Figure 4A**). The expression levels of the 33 eGenes (**Supplementary Table 3**) that were shared across all phenotypes are associated with eQTLs that are located within three loci (chr3:52256696-53455568, chr6:25177507-32914725, and chr10:103816827-105039240; **Supplementary Figure 5**).

Functional assignments for phenotype-associated SNPs are typically made to the closest gene to the phenotype-associated variant (9, 21, 34–36). Our analysis of eQTL-eGenes spatial connections showed that 5–7% are explained by associations with the closest gene (**Supplementary Spreadsheet 2**). For example, ADHD, BD, UD, and SCZ (34) -associated rs2535629, which is located within an intron of *ITIH3*, is an eQTL for *ITIH4* in the brain, adipose tissues, and the cardiovascular system. However, rs2535629 is also associated with spatial regulatory interactions with another 12 eGenes (*ITIH3*, *PPM1M*, *GNL3*, *RAF1*, *MUSTN1*, *NEK4*, *NT5DC2*, *PBRM1*, *RBMS3*, *SFMBT1*, *TMEM110*, *WDR82*; **Supplementary Table 4**). Thus, incorporating data on spatial genome organization enables the identification of local and distal eQTL-gene connections that can potentially contribute to multimorbid phenotypes.

Psychiatric and Cognitive Phenotypes Share Common Biological Pathways

The pleiotropic effects of eQTLs and co-occurrence of the affected eGenes within biological processes and pathways could contribute to the underlying multimorbid conditions (**Figure 1**). Gene ontology (GO) analysis of the 33 shared eGenes identified significant enrichment (adjusted $p < 0.05$) in gene expression, transcription, metabolic, biosynthetic, and regulatory processes (**Supplementary Figure 6**). Notably, ontological analyses of eGenes specific for each phenotype revealed associations with neurodevelopment (e.g., "nervous system development"), immune system processes (e.g., "immune response"), responses to environmental stimuli, and signal transduction that were common to all phenotypes (**Supplementary Spreadsheet 3**).

Pathway analysis using eGenes (shared and specific for each phenotype) identified 61 biological pathways that are shared by all phenotypes (**Figure 3D**). These shared pathways are associated with human diseases, signal transduction, neurodevelopment, learning, and immunity (**Supplementary Spreadsheet 4, Supplementary Table 5**). Bootstrapping ($n = 10,000$) confirmed that the observed overlap was highly significant ($p < 0.001$; **Supplementary Figure 7**) and not an artifact of the presence of shared eGenes within these pathways (**Supplementary Figure 8**).

FIGURE 1 | Genetic model of multimorbidity and the SNP-phenotype relationship. Phenotype-specific genetic variants alter tissue-specific gene expression by changing regulatory connections within the 3D dimensional organization of the genome. The gene products, whose expression is altered, interact within biological pathways. Multimorbidity results when affected gene products co-occur within pathways. The co-occurrence of affected gene products within shared pathways changes the way pathways respond to environmental signals and thus affects cellular activities at tissue and system levels. Orange—genetic variants, genes, and pathways specific to phenotype A. Green—genetic variants, genes, and pathways specific to phenotype B. Gray—genetic variants, genes, and pathways shared between phenotypes A and B. White—genes that are not specific to either phenotype A or phenotype B.
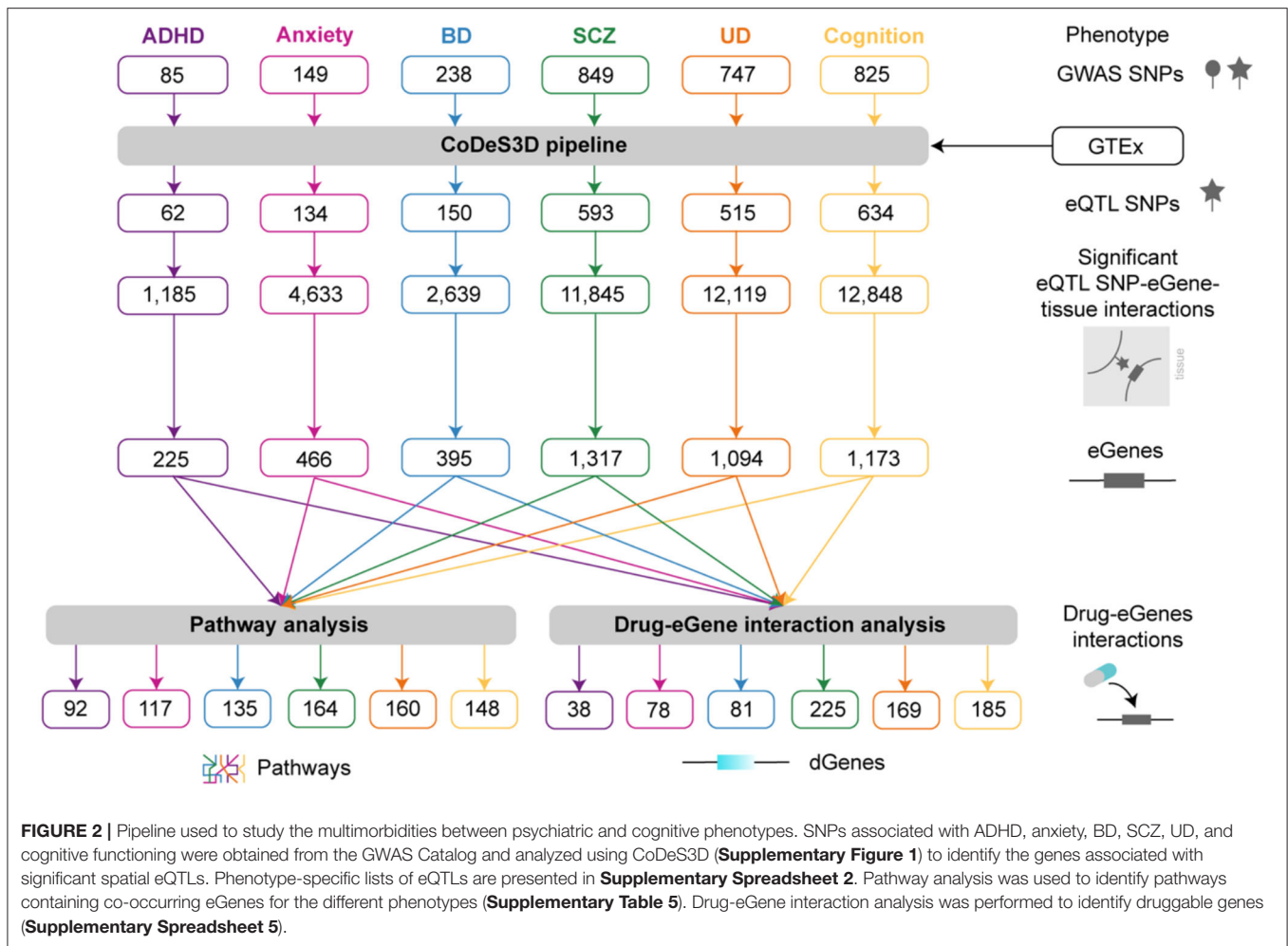
Notably, despite a high prevalence of comorbidity and substantial genetic correlation between anxiety and UD, only 6 pathways were shared between these two conditions (37–39).

The neurotrophin signaling pathway is important in developmental neurobiology (40). This pathway contained eGenes associated with eQTLs from all phenotypes (**Figure 4**). Most of the eGenes within the neurotrophin signaling pathway were regulated by trans-acting eQTLs (**Figure 4**). Dysregulation in the neurotrophin signaling cascade can impact downstream pathways, e.g., axon guidance and long-term potentiation pathways, which also contained co-occurring eGenes associated with cognition and psychiatric phenotypes (**Supplementary Figures 9**, **10**).

We also identified 12 biological pathways that were impacted by eGenes that were specific to individual psychiatric and cognitive phenotypes (**Figure 3D**, **Supplementary Table 5**).

Drug-Gene Interactions Identify Pharmaceutical Impacts on Multimorbidity

Drugs to prevent, stabilize, or slow the progression of psychiatric and cognitive conditions often have side effects consistent with known multimorbidities (41). Between 15–21% of the eGenes we identified encode proteins that are targeted by existing drugs (**Supplementary Spreadsheet 5**). Four eGenes (i.e., *AS3MT*, *FLOT1*, *HLA-A*, and *PBRM1*) are affected by eQTLs from all tested phenotypes and are current therapeutic targets for psychiatric disorders (**Supplementary Figure 11**). For example, everolimus targets PBRM1 (polybromo 1 protein, subunit of chromatin remodeling complex) that can be important in neural development (42). Notably, everolimus improves memory and learning but simultaneously aggravates depression and anxiety in mice (43). Therefore, we contend that these observations provide insights into potential mechanisms for drug side-effects associated with multimorbidities among psychiatric disorders and cognition.

FIGURE 2 | Pipeline used to study the multimorbidities between psychiatric and cognitive phenotypes. SNPs associated with ADHD, anxiety, BD, SCZ, UD, and cognitive functioning were obtained from the GWAS Catalog and analyzed using CoDeS3D (**Supplementary Figure 1**) to identify the genes associated with significant spatial eQTLs. Phenotype-specific lists of eQTLs are presented in **Supplementary Spreadsheet 2**. Pathway analysis was used to identify pathways containing co-occurring eGenes for the different phenotypes (**Supplementary Table 5**). Drug-eGene interaction analysis was performed to identify druggable genes (**Supplementary Spreadsheet 5**).

Spatial Regulatory eQTL Effects Are Tissue-Specific

Brain structural changes are commonly considered to be relevant to the development of psychiatric and cognitive phenotypes. However, these phenotypes are also associated with physiological changes at the level of the whole body [e.g., with impaired functioning of endocrine (44), immune (45, 46), and cardiometabolic (47) systems]. Therefore, the tissue-specificity of phenotype-associated spatial regulatory impacts may provide important insights into psychiatric and cognitive multimorbidities. The cis- and trans-acting eQTLs we identified were widely distribution across human tissues (**Supplementary Figure 12**). The numbers of cis and intra-chromosomal spatial eQTL SNP-eGene regulatory interactions correlated with the GTEx tissue sample size (**Supplementary Figure 13**). By contrast, inter-chromosomal eQTLs were not correlated with tissue sample size (**Supplementary Figure 13**). We detected greater numbers of cis (<1 Mb) and intra-chromosomal (≥1 Mb) eQTL-eGene interactions in thyroid and brain cerebellum tissues, than were predicted by the correlation curve (**Supplementary Spreadsheet 6**, **Supplementary Figure 13**). This is consistent with reported roles for the thyroid in the development of cognitive functions and psychiatric disorders (44, 48).

Brain-Specific Regulatory Associations Are Mostly Unique to Individual Phenotypes

ADHD, anxiety, BD, UD, SCZ, and cognitive functioning are widely regarded as brain-specific phenotypes. Therefore, we identified brain-specific eQTL-eGene regulatory interactions that are supported by spatial connections within brain-specific Hi-C datasets (**Supplementary Spreadsheet 7**). The brain-specific eQTLs and eGenes we identified were mostly unique to the individual psychiatric and cognitive phenotypes (**Figures 5A,B**). Consistent with our earlier findings across all tissues, we observed many shared pathways across different phenotype combinations. However, we observed more pairwise combinations and individualized pathways for the
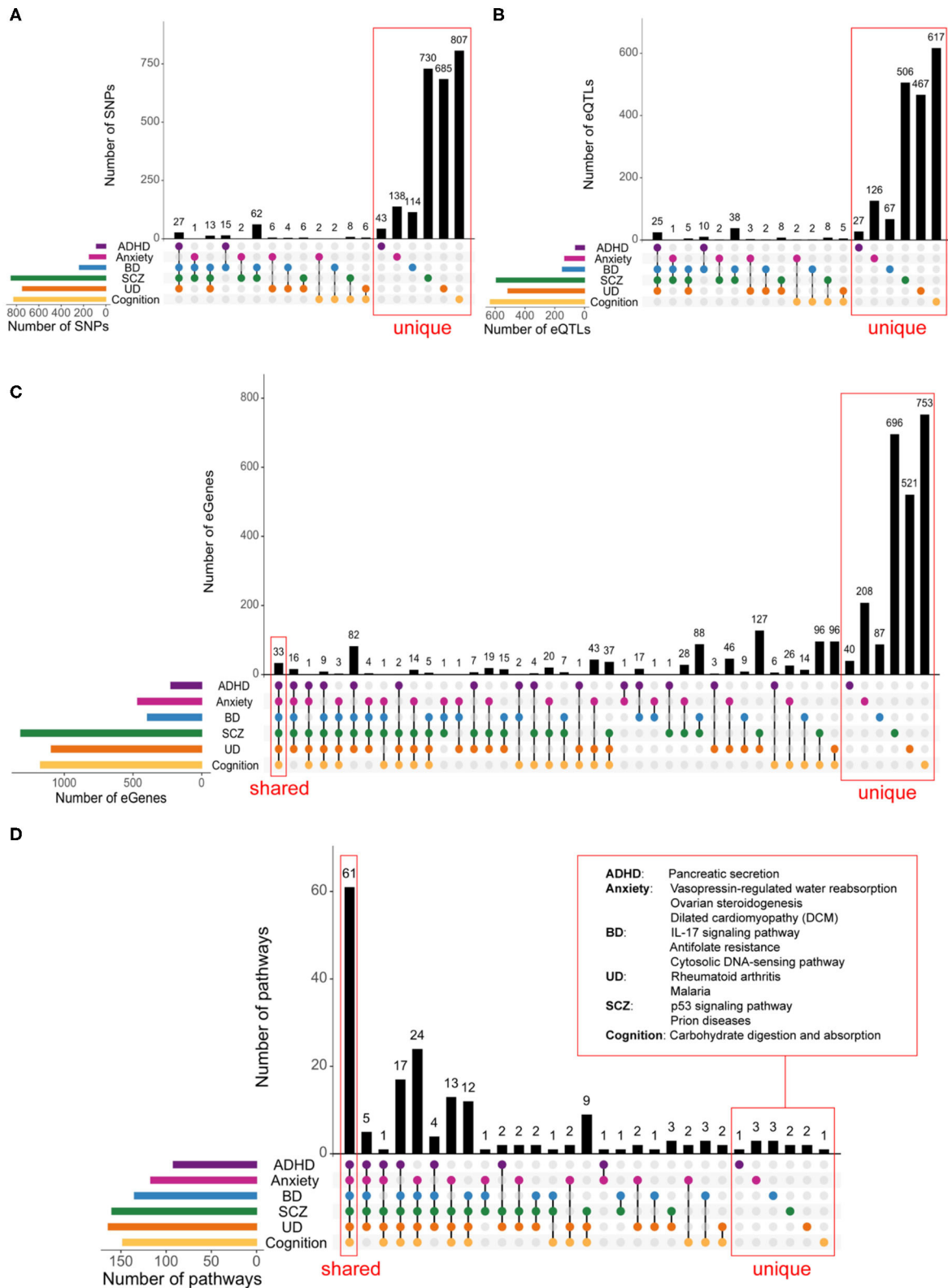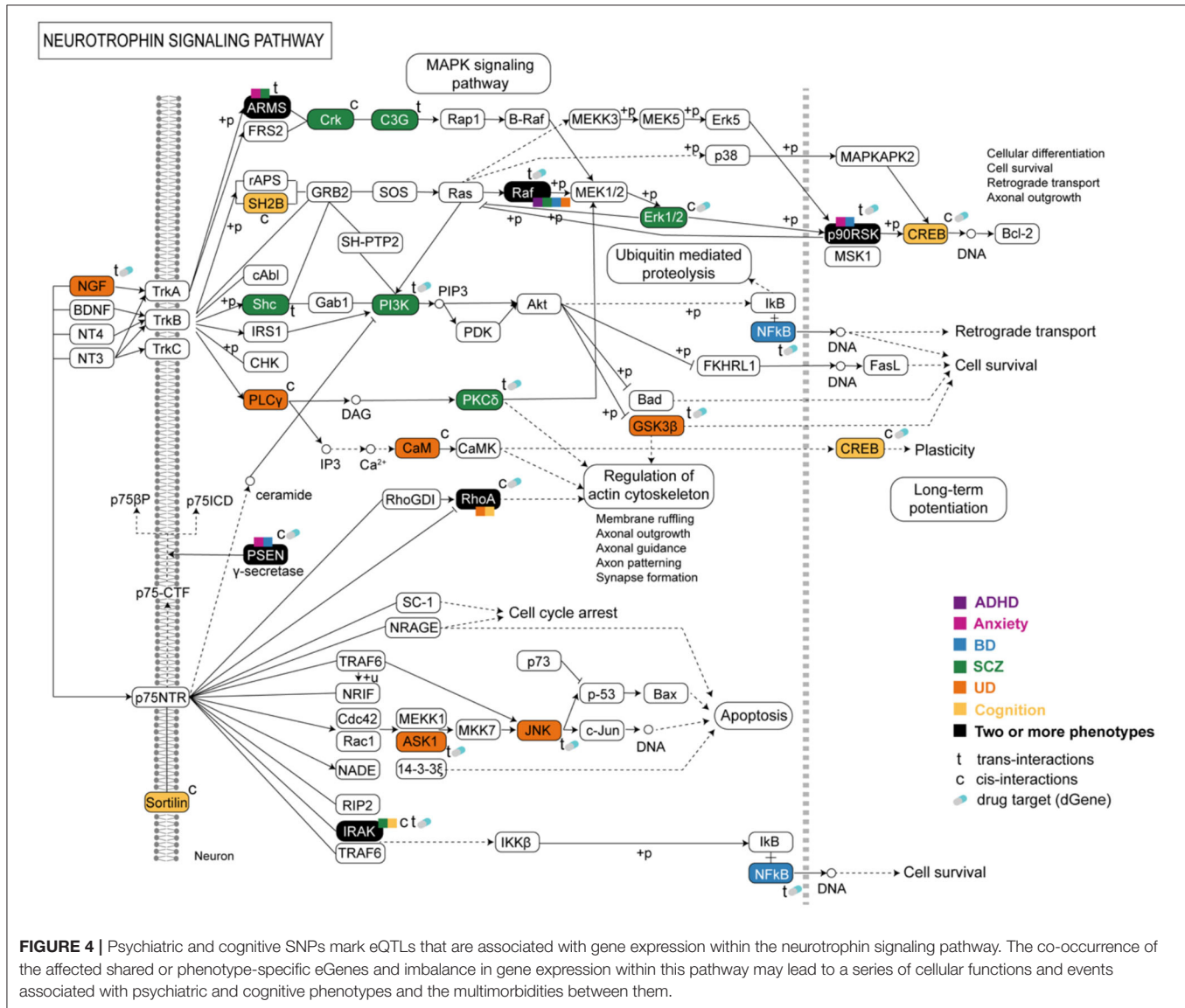
FIGURE 3 | Shared biological pathways link psychiatric disorders and cognition. Psychiatric disorders and cognitive functions have low levels of genetic similarity at the SNPs **(A)**, eQTLs **(B)**, and eGene **(C)** levels. A FDR-adjusted $p < 0.05$ was used to identify eQTLs and eGenes. **(D)** Psychiatric disorders and cognition share a

(Continued)

FIGURE 3 | large degree of commonality at the biological pathways level. Biological pathways containing eGenes for each phenotype were identified using iPathwayGuide (**Supplementary Spreadsheet 4**). Among the most impacted pathways, 61 were shared between psychiatric disorders and cognition, 66—across all five psychiatric disorders. Only one pathway (i.e., Pancreatic secretion) was unique to ADHD, three pathways [i.e., Vasopressin-regulated water reabsorption, Ovarian steroidogenesis, and Dilated cardiomyopathy (DCM)] were specific to anxiety, three were unique to BD (i.e., IL-17 signaling pathway, Antifolate resistance, and Cytosolic DNA-sensing pathway), 2—to SCZ (i.e., p53 signaling pathway and Prion diseases), 2—to UD (i.e., Rheumatoid arthritis and Malaria) and 1—to cognition (i.e., Carbohydrate digestion and absorption). The full list of shared and unique pathways is in **Supplementary Spreadsheet 5**.

FIGURE 4 | Psychiatric and cognitive SNPs mark eQTLs that are associated with gene expression within the neurotrophin signaling pathway. The co-occurrence of the affected shared or phenotype-specific eGenes and imbalance in gene expression within this pathway may lead to a series of cellular functions and events associated with psychiatric and cognitive phenotypes and the multimorbidities between them.

psychiatric and cognitive phenotypes in the brain (**Figure 5C**, **Supplementary Table 6**, **Supplementary Spreadsheet 8**). Whilst there were no brain-specific pathways that were shared across all phenotypes, the psychiatric disorders have between 6 and 24 shared pathways with the cognitive phenotypes (eTable 15 in Supplement 2). Again, ADHD was an exception and did not share any pathways with the cognitive phenotypes (**Figure 5C**). Bootstrapping ($n = 10,000$) confirmed the significance ($p < 0.01$) of the observed brain-specific eGene

and pathway overlaps, except for the anxiety+BD+SCZ overlap (**Supplementary Figure 14**).

DISCUSSION

There is increasing evidence that psychiatric disorders and cognitive functioning share an underlying cause, the so-called "p factor" (11). The existence of the p factor is further supported
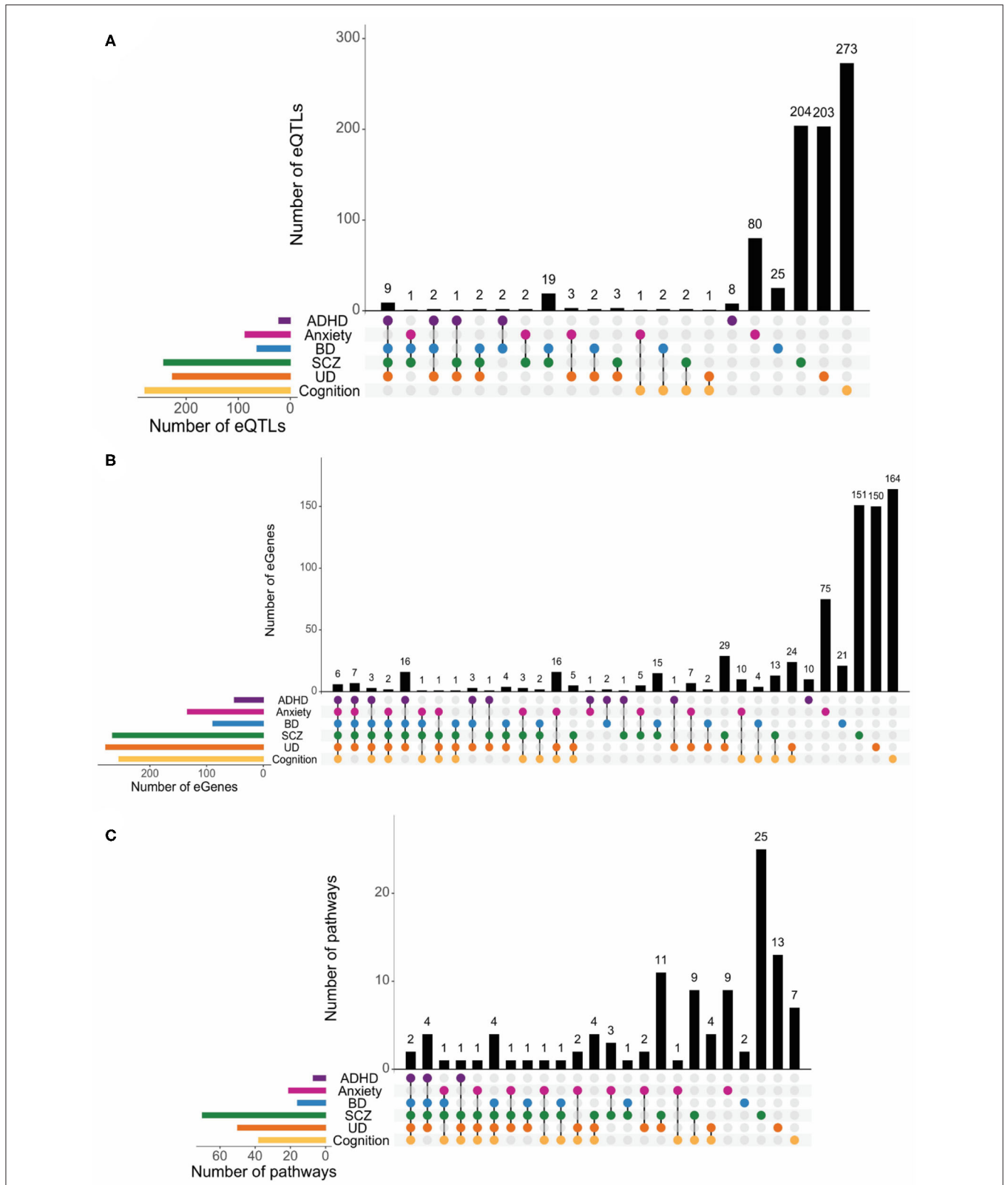
FIGURE 5 | Psychiatric disorders and cognition show no shared pathways among all phenotypes in brain tissues. Psychiatric disorders and cognitive functions have low levels of genetic similarity at the eQTLs **(A)** and eGene **(B)** levels in brain tissues. A FDR-adjusted $p < 0.05$ was used to identify eQTLs and eGenes. More pairwise brain-specific pathway overlaps **(C)** and individualized pathways were identified for psychiatric and cognitive phenotypes.

by the identification of genetic overlaps between psychiatric disorders (12). Despite the growing recognition of the p factor for psychiatric conditions, how the p factor manifests biologically remains unclear. We hypothesized that the p factor results from the action of genetic variants associated with psychiatric and cognitive phenotypes on tissue-specific expression of genes within shared biological pathways. Consistent with our hypothesis, we identified relatively little common genetic risk at the level of SNPs and eQTLs, but extensive biological pathway overlaps. We identified 61 biological pathways that were shared amongst the tested psychiatric and cognitive phenotypes.

Even though the relevance of the pathways unique to individual psychiatric or cognitive phenotypes may not be immediately obvious (e.g., pancreatic secretion and ADHD), there is increasing evidence that psychiatric disorders represent complex multisystem conditions (49–51). Moreover, it is obvious that the brain is a component within a much larger system and is subject to both direct and indirect contributions from other tissues that impact psychiatric and cognitive phenotypes. This is epitomized by hypo and hyperthyroidism, whereby alterations to thyroid hormone status are associated with development of psychiatric conditions (e.g., SCZ, BD, anxiety and depression) (44, 52). Notably, we observed that the thyroid hormone signaling pathway was impacted by genetic variants specific to all six phenotypes.

Infections during pregnancy, at birth and in early childhood increase the risk of ADHD (53), BD (54), and SCZ (54). We identified changes to gene regulation within pathways associated with infectious diseases (e.g., HTLV-I infection, and hepatitis B) that involved eQTLs for all phenotypes. Phenotype specific overlaps in immune related pathways were retained in the brain specific analysis. For example, gene expression within the natural killer cell mediated cytotoxicity and Th1 and Th2 cell differentiation pathways was affected by genetic variants associated with ADHD, BD, UD, SCZ, and cognition. We also identified regulatory changes to genes within signaling and neurodevelopmental pathways. Collectively, these results are consistent with changes to gene regulation within signaling, immune, and neurodevelopmental pathways combining to affect the pathophysiology and development of ADHD, anxiety, BD, UD, SCZ, and their association with cognitive functions. Of course, the convergence of phenotype specific genetic impacts on shared biological pathways does not explain all the features of the observed multimorbidities between the tested phenotypes. Clearly, there remains a role for dysregulation within pathways unique to individual psychiatric and cognitive phenotypes. Moreover, it is important to remember that pathways represent part of many cell-type and tissue-specific networks and thus can exert their effects on the multimorbidity risk through these networks. The impacts of cell-type and tissue-specific networks were not explicitly considered in this study.

We identified drugs that target gene products encoded by genes impacted by eQTL SNPs associated with each of the tested psychiatric and cognitive phenotypes. We contend that the therapeutic targeting of shared genes helps to explain the psychiatric side effects that are observed for these drugs. This hypothesis could be tested by categorizing patients who exhibit treatment associated side-effects according to the presence or absence of the allele associated with the side-effect relevant phenotype. However, genes in the shared set can only be used in patient stratification for personalized treatment if the eQTL effects occur in the same direction. Arguably, these decisions will only affect clinical outcomes if the variants are of high penetrance. However, the validity of these assumptions remains to be determined.

The development of psychiatric and cognitive phenotypes depends on a complex, often non-linear, dynamic interplay between genetic and environmental factors. As such, we are aware that our study has several limitations. Firstly, increasing the number of GWAS studies on certain phenotypes (e.g., ADHD, anxiety, etc.) will result in the identification of additional novel SNP loci. Secondly, common SNPs do not explain all of the estimated heritability in psychiatric and cognitive phenotypes, suggesting that other factors (e.g., rare variants, indels, structural variation etc …) also contribute (55). Thirdly, the GTEx eQTL data used in this study was largely from European individuals aged 40 years and older (56). Thus, robust predictions of the regulatory mechanisms of psychopathology and cognitive functioning across different populations and developmental stages requires additional data sets. Moreover, the reduced numbers of brain samples may have made the eQTLs and thus pathway overlaps more difficult to detect, thus accounting for the observed reduction in phenotype overlap. Fourthly, SNPs associated with diverse cognitive functions were combined into one general "cognitive functioning" phenotype, and do not represent common factors underlying all cognitive traits (**Supplementary Spreadsheet 1**). As such further research is needed to look more precisely at specific aspects of cognitive functioning (e.g., mathematical ability, language performance etc…) and their relationship with psychiatric disorders. Fifthly, it should be noted that the way we identify eGenes (restriction fragments do not have to contain the eGene promoter) may lead to potential eGenes over-identification. Sixthly, we are also aware that there is a potential bias in the tools, datasets, and databases used in this study, therefore our findings do not necessarily represent the full picture. For example, the usage of Hi-C datasets for HeLa cells may introduce unwanted bias due to their aberrant karyotype. Moreover, our study of tissue-specific regulatory interactions was potentially confounded by the tissue specific expression and Hi-C datasets not being paired. This is particularly relevant if there is an interaction with RNA degradation during sample preparation. Tissue-specific eQTL SNP effects mediating the eGene overlaps may not be independent and represent side effects of chromatin-level regulation. Incorporating other tissue-specific multi-omics data [e.g., from ENCODE consortium (57)] could provide a more comprehensive and accurate identification of tissue-specific regulatory changes associated with psychiatric and cognitive phenotypes. Finally, despite the statistical significance of the observed enrichment (FDR < 0.05), it can be argued that these results are unrealistically positive and that a stricter cut-off (FDR < 0.01) should have been employed to further reduce the chances of identifying false positives. Despite these limitations, our analysis provides a starting point for further mechanistic and

functional investigation. Replication analyses will increase the robustness of the results and provide a clearer indication of cross-phenotype overlap and phenotype-specific genetic architecture.

In conclusion, we have described biological pathways for multimorbidity and identified drug-gene interactions that may be clinically relevant for the treatment and prevention of psychiatric and cognitive phenotypes. Our results provide support for a shift from a gene-centric approach to identifying pathways for the treatment of multimorbid psychiatric and cognitive conditions. Future applications of the spatial genetic approach we used to other phenotypes will cross the molecular, cellular, tissue, and system levels to define personalized disease risk profiles, therapeutic targets and drug side-effects.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

EG performed the analyses and wrote the manuscript. MV co-supervised EG and commented on the manuscript. CE aided results interpretation and commented on the manuscript. JO'S supervised EG and co-wrote the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyt.2020.560751/full#supplementary-material

REFERENCES

1. James SL, Abate D, Abate KH, Abay SM, Abbafati C, Abbasi N, et al. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: a systematic analysis for the global burden of disease study 2017. *Lancet.* (2018) 392:1789–858. doi: 10.1016/S0140-6736(18)32279-7

2. Katzman MA, Bilkey TS, Chokka PR, Fallu A, Klassen LJ. Adult ADHD and comorbid disorders: clinical implications of a dimensional approach. *BMC Psychiatry.* (2017) 17:302. doi: 10.1186/s12888-017-1463-3

3. Essau CA, Lewinsohn PM, Lim JX, Ho M-HR, Rohde P. Incidence, recurrence and comorbidity of anxiety disorders in four major developmental stages. *J Affect Disord.* (2018) 228:248–53. doi: 10.1016/j.jad.2017.12.014

4. Buckley PF, Miller BJ, Lehrer DS, Castle DJ. Psychiatric comorbidities and schizophrenia. *Schizophr Bull.* (2009) 35:383–402. doi: 10.1093/schbul/sbn135

5. Kessler RC, Gruber M, Hettema JM, Hwang I, Sampson N, Yonkers KA. Co-morbid major depression and generalized anxiety disorders in the national comorbidity survey follow-up. *Psychol Med.* (2008) 38:365–74. doi: 10.1017/S0033291707002012

6. Brainstorm Consortium, Anttila V, Bulik-Sullivan B, Finucane HK, Walters RK, Bras J, et al. Analysis of shared heritability in common disorders of the brain. *Science.* (2018) 360:eaap8757. doi: 10.1126/science.aap8757

7. Smeland OB, Bahrami S, Frei O, Shadrin A, O'Connell K, Savage J, et al. Genome-wide analysis reveals extensive genetic overlap between schizophrenia, bipolar disorder, and intelligence. *Mol Psychiatry.* (2019) 25:914. doi: 10.1038/s41380-019-0456-7

8. Cross-Disorder Group of the Psychiatric Genomics Consortium. Genomic relationships, novel loci, and pleiotropic mechanisms across eight psychiatric disorders. *Cell.* (2019) 179:1469–82.e11. doi: 10.1016/j.cell.2019.11.020

9. Xia L, Xia K, Weinberger DR, Zhang F. Common genetic variants shared among five major psychiatric disorders: a large-scale genome-wide combined analysis. *Glob Clin Transl Res.* (2019) 1:21–30. doi: 10.36316/gcatr.01.0003

10. Schork AJ, Won H, Appadurai V, Nudel R, Gandal M, Delaneau O, et al. A genome-wide association study of shared risk across psychiatric disorders implicates gene regulation during fetal neurodevelopment. *Nat Neurosci.* (2019) 22:353–61. doi: 10.1038/s41593-018-0320-0

11. Caspi A, Houts RM, Belsky DW, Goldman-Mellor SJ, Harrington H, Israel S, et al. The p factor: one general psychopathology factor in the structure of psychiatric disorders? *Clin Psychol Sci.* (2014) 2:119–37. doi: 10.1177/2167702613497473

12. Selzam S, Coleman JRI, Caspi A, Moffitt TE, Plomin R. A polygenic p factor for major psychiatric disorders. *Transl Psychiatry.* (2018) 8:205. doi: 10.1038/s41398-018-0217-4

13. Maurano MT, Humbert R, Rynes E, Thurman RE, Haugen E, Wang H, et al. Systematic localization of common disease-associate variation in regulatorty DNA. *Science.* (2012) 337:1190–5. doi: 10.1126/science.1222794

14. van Arensbergen J, Pagie L, FitzPatrick VD, de Haas M, Baltissen MP, Comoglio F, et al. High-throughput identification of human SNPs affecting

regulatory element activity. *Nat Genet.* (2019) 51:1160–9. doi: 10.1101/460402

15. Tam V, Patel N, Turcotte M, Bossé Y, Paré G, Meyre D. Benefits and limitations of genome-wide association studies. *Nat Rev Genet.* (2019) 20:467–84. doi: 10.1038/s41576-019-0127-1

16. Marbach D, Lamparter D, Quon G, Kellis M, Kutalik Z, Bergmann S. Tissue-specific regulatory circuits reveal variable modular perturbations across complex diseases. *Nat Methods.* (2016) 13:366–70. doi: 10.1038/nmeth.3799

17. Sonawane AR, Platig J, Fagny M, Chen C-Y, Paulson JN, Lopes-Ramos CM, et al. Understanding Tissue-Specific Gene Regulation. *Cell Rep.* (2017) 21:1077–88. doi: 10.1016/j.celrep.2017.10.001

18. Nott A, Holtman IR, Coufal NG, Schlachetzki JCM, Yu M, Hu R, et al. Brain cell type-specific enhancer-promoter interactome maps and disease-risk association. *Science.* (2019) 366:1134–9. doi: 10.1126/science.aay0793

19. Schmitt AD, Hu M, Jung I, Xu Z, Qiu Y, Tan CL, et al. A compendium of chromatin contact maps reveals spatially active regions in the human genome. *Cell Rep.* (2016) 17:2042–59. doi: 10.1016/j.celrep.2016.10.061

20. Huo Y, Li S, Liu J, Li X, Luo X-J. Functional genomics reveal gene regulatory mechanisms underlying schizophrenia risk. *Nat Commun.* (2019) 10:670. doi: 10.1038/s41467-019-08666-4

21. Li M, Jaffe AE, Straub RE, Tao R, Shin JH, Wang Y, et al. A human-specific AS3MT isoform and BORCS7 are molecular risk factors in the 10q24.32 schizophrenia-associated locus. *Nat Med.* (2016) 22:649–56. doi: 10.1038/nm.4096

22. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* (2010) 38:e164. doi: 10.1093/nar/gkq603

23. Chang X, Wang K. wANNOVAR: annotating genetic variants for personal genomes via the web. *J Med Genet.* (2012) 49:433–6. doi: 10.1136/jmedgenet-2012-100918

24. Rao SSP, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, et al. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell.* (2014) 159:1665–80. doi: 10.1016/j.cell.2014.11.021

25. Won H, de la Torre-Ubieta L, Stein JL, Parikshak NN, Huang J, Opland CK, et al. Chromosome conformation elucidates regulatory relationships in developing human brain. *Nature.* (2016) 538:523–7. doi: 10.1038/nature19847

26. Dixon JR, Jung I, Selvaraj S, Shen Y, Antosiewicz-Bourget JE, Lee AY, et al. Chromatin architecture reorganization during stem cell differentiation. *Nature.* (2015) 518:331–6. doi: 10.1038/nature14222

27. Durand NC, Shamim MS, Machol I, Rao SSP, Huntley MH, Lander ES, et al. Juicer provides a one-click system for analyzing loop-resolution hi-c experiments. *Cell Systems.* (2016) 3:95–8. doi: 10.1016/j.cels.2016.07.002

28. Fadason T, Ekblad C, Ingram JR, Schierding WS, O'Sullivan JM. Physical interactions and expression quantitative traits loci identify regulatory connections for obesity and type 2 diabetes associated SNPs. *Front Genet.* (2017) 8:150. doi: 10.3389/fgene.2017.00150

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

30. Reimand J, Arak T, Adler P, Kolberg L, Reisberg S, Peterson H, et al. g:Profiler-a web server for functional interpretation of gene lists (2016 update). *Nucleic Acids Res.* (2016) 44:W83–9. doi: 10.1093/nar/gkw199

31. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* (2000) 28:27–30. doi: 10.1093/nar/28.1.27

32. Cotto KC, Wagner AH, Feng Y-Y, Kiwala S, Coffman AC, Spies G, et al. DGIdb 3.0: a redesign and expansion of the drug-gene interaction database. *Nucleic Acids Res.* (2018) 46:D1068–D73. doi: 10.1093/nar/gkx1143

33. Cross-Disorder Group of the Psychiatric Genomics Consortium, Lee SH, Ripke S, Neale BM, Faraone SV, Purcell SM, et al. Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. *Nat Genet.* (2013) 45:984–94. doi: 10.1038/ng.2711

34. Cross-Disorder Group of the Psychiatric Genomics Consortium. Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis.

35. Ohi K, Shimada T, Nitta Y, Kihara H, Okubo H, Uehara T, et al. Schizophrenia risk variants in ITIH4 and CALN1 regulate gene expression in the dorsolateral prefrontal cortex. *Psychiatr Genet.* (2016) 26:142–3. doi: 10.1097/YPG.0000000000000127

36. Liu J, Chen J, Perrone-Bizzozero NI, Turner JA, Calhoun VD. Regional enrichment analyses on genetic profiles for schizophrenia and bipolar disorder. *Schizophr Res.* (2018) 192:240–6. doi: 10.1016/j.schres.2017.04.033

37. Zhou Y, Cao Z, Yang M, Xi X, Guo Y, Fang M, et al. Comorbid generalized anxiety disorder and its association with quality of life in patients with major depressive disorder. *Sci Rep.* (2017) 7:40511. doi: 10.1038/srep40511

38. Hettema JM. What is the genetic relationship between anxiety and depression? *Am J Med Genet C Semin Med Genet.* (2008) 148C:140–6. doi: 10.1002/ajmg.c.30171

39. Davies MR, Kalsi G, Armour C, Jones IR, McIntosh AM, Smith DJ, et al. The genetic links to anxiety and depression (GLAD) study: online recruitment into the largest recontactable study of depression and anxiety. *Behav Res Ther.* (2019) 123:103503. doi: 10.1016/j.brat.2019.103503

40. Mitre M, Mariga A, Chao MV. Neurotrophin signalling: novel insights into mechanisms and pathophysiology. *Clin Sci.* (2017) 131:13–23. doi: 10.1042/CS20160044

41. Kahl KG, Westhoff-Bleck M, Krüger THC. Effects of psychopharmacological treatment with antipsychotic drugs on the vascular system. *Vascul Pharmacol.* (2018) 100:20–5. doi: 10.1016/j.vph.2017.09.001

42. Ronan JL, Wu W, Crabtree GR. From neural development to cognition: unexpected roles for chromatin. *Nat Rev Genet.* (2013) 14:347–59. doi: 10.1038/nrg3413

43. Russo E, Leo A, Crupi R, Aiello R, Lippiello P, Spiga R, et al. Everolimus improves memory and learning while worsening depressive- and anxiety-like behavior in an animal model of depression. *J Psychiatr Res.* (2016) 78:1–10. doi: 10.1016/j.jpsychires.2016.03.008

44. Noda M. Possible role of glial cells in the relationship between thyroid dysfunction and mental disorders. *Front Cell Neurosci.* (2015) 9:194. doi: 10.3389/fncel.2015.00194

45. Bauer ME, Teixeira AL. Inflammation in psychiatric disorders: what comes first? *Ann N Y Acad Sci.* (2019) 1437:57–67. doi: 10.1111/nyas.13712

46. Savitz J, Harrison NA. Interoception and inflammation in psychiatric disorders. *Biol Psychiatry Cogn Neurosci Neuroimaging.* (2018) 3:514–24. doi: 10.1016/j.bpsc.2017.12.011

47. De Hert M, Detraux J, Vancampfort D. The intriguing relationship between coronary heart disease and mental disorders. *Dialogues Clin Neurosci.* (2018) 20:31–40.

48. Phillips JR, Hewedi DH, Eissa AM, Moustafa AA. The cerebellum and psychiatric disorders. *Front Public Health.* (2015) 3:66. doi: 10.3389/fpubh.2015.00066

49. Leboyer M, Soreca I, Scott J, Frye M, Henry C, Tamouza R, et al. Can bipolar disorder be viewed as a multi-system inflammatory disease? *J Affect Disord.* (2012) 141:1–10. doi: 10.1016/j.jad.2011.12.049

50. Pillinger T, D'Ambrosio E, McCutcheon R, Howes OD. Is psychosis a multisystem disorder? A meta-review of central nervous system, immune, cardiometabolic, and endocrine alterations in first-episode psychosis and perspective on potential models. *Mol Psychiatry.* (2019) 24:776–94. doi: 10.1038/s41380-018-0058-9

51. de Lange SC, Scholtens LH, Alzheimer's Disease Neuroimaging Initiative, van den Berg LH, Boks MP, Bozzali M, et al. Shared vulnerability for connectome alterations across psychiatric and neurological brain disorders. *Nat Hum Behav.* (2019) 3:988–98. doi: 10.1101/360586

52. Ghassabian A, Trasande L. Disruption in thyroid signaling pathway: a mechanism for the effect of endocrine-disrupting chemicals on child neurodevelopment. *Front Endocrinol.* (2018) 9:204. doi: 10.3389/fendo.2018.00204

53. Millichap JG. Etiologic classification of attention-deficit/hyperactivity disorder. *Pediatrics.* (2008) 121:e358–65. doi: 10.1542/peds.2007-1332

54. Brown AS. The kraepelinian dichotomy from the perspective of prenatal infectious and immunologic insults. *Schizophr Bull.* (2015) 41:786–91. doi: 10.1093/schbul/sbv063

Lancet. (2013) 381:1371–9. doi: 10.1016/S0140-6736(12)62129-1

55. Geschwind DH, Flint J. Genetics and genomics of psychiatric disease. *Science.* (2015) 349:1489–94. doi: 10.1126/science.aaa8954

56. Gamazon ER, Segrè AV, van de Bunt M, Wen X, Xi HS, Hormozdiari F, et al. Using an atlas of gene regulation across 44 human tissues to inform complex disease- and trait-associated variation. *Nat Genet.* (2018) 50:956–67. doi: 10.1038/s41588-018-0154-4

57. Davis CA, Hitz BC, Sloan CA, Chan ET, Davidson JM, Gabdank I, et al. The encyclopedia of DNA elements (ENCODE): data portal update. *Nucleic Acids Res.* (2018) 46:D794–801. doi: 10.1093/nar/gkx1081

58. Golovina E, Vickers MH, Erb CD, O'Sullivan JM. Decoding the multimorbidities among psychiatric disorders and cognitive functioning. *bioRxiv [Preprint].* (2019)