Dissecting Schizophrenia Biology Using Pleiotropy with Cognitive Genomics

Upasana Bhattacharyya*[1,2], Jibin John*[1,2], Todd Lencz[§1,2,3], Max Lam[§1,4,5]

[1]Institute of Behavioral Science, Feinstein Institutes for Medical Research, Manhasset, NY

[2]Division of Psychiatry Research, Zucker Hillside Hospital, Glen Oaks, NY

[3]Departments of Psychiatry and Molecular Medicine, Zucker School of Medicine at Hofstra/Northwell, Hempstead, NY

[4]Institute of Mental Health, Singapore

[5]Lee Kong Chian School of Medicine, Population and Global Health, Nanyang Technological University

*These authors contributed equally

[§]These authors contributed equally

Corresponding Authors: tlencz@northwell.edu; max.lam@ntu.edu.sg

Abstract

Given the increasingly large number of loci discovered by psychiatric GWAS, specification of the key biological pathways underlying these loci has become a priority for the field. We have previously leveraged the pleiotropic genetic relationships between schizophrenia and two cognitive phenotypes (educational attainment and cognitive task performance) to differentiate two subsets of illness-relevant SNPs: (1) those with "concordant" alleles, which are associated with reduced cognitive ability/education and increased schizophrenia risk; and (2) those with "discordant" alleles linked to reduced educational and/or cognitive levels but lower schizophrenia susceptibility. In the present study, we extend our prior work, utilizing larger input GWAS datasets and a more powerful statistical approach to pleiotropic meta-analysis, the Pleiotropic Locus Exploration and Interpretation using Optimal test (PLEIO). Our pleiotropic meta-analysis of schizophrenia and the two cognitive phenotypes revealed 768 significant loci (159 novel). Among these, 347 loci harbored concordant SNPs, 270 encompassed discordant SNPs, and 151 “dual” loci contained concordant and discordant SNPs. Competitive gene-set analysis using MAGMA related concordant SNP loci with neurodevelopmental pathways (e.g., neurogenesis), whereas discordant loci were associated with mature neuronal synaptic functions. These distinctions were also observed in BrainSpan analysis of temporal enrichment patterns across developmental periods, with concordant loci containing more prenatally expressed genes than discordant loci. Dual loci were enriched for genes related to mRNA translation initiation, representing a novel finding in the schizophrenia literature.

Keywords: schizophrenia, cognition, educational attainment, GWAS, pleiotropy

Schizophrenia is a complex neuropsychiatric disorder affecting approximately 1% of the global population, contributing significantly to the worldwide burden of disease[1]. Over the past decade, genome-wide association studies (GWAS) have identified many genomic loci associated with schizophrenia, revealing its heterogeneous and multifaceted nature[2–6]. However, gleaning biological insights from GWAS results has proven challenging, with downstream analyses pointing broadly toward neuronal and synaptic mechanisms without clear mechanistic differentiation[2,4,7]. Pleiotropic analyses, combining GWAS of schizophrenia with other genetically correlated forms of psychopathology, have boosted the power of gene set enrichment analyses; such studies have implicated a range of neurodevelopmental, synaptic, and other molecular pathways[8,9] but have not adequately parsed the association of individual pathways with specific phenotypic features.

In two recent studies[10,11], we have demonstrated that pleiotropic analysis of schizophrenia GWAS with cognitive endophenotypes may permit well-powered, fine-grained, mechanistic differentiation of molecular pathways associated with specific phenotypic constellations. Impaired cognitive ability is a core dimension of schizophrenia pathology[12]; cognitive deficits often emerge early, frequently preceding schizophrenia's onset by years and contributing to diminished functional outcomes, including educational attainment[13]. Moreover, recent GWAS of cognition and educational attainment demonstrate significant overlap with schizophrenia risk loci[10,14–17]. However, genetic correlation analyses reveal a nuanced relationship between cognitive function, educational attainment, and schizophrenia risk. While there is a significant negative genetic correlation between reduced cognitive function and increased schizophrenia susceptibility ($r_g \approx -0.20$), there exists a paradoxical, positive genetic correlation between greater educational attainment and schizophrenia risk ($r_g \approx 0.10$), despite a strong genetic overlap between educational attainment and cognitive ability ($r_g \approx 0.70$).

Thus, while educational attainment is often viewed as a proxy for cognitive ability, the imperfect genetic concordance between the two can be exploited in pleiotropic analyses of schizophrenia. In our previous work[10,11], we identified two distinct subsets of schizophrenia-related genetic loci: those with "concordant" alleles that align with expectations—lower cognition/education and increased illness risk—and "discordant" SNPs that show higher educational attainment and/or cognitive ability alongside greater schizophrenia susceptibility. This distinction is obscured when pleiotropy is examined globally, using only genomewide correlational analysis (r_g). Downstream analyses of concordant loci implicated neurodevelopmental pathways in schizophrenia, while discordant loci were associated with postnatally expressed synaptic mechanisms underlying the disorder.

The present study was designed to replicate and extend our prior pleiotropic meta-analysis, utilizing substantially larger input GWAS for schizophrenia, cognitive task performance, and educational attainment than were available in our prior report[10]. Moreover, in the current investigation, we applied a novel, more powerful meta-analytic approach to pleiotropy, the 'Pleiotropic Locus Exploration and Interpretation using Optimal test' (PLEIO)[18]. This methodology permits a direct assessment of the directional associations among these phenotypes. Based on our prior report, we anticipate the identification of novel loci previously unreported in discovery GWAS datasets. Additionally, post-GWAS annotation of these pleiotropically identified loci is expected to extend beyond neurodevelopmental and synaptic pathways, offering deeper insights into the biological mechanisms underpinning these associations (**Figure 1**).

Building on previous findings, our current study integrates the latest GWAS on schizophrenia (N= 53,386 cases and 77,258 controls; GWAS mean χ^2 =2.0126; total SNPs = 7,659,767)[2], and cognitive function (N=373,617; GWAS mean χ^2 =2.11; total SNPs = 8,050,310)[11] to deepen our understanding of its genetic underpinnings. We chose to utilize the educational attainment data set from Lee et al., 2018 (N=766,345; mean GWAS χ^2=2.647; total SNPs=10,101,242)[19], rather than the most recent study of >3M individuals[20], which could potentially introduce bias into our pleiotropy analyses due to imbalance in sample size and power relative to the schizophrenia and cognitive GWASs. All the GWASs used here are generated from populations of European ancestry.

In the current study, we applied rigorous quality control (QC) to each set of summary statistics utilizing the MungeSumstat tool[21], which we have optimized as a Python wrapper (accessible via GitHub). We employed the default QC parameters of MungeSumstats[21], which include removing non-biallelic SNPs and strand-ambiguous SNPs, ensuring consistent direction of reported SNP effects, and ensuring that all variants have non-zero effect sizes and/or standard errors. These QC procedures resulted in the retention of 6,803,445 SNPs common to all three phenotypes, with INFO scores greater than 0.3 and minor allele frequencies above 0.01 (See Supplementary Information).

Two core considerations guided our approach to pleiotropic meta-analysis across the three GWAS summary statistics: (i) We required a method that can account for cryptic sample overlap across GWAS to prevent statistical inflation during meta-analytic procedures; and (ii) we required an approach that could account for SNPs with allelic effects of opposite signs across phenotypes, as our primary aim was to identify “Discordant” as well as “Concordant” SNPs. To our knowledge, only two methods exist

that adequately support study objectives: ASSET (Association analysis based on subsets)[22] and PLEIO (Pleiotropic Locus Exploration and Interpretation using Optimal test)[18]. PLEIO overcomes limitations associated with fixed-effect model approaches by employing a random-effect model to capture genetic correlations and heritabilities across trait pairs[18]. Relative to ASSET, PLEIO is more computationally efficient and provides greater power in simulations across a range of different heritabilities[18]. Linkage disequilibrium score regression (LDSC) implemented within PLEIO estimates genetic covariance and environmental correlation matrices across traits to account for potential sample overlap.

Consistent with our earlier study, we reversed the effect direction per variant for schizophrenia to align with the interpretation of allelic effects on cognition and education[10]; in other words, decreased risk for schizophrenia was coded by positive allelic effects, just as higher scores on cognitive task performance and educational attainment are coded with positive allelic effects. By applying PLEIO, we identified two subsets of variants, concordant and discordant, defined as in our prior study[8]. In the notation that follows, the symbol ∩ indicates variant subsets with the same effect direction across the traits, and the symbol | represents traits whose effect sizes exhibit the opposite direction compared to the other two traits. Thus, the concordant subset includes variants with the same effect direction (i.e., scz ∩ edu ∩ cog, after reversal of sign for schizophrenia). By contrast, the discordant subset includes three distinct sets of variants: (i) edu ∩ cog | scz (schizophrenia outliers, i.e., same allele is associated with an increase in cognitive ability, educational attainment, and schizophrenia risk); (ii) scz ∩ cog | edu (education outliers, i.e., variants associated with an increase in schizophrenia risk and reduced cognitive ability, but increased educational attainment); and (iii) scz ∩ edu | cog (cognition outliers, i.e., variants associated with an increase in schizophrenia risk and reduced educational attainment but

increased cognitive ability). Independent genome-wide significant loci ($p<5x10^{-8}$) were identified using the Functional Mapping and Annotation (FUMA) pipeline[23] (See Data Availability). We excluded the MHC region (chr6:25000000-3500000) from further analysis due to its complex LD patterns, opting to retain a single top SNP from this region.

After harmonization in FUMA (see supplementary methods), we identified three distinct sets of genome-wide significant loci (**Figure 2**): (i) 347 independent loci with exclusively concordant genome-wide significant SNPs (Supplementary Table 1); (ii) 270 independent genomic loci in which all genome-wide significant loci were discordant (Supplementary Table 2); and (iii) 151 "dual" loci, containing both concordant and discordant variants (Supplementary Methods; Supplementary Table 3). The 'dual' variant set is novel, ostensibly due to the increased power of the larger input GWAS sample sizes. Next, we compared these results with our prior report[8], where we reported only Concordant (89 loci) and Discordant (65 loci) with earlier and more modestly sampled versions of the input GWAS.

To identify novel loci emerging from our PLEIO meta-analysis, we merged the loci reported for schizophrenia, cognitive task performance, and educational attainment in the original input GWAS publications, along with loci obtained from our pleiotropic meta-analysis using the "merge" option in 'bedtools'[24]. (We first employed FUMA to map significant variants to loci for education since the original publication lacked comparable locus information.) This process yielded 739 genome-wide significant ($p < 5x10^{-8}$) loci emerging from the pleiotropic meta-analysis, as well as 280 loci for schizophrenia, 432 for education, and 292 for cognitive task performance; these numbers differ slightly from the original reports due to the merging of partially overlapping loci in bedtools (see

Supplementary Methods). A Venn diagram of these results is presented in **Figure 3**, revealing that 166 PLEIO loci are not being identified in any of the three input GWASs. By contrast, several loci remained specific to the input GWASs and did not reach genome-wide significance in the pleiotropic meta-analysis (schizophrenia: 94; education: 46; cognitive task performance: 52), presumably indicating that these loci were specific to one phenotype and not pleiotropic.

Comparing results with our earlier pleiotropic meta-analytic study[8], we identified 110 novel loci. Notably, 89 of these loci (comprising 80.9%) were validated as significant in subsequent, larger GWAS datasets for the analyzed traits, underscoring the efficacy of pleiotropic analysis in uncovering novel genetic loci[2,11,19]. In the present report, we employed the most recent GWAS data for schizophrenia and cognitive task performance, leaving us without newer studies for direct comparisons. Nevertheless, we can cross-reference a more recent study on educational attainment[20] with sample sizes nearly three times larger than our input GWAS[19]. Because the complete summary statistics of the recent education GWAS are not available in the public domain, we have used the restricted summary statistics for the complete additive autosomal GWAS that contains clumped results for 8,618 variants with $P < 1x10^{-5}$ (Sample size = 3,037,499). To conduct the look-up, we checked if any loci identified from our analysis overlap with 20kb upstream or downstream of any reported variants. Among the 166 novel loci identified, 111 (66.86%) were confirmed in the latest educational GWAS, validating the PLEIO methodology's effectiveness and highlighting its increased capacity for discovering novel genetic loci across phenotypes (Supplementary Table 4). It is important to note that we exclusively considered variants achieving genome-wide significance (adjusted $P < 1.28x10^{-8}$) to ensure compatibility with our dataset.

MAGMA (version 1.10) (Multi-marker Analysis of GenoMic Annotation)[25] gene set analysis was conducted to identify specific biological processes underpinning Concordant, Discordant and 'Dual' variant sets. Variants were annotated to gene regions, setting boundaries 35 kb upstream and 10 kb downstream to include likely regulatory regions associated with each gene, as suggested by previous publications[4,26–28]. Following annotation, gene-based genome-wide association (MAGMA-GBGWA) analysis was conducted using the SNP-wise mean method to compute the p-value for each gene. The gene-based results were subsequently entered into the MAGMA competitive gene set analysis. Gene sets from MsigDb C5 collection (GO-basic obo file released on 2023-07-27) that contain all the Gene Ontology gene sets[29] (GO Biological Process ontology, GO Cellular Component ontology, GO Molecular Function ontology) were utilized as annotations for gene set analysis. We excluded genes within chromosome X or Y or extended MHC region. Any gene set that contained less than 10 genes was also excluded (See Supplementary Methods). Competitive gene set analyses were performed separately for Concordant, Discordant, and Dual variants against 7,353 gene sets (See Data Availability).

MAGMA competitive gene set analysis for the concordant loci revealed 30 GO annotations enriched at FDR<0.05 (Supplementary Table 5). Results were broadly consistent with our previous observation that concordant loci were enriched for transcriptional regulation pathways and neurodevelopmental genes[8]. With the current study's enhanced power, we could identify multiple RNA regulatory pathways implicated at concordant loci beyond the chromatin regulation discussed in our prior report. These include a broader range of gene sets involved with chromosome organization,

chromatin formation, and methylation. Additionally, we identified significant gene sets regulating the spliceosome, which has emerged as a critical focus of abnormalities in neuropsychiatric disorders[30].

The present results also provided additional details on the nature of neurodevelopmental processes implicated by concordant genes, including macro-level processes such as head development as well as cellular processes of neurogenesis. Notably, two FDR-significant gene sets specifically implicated forebrain development. As the forebrain develops into the cerebral cortex, neurodevelopmental models of schizophrenia have long focused on this process[31]. Converging lines of evidence demonstrate that illness-related cognitive impairment in schizophrenia is neurodevelopmental in origin and long predates overt symptomatology[32,33]. Forebrain developmental abnormalities may provide a genetic basis for the increased rate of craniofacial abnormalities in patients with schizophrenia, which are sometimes also observed in their first-degree relatives[34]. Indeed, it has been suggested that refined measurement of craniofacial dysmorphology may represent a readily accessible index of neurodevelopmental abnormalities in schizophrenia[35,36]. Additionally, very recent studies have recapitulated gene expression abnormalities in schizophrenia using forebrain organoids derived from induced pluripotent stem cells of patients with schizophrenia[34].

MAGMA competitive gene set analysis at discordant loci revealed 19 GO annotations enriched at FDR<0.05 (Supplementary Table 5). As in our prior study[8], the discordant subset demonstrated significant enrichment in synaptic pathways, especially the postsynaptic density, which has emerged as perhaps the strongest gene set emerging from schizophrenia GWAS[2]. The present results suggest that these pathways are distinct from those causing the cognitive deficits related to schizophrenia and are

instead associated with the counter-intuitive positive genetic correlation of schizophrenia with educational attainment. As noted in our prior report[8], these results do not point towards a subset of schizophrenia patients marked by high educational attainment but rather may indicate an inverted-U function: post-synaptic processes that underlie successful academic performance may also lead to illness if taken too far.

With the enhanced power of the present study, we identified several novel gene sets enriched in the discordant loci, most notably several involved in neurodevelopment specific to the hindbrain/cerebellum. While hindbrain development has not been a primary focus in schizophrenia research, abnormalities in the mature cerebellum have been identified using both structural and functional MRI[37]. Traditionally linked to motor control, the cerebellum is increasingly recognized for its involvement in cognitive and emotional functions[38]. Identification of cerebellar abnormalities, in conjunction with synaptic deficits, suggests a previously under-appreciated role for metabotropic[39] and delta ionotropic[40] glutamate receptors in the illness process.

Several additional gene sets were enriched in discordant loci, pointing towards novel targets for future study. The regulation of the Notch signaling pathway plays a pivotal role in governing cell fate decisions during neurodevelopment but is increasingly understood as subserving learning processes in mature neurons[41]. The unfolded protein response pathway maintains homeostasis within the endoplasmic reticulum (ER), in response to ER stress[42]; molecular evidence of dysfunction in this system has recently been observed in postmortem brains of patients with schizophrenia[43].

In gene set enrichment analysis for "dual" loci, pathways related to translation initiation were the primary significant result (Supplementary Table 5). Intriguingly, one of the genes driving this association was *EIF3C*, which neighbors (and is often included in) the schizophrenia-associated copy number variant at distal 16p11.2[44]. The translation initiation pathway is fundamental to many biological processes, as the effective operation of diverse cell types hinges on the precise and dynamic regulation of mRNA translation to proteins. Neurons, characterized by their highly polarized morphology, particularly depend on the spatial organization of mRNA translation machinery[45]. Moreover, temporal control of mRNA translation is pivotal for neurons responding promptly to environmental changes by modifying synaptic protein composition; this process is essential to experience-dependent long-term synaptic plasticity[46]. A recent study directly investigated the translational control of protein synthesis in olfactory neurosphere-derived (ONS) cells drawn from patients with schizophrenia and controls. The study found 48 differentially expressed eIF2 pathway proteins and mRNA transcripts associated with schizophrenia; additionally, mTOR and eIF4 signaling pathways, known regulators of protein synthesis, were implicated[47]. Thus, our finding that translation-related gene sets are enriched at the "dual" loci highlights the need for more studies to understand the role of protein translation in schizophrenia.

Finally, we utilized the publicly accessible BrainSpan dataset[48] to analyze temporal gene expression patterns concerning the three sets of pleiotropic loci. We hypothesized that genes at concordant loci, primarily involved in neurodevelopment, would peak in expression prenatally. In contrast, genes linked to adult synaptic regulation (discordant genes) would be more active during later developmental stages. We transformed neurodevelopmental stages into weeks to uncover more intricate gene expression profiles over time. As hypothesized, we found substantial time by locus-type

interaction effects among gene sets, with discordant genes demonstrating an upward slope over time relative to concordant genes (Concordant vs Discordant, b = 1.92×10^{-4}, s.e. 4.79×10^{-5}, p = 6.41×10^{-5}) (**Figure 4**). Somewhat surprisingly, genes at the dual loci demonstrated an even greater downward slope relative even to concordant genes (Concordant vs Dual, b = 2.45×10^{-4}, s.e. = 4.76×10^{-4}, p = 3.14×10^{-7}; Discordant v Dual, b = 4.37×10^{-4}, s.e. = 5.47×10^{-5}, p = 3.63×10^{-15}). However, the main effect for developmental weeks was not statistically significant for any of the three sets of genes individually (Concordant, b = -6.49×10^{-5}, s.e. = 2.5×10^{-4} p = 0.796; Discordant, b = 0.000138, s.e. = 0.000243, p = 0.572; Dual, b = -0.000375, s.e.= 0.000238, p = 0.124).

In conclusion, our investigation has yielded several key findings that replicate and extend our prior report[8] on the pleiotropic relationship between schizophrenia, cognitive task performance, and educational attainment. First, applying pleiotropic analysis enabled us to harness the collective power of multiple GWAS datasets for enhanced locus discovery. Second, our findings reaffirm the divergence between forebrain neurodevelopmental pathways and synaptic regulation in schizophrenia. While these pathways have previously emerged as the primary results from gene set analysis of common[4] and rare[49] genetic variation in schizophrenia, our work indicates that these can be parsed with respect to their endophenotypic manifestations. The larger sample sizes in the current report permit us to expand and refine these pathways, newly demonstrating roles for altered gene splicing at concordant loci and notch signaling at discordant loci. Third, identifying newly emergent gene set categories related to cerebellar development at discordant loci and translational initiation factors at dual loci provides clues to schizophrenia pathophysiology and potential treatment targets not previously reported in genetics literature.

Our pleiotropic integration of psychiatric GWAS with endophenotypic GWAS was designed to parse biological mechanisms underlying schizophrenia. While our primary results contribute to our understanding of the etiopathophysiology of the disorder, pointing towards novel treatment targets, we propose that the resulting data may, in the future, have potential clinical applicability in the form of more refined polygenic risk scores (PRS) reflecting these distinct biological pathways. Such an approach has been successfully developed in recent studies of Type 2 diabetes (T2D), in which the genetic heterogeneity of that complex disease has been parsed into several dissociable biological processes (e.g., lipodystrophy, insulin synthesis, hepatic lipid metabolism, etc.)[50]. The model, as proposed by Udler[51] and McCarthy[52], is not that there are multiple distinct subtypes of T2D, but rather that the observed diagnostic entity is comprised of several underlying dimensions and that any individual diagnosed with T2D has some combination of risk along these dimensions. Each dimension of risk is quantifiable by subsets of SNPs drawn from the overall PRS for T2D, corresponding to genetic regions relevant to the underlying biological processes, and a patient's PRS profile along these various dimensions can more accurately differentiate outcomes in T2D[50]. Recent endophenotype studies in schizophrenia have been guided by a categorical approach to the identification of 'biotypes,' such as a subtype of psychosis marked by cognitive deficits[53] and distinct brain dysmorphism[54]. While the present results are consistent with the evidence of such endophenotypic manifestations in some individuals with schizophrenia, our approach suggests that these represent dimensions of abnormality rather than distinct subgroups of patients.

Figures

Figure 1. Study strategy and workflow
Methods underlying the present report are broken down into (i) GWAS summary statistics QC phase using 'mungesumstats,' (ii) PLEIO pleiotropy meta-analysis, (iii) FUMA GWAS annotation and (iv) MAGMA gene set analysis and further downstream analytic procedures.

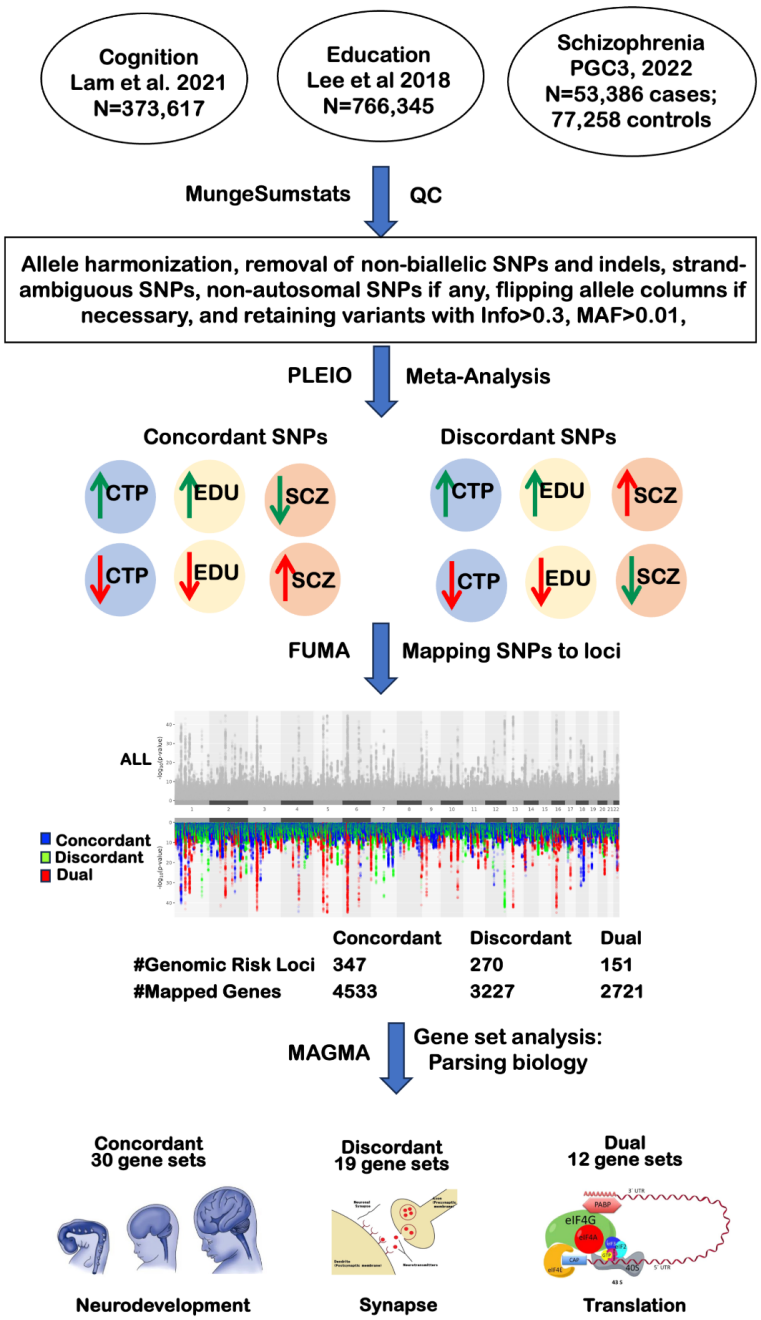


Figure 2: Miami plot of PLEIO meta-analysis results
Top panel: visualization of variant subsets derived from –blup recalibrated effect size directions: (i) Concordant variants (red) (ii) Discordant variants (green) and (iii) Dual variants (blue). Bottom panel: Complete pleiotropic meta-analysis results from the PLEIO tool, NPR: novel loci (orange), Not Previously Reported; Novel index variants reaching genome-wide significance $5x10^{-8}$ are represented as black dots.

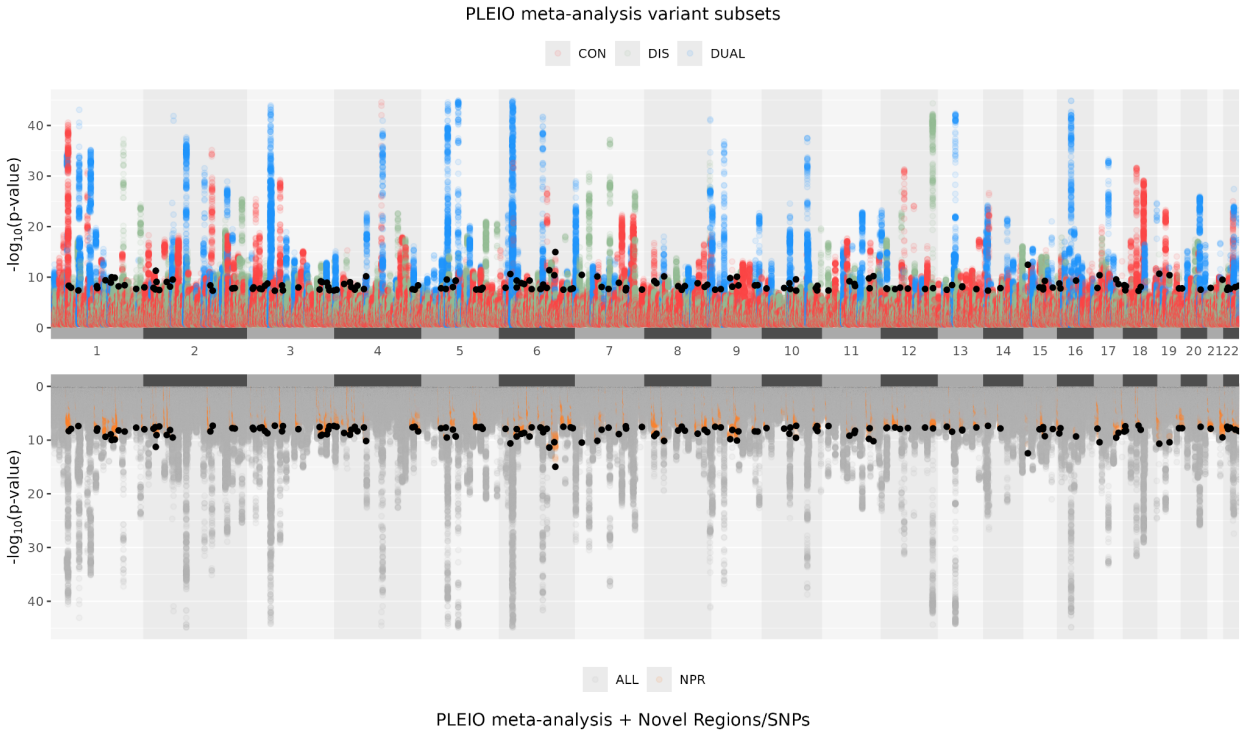


Figure 3: Venn Diagram Comparing Significant PLEIO Loci to Significant Loci from Input GWASs Pleio_meta: results from PLEIO meta-analysis; SCZ_PGC3: Schizophrenia GWAS input (PGC Schizophrenia Working Group[2]); Edu_Lee: Education GWAS[19]; CTP_Lam: Cognitive task performance GWAS[11].

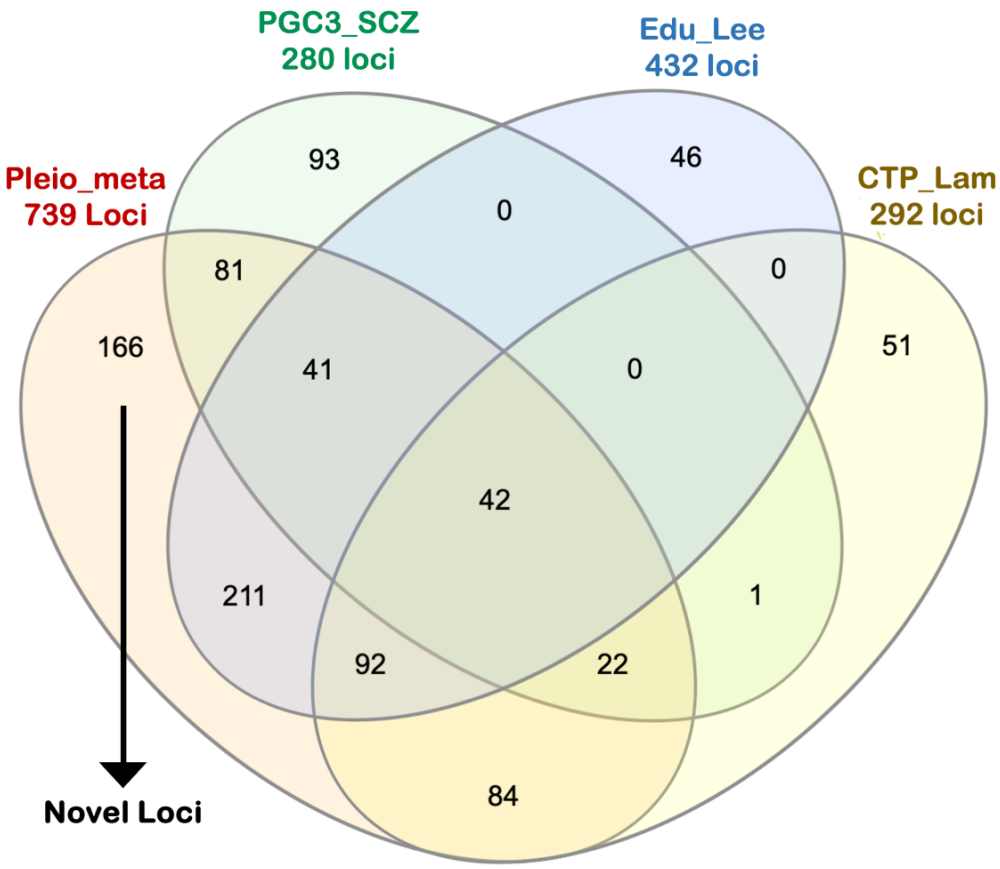
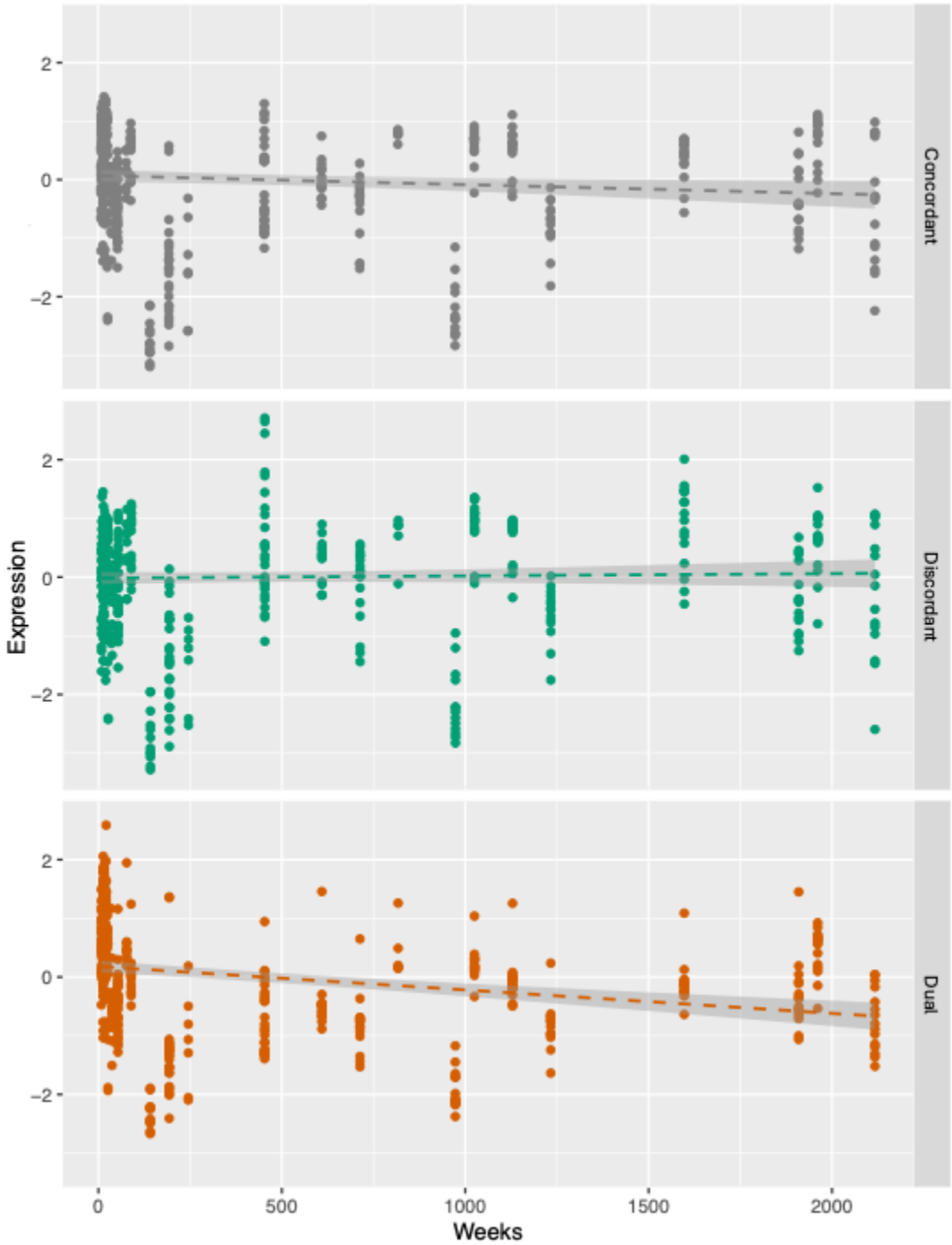


Figure 4. Brainspan Temporal Gene Expression Profiles
Top panel: Brainspan gene expression profiles for 'Concordant' genes. Middle panel: Brainspan gene expression profiles for 'Discordant' genes. Bottom panel: Brainspan gene expression profiles for 'Dual' genes. x-axis: Development time in Weeks. y-axis: standardized (z-scores) of normalized gene expression values in BrainSpan.



Data and Code Availability

1. GWAS summary statistics for pleiotropy analysis and gene set analysis
    a. GWAS: gs://pleiotropy-proj-1/02-interim-testing/MTAG_PLEIO/PLEIO_EDU_Cog_SCZ_ldsc_vcf/PLEIO_Output
    b. FUMA: gs://pleiotropy-proj-1/02-interim-testing/Post_GWAS_Anlalysis/FUMA_results/PLEIO.PLEIO_FUMA_job240552
    c. MAGMA: gs://pleiotropy-proj-1/NoMHC_Extended_25Mb_35Mb
    d. Summary statistics and other auxiliary files to be made available upon acceptance of the manuscript
2. Github links
    a. https://github.com/mlamcogent/cogent-data-curation/tree/main
3. Publicly available GWAS summary statistics
    a. SSGAC https://thessgac.com
    b. Cognitive Task Performance (Lam et al., 2022) https://storage.googleapis.com/broad_institute_mlam/brainstorm-v2-local-gencor-1/03_quality_control_sumstatsqc/07_Data_Release_GWAS_Catalog_01/Lam_et_al_2021_CognitiveTaskPerformance.tsv.gz
    c. PGC3 https://pgc.unc.edu/for-researchers/download-results/
4. Bioinformatic Tools & Resources
    a. Bedtools https://github.com/arq5x/bedtools2?tab=readme-ov-file
    b. Mungesumstats https://github.com/neurogenomics/MungeSumstats
    c. PLEIO https://github.com/cuelee/pleio
    d. MSigDB https://www.gsea-msigdb.org/gsea/msigdb/
    e. BrainSpan https://www.brainspan.org/

Acknowledgments

This work was supported by the National Institute of Mental Health of the National Institutes of Health (NIH) under award no. R01MH117646 (T.L., principal investigator). The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

Competing interests

The authors declare no competing interests.

References

1. Charlson, F.J., Ferrari, A.J., Santomauro, D.F., Diminic, S., Stockings, E., Scott, J.G., McGrath, J.J., and Whiteford, H.A. (2018). Global Epidemiology and Burden of Schizophrenia: Findings From the Global Burden of Disease Study 2016. Schizophr. Bull. *44*, 1195–1203.

2. Trubetskoy, V., Pardiñas, A.F., Qi, T., Panagiotaropoulou, G., Awasthi, S., Bigdeli, T.B., Bryois, J., Chen, C.-Y., Dennison, C.A., Hall, L.S., et al. (2022). Mapping genomic loci implicates genes and synaptic biology in schizophrenia. Nature *604*, 502–508.

3. Schizophrenia Working Group of the Psychiatric Genomics Consortium (2014). Biological insights from 108 schizophrenia-associated genetic loci. Nature *511*, 421–427.

4. Pardiñas, A.F., Holmans, P., Pocklington, A.J., Escott-Price, V., Ripke, S., Carrera, N., Legge, S.E., Bishop, S., Cameron, D., Hamshere, M.L., et al. (2018). Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. Nat. Genet. *50*, 381–389.

5. Ripke, S., O'Dushlaine, C., Chambert, K., Moran, J.L., Kähler, A.K., Akterin, S., Bergen, S.E., Collins, A.L., Crowley, J.J., Fromer, M., et al. (2013). Genome-wide association analysis identifies 13 new risk loci for schizophrenia. Nat. Genet. *45*, 1150–1159.

6. Schizophrenia Psychiatric Genome-Wide Association Study (GWAS) Consortium (2011). Genome-wide association study identifies five new schizophrenia loci. Nat. Genet. *43*, 969–976.

7. Schijven, D., Kofink, D., Tragante, V., Verkerke, M., Pulit, S.L., Kahn, R.S., Veldink, J.H., Vinkers, C.H., Boks, M.P., and Luykx, J.J. (2018). Comprehensive pathway analyses of schizophrenia risk loci point to dysfunctional postsynaptic signaling. Schizophr. Res. *199*, 195–202.

8. Network and Pathway Analysis Subgroup of Psychiatric Genomics Consortium (2015). Psychiatric genome-wide association study analyses implicate neuronal, immune and histone pathways. Nat. Neurosci. *18*, 199–209.

9. Schork, A.J., Won, H., Appadurai, V., Nudel, R., Gandal, M., Delaneau, O., Revsbech Christiansen, M., Hougaard, D.M., Bækved-Hansen, M., Bybjerg-Grauholm, J., et al. (2019). A genome-wide association study of shared risk across psychiatric disorders implicates gene regulation during fetal neurodevelopment. Nat. Neurosci. *22*, 353–361.

10. Lam, M., Hill, W.D., Trampush, J.W., Yu, J., Knowles, E., Davies, G., Stahl, E., Huckins, L., Liewald, D.C., Djurovic, S., et al. (2019). Pleiotropic Meta-Analysis of Cognition, Education, and Schizophrenia Differentiates Roles of Early Neurodevelopmental and Adult Synaptic Pathways. Am. J. Hum. Genet. *105*, 334–350.

11. Lam, M., Chen, C.-Y., Hill, W.D., Xia, C., Tian, R., Levey, D.F., Gelernter, J., Stein, M.B., Hatoum, A.S., Huang, H., et al. (2022). Collective genomic segments with differential pleiotropic patterns between cognitive dimensions and psychopathology. Nat. Commun. *13*, 6868.

12. Harvey, P.D., Bosia, M., Cavallaro, R., Howes, O.D., Kahn, R.S., Leucht, S., Müller, D.R., Penadés, R., and Vita, A. (2022). Cognitive dysfunction in schizophrenia: An expert group paper on the current state of the art. Schizophr Res Cogn *29*, 100249.

13. McCutcheon, R.A., Keefe, R.S.E., and McGuire, P.K. (2023). Cognitive impairment in schizophrenia: aetiology, pathophysiology, and treatment. Mol. Psychiatry *28*, 1902–1918.

14. Bansal, V., Mitjans, M., Burik, C.A.P., Linnér, R.K., Okbay, A., Rietveld, C.A., Begemann, M., Bonn, S., Ripke, S., de Vlaming, R., et al. (2018). Genome-wide association study results for educational attainment aid in identifying genetic heterogeneity of schizophrenia. Nat. Commun. *9*, 3078.

15. Koch, E., Rosenthal, B., Lundquist, A., Chen, C.-H., and Kauppi, K. (2020). Interactome overlap between schizophrenia and cognition. Schizophr. Res. *222*, 167–174.

16. Lencz, T., Knowles, E., Davies, G., Guha, S., Liewald, D.C., Starr, J.M., Djurovic, S., Melle, I., Sundet, K., Christoforou, A., et al. (2014). Molecular genetic evidence for overlap between general cognitive ability and risk for schizophrenia: a report from the Cognitive Genomics consorTium (COGENT). Mol. Psychiatry *19*, 168–174.

17. Smeland, O.B., Frei, O., Kauppi, K., Hill, W.D., Li, W., Wang, Y., Krull, F., Bettella, F., Eriksen, J.A., Witoelar, A., et al. (2017). Identification of Genetic Loci Jointly Influencing Schizophrenia Risk and the Cognitive Traits of Verbal-Numerical Reasoning, Reaction Time, and General Cognitive Function. JAMA Psychiatry *74*, 1065–1075.

18. Lee, C.H., Shi, H., Pasaniuc, B., Eskin, E., and Han, B. (2021). PLEIO: a method to map and interpret pleiotropic loci with GWAS summary statistics. Am. J. Hum. Genet. *108*, 36–48.

19. Lee, J.J., Wedow, R., Okbay, A., Kong, E., Maghzian, O., Zacher, M., Nguyen-Viet, T.A., Bowers, P., Sidorenko, J., Karlsson Linnér, R., et al. (2018). Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. Nat. Genet. *50*, 1112–1121.

20. Okbay, A., Wu, Y., Wang, N., Jayashankar, H., Bennett, M., Nehzati, S.M., Sidorenko, J., Kweon, H., Goldman, G., Gjorgjieva, T., et al. (2022). Polygenic prediction of educational attainment within and between families from genome-wide association analyses in 3 million individuals. Nat. Genet. *54*, 437–449.

21. Murphy, A.E., Schilder, B.M., and Skene, N.G. (2021). MungeSumstats: a Bioconductor package for the standardization and quality control of many GWAS summary statistics. Bioinformatics *37*, 4593–4596.

22. Bhattacharjee, S., Rajaraman, P., Jacobs, K.B., Wheeler, W.A., Melin, B.S., Hartge, P., GliomaScan Consortium, Yeager, M., Chung, C.C., Chanock, S.J., et al. (2012). A subset-based approach improves power and interpretation for the combined analysis of genetic association studies of heterogeneous traits. Am. J. Hum. Genet. *90*, 821–835.

23. Watanabe, K., Taskesen, E., van Bochoven, A., and Posthuma, D. (2017). Functional mapping and annotation of genetic associations with FUMA. Nat. Commun. *8*, 1826.

24. Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841–842.

25. de Leeuw, C.A., Mooij, J.M., Heskes, T., and Posthuma, D. (2015). MAGMA: generalized gene-set analysis of GWAS data. PLoS Comput. Biol. *11*, e1004219.

26. Network and Pathway Analysis Subgroup of the Psychiatric Genomics Consortium (2015). Corrigendum: Psychiatric genome-wide association study analyses implicate neuronal, immune and histone pathways. Nat. Neurosci. *18*, 926.

27. Sey, N.Y.A., Hu, B., Mah, W., Fauni, H., McAfee, J.C., Rajarajan, P., Brennand, K.J., Akbarian, S., and Won, H. (2020). A computational tool (H-MAGMA) for improved prediction of brain-disorder risk genes by incorporating brain chromatin interaction profiles. Nat. Neurosci. *23*, 583–593.

28. Singh, T., Walters, J.T.R., Johnstone, M., Curtis, D., Suvisaari, J., Torniainen, M., Rees, E., Iyegbe, C., Blackwood, D., McIntosh, A.M., et al. (2017). The contribution of rare variants to risk of schizophrenia in individuals with and without intellectual disability. Nat. Genet. *49*, 1167–1173.

29. Gene Ontology resource.

30. Bhattacharya, A., Vo, D.D., Jops, C., Kim, M., Wen, C., Hervoso, J.L., Pasaniuc, B., and Gandal, M.J. (2023). Isoform-level transcriptome-wide association uncovers genetic risk mechanisms for neuropsychiatric disorders in the human brain. Nat. Genet. *55*, 2117–2128.

31. Murray, R.M., Jones, P., O'Callaghan, E., Takei, N., and Sham, P. (1992). Genes, viruses and neurodevelopmental schizophrenia. J. Psychiatr. Res. *26*, 225–235.

32. Bora, E. (2015). Neurodevelopmental origin of cognitive impairment in schizophrenia. Psychol. Med. *45*, 1–9.

33. Kahn, R.S., Sommer, I.E., Murray, R.M., Meyer-Lindenberg, A., Weinberger, D.R., Cannon, T.D., O'Donovan, M., Correll, C.U., Kane, J.M., van Os, J., et al. (2015). Schizophrenia. Nat Rev Dis Primers *1*, 15067.

34. Xu, T., Chan, R.C.K., and Compton, M.T. (2011). Minor physical anomalies in patients with schizophrenia, unaffected first-degree relatives, and healthy controls: a meta-analysis. PLoS One *6*, e24129.

35. Katina, S., Kelly, B.D., Rojas, M.A., Sukno, F.M., McDermott, A., Hennessy, R.J., Lane, A., Whelan, P.F., Bowman, A.W., and Waddington, J.L. (2020). Refining the resolution of craniofacial dysmorphology in bipolar disorder as an index of brain dysmorphogenesis. Psychiatry Res. *291*, 113243.

36. Compton, M.T., and Walker, E.F. (2009). Physical manifestations of neurodevelopmental disruption: are minor physical anomalies part of the syndrome of schizophrenia? Schizophr. Bull. *35*, 425–436.

37. Moberget, T., and Ivry, R.B. (2019). Prediction, psychosis, and the cerebellum. Biol. Psychiatry Cogn. Neurosci. Neuroimaging *4*, 820–831.

38. Zang, Y., and De Schutter, E. (2023). Recent data on the cerebellum require new models and theories. Curr. Opin. Neurobiol. *82*, 102765.

39. Nicoletti, F., Di Menna, L., Iacovelli, L., Orlando, R., Zuena, A.R., Conn, P.J., Dogra, S., and Joffe, M.E. (2023). GPCR interactions involving metabotropic glutamate receptors and their relevance to the pathophysiology and treatment of CNS disorders. Neuropharmacology *235*, 109569.

40. Naur, P., Hansen, K.B., Kristensen, A.S., Dravid, S.M., Pickering, D.S., Olsen, L., Vestergaard, B., Egebjerg, J., Gajhede, M., Traynelis, S.F., et al. (2007). Ionotropic glutamate-like receptor delta2 binds D-serine and glycine. Proc. Natl. Acad. Sci. U. S. A. *104*, 14116–14121.

41. Salazar, J.L., Yang, S.-A., and Yamamoto, S. (2020). Post-Developmental Roles of Notch Signaling in the Nervous System. Biomolecules *10*,.

42. Kim, P. (2024). Understanding the Unfolded Protein Response (UPR) Pathway: Insights into Neuropsychiatric Disorders and Therapeutic Potentials. Biomol. Ther. *32*, 183–191.

43. Kim, P., Scott, M.R., and Meador-Woodruff, J.H. (2021). Dysregulation of the unfolded protein response (UPR) in the dorsolateral prefrontal cortex in elderly patients with schizophrenia. Mol. Psychiatry *26*, 1321–1331.

44. Guha, S., Rees, E., Darvasi, A., Ivanov, D., Ikeda, M., Bergen, S.E., Magnusson, P.K., Cormican, P., Morris, D., Gill, M., et al. (2013). Implication of a rare deletion at distal 16p11.2 in schizophrenia. JAMA Psychiatry *70*, 253–260.

45. Kapur, M., Monaghan, C.E., and Ackerman, S.L. (2017). Regulation of mRNA Translation in Neurons-A Matter of Life and Death. Neuron *96*, 616–637.

46. Costa-Mattioli, M., Sossin, W.S., Klann, E., and Sonenberg, N. (2009). Translational control of long-lasting synaptic plasticity and memory. Neuron *61*, 10–26.

47. English, J.A., Fan, Y., Föcking, M., Lopez, L.M., Hryniewiecka, M., Wynne, K., Dicker, P., Matigian, N., Cagney, G., Mackay-Sim, A., et al. (2015). Reduced protein synthesis in schizophrenia patient-derived olfactory cells. Transl. Psychiatry *5*, e663.

48. Miller, J.A., Ding, S.-L., Sunkin, S.M., Smith, K.A., Ng, L., Szafer, A., Ebbert, A., Riley, Z.L., Royall, J.J., Aiona, K., et al. (2014). Transcriptional landscape of the prenatal human brain. Nature *508*, 199–206.

49. Singh, T., Poterba, T., Curtis, D., Akil, H., Al Eissa, M., Barchas, J.D., Bass, N., Bigdeli, T.B., Breen, G., Bromet, E.J., et al. (2022). Rare coding variants in ten genes confer substantial risk for schizophrenia. Nature *604*, 509–516.

50. DiCorpo, D., LeClair, J., Cole, J.B., Sarnowski, C., Ahmadizar, F., Bielak, L.F., Blokstra, A., Bottinger, E.P., Chaker, L., Chen, Y.-D.I., et al. (2022). Type 2 Diabetes Partitioned Polygenic Scores Associate With Disease Outcomes in 454,193 Individuals Across 13 Cohorts. Diabetes Care *45*, 674–683.

51. Udler, M.S. (2019). Type 2 Diabetes: Multiple Genes, Multiple Diseases. Curr. Diab. Rep. *19*, 55.

52. McCarthy, M.I. (2017). Painting a new picture of personalised medicine for diabetes. Diabetologia *60*, 793–799.

53. Clementz, B.A., Sweeney, J.A., Hamm, J.P., Ivleva, E.I., Ethridge, L.E., Pearlson, G.D., Keshavan, M.S., and Tamminga, C.A. (2016). Identification of Distinct Psychosis Biotypes Using Brain-Based Biomarkers. Am. J. Psychiatry *173*, 373–384.

54. Koen, J.D., Lewis, L., Rugg, M.D., Clementz, B.A., Keshavan, M.S., Pearlson, G.D., Sweeney, J.A., Tamminga, C.A., and Ivleva, E.I. (2023). Supervised machine learning classification of psychosis biotypes based on brain structure: findings from the Bipolar-Schizophrenia network for intermediate phenotypes (B-SNIP). Sci. Rep. *13*, 12980.