Analyses of GWAS and sub-threshold loci lead to the discovery of dendrite development and morphology dysfunction underlying schizophrenia genetic risk

Rui Chen
Vanderbilt University
Quan Wang
Vanderbilt University https://orcid.org/0000-0002-7056-0519
Chongchong Xu
Emory University School of Medicine
Qiang Wei
Vanderbilt University
Hai Yang
Department of Computer Science and Engineering
James Sutcliffe
Vanderbilt University
Yi Jiang
School of Public Health, Tongji Medical College, Huazhong University of Science and Technology Department of Epidemiology and Biostatistics, , , China.
Ying Ji
Vanderbilt University
Feixiong Cheng
Cleveland Clinic
Edwin Cook
University of Illinois at Chicago
Xue Zhong
Vanderbilt University Medical Center
Nancy Cox
Vanderbilt University Medical Center
Zhexing Wen
Emory University https://orcid.org/0000-0002-1518-3845
Bingshan Li (bingshan.li@vanderbilt.edu)
Vanderbilt University

Genetics Article

Keywords: Genetic Risk Pathways, Heritability, Cell Type Specificity, eQTL Analysis, Induced Pluripotent Stem Cell Lines, Neuronal Development

Posted Date: March 2nd, 2021

DOI: https://doi.org/10.21203/rs.3.rs-240477/v1

Abstract

Schizophrenia (SCZ) is highly polygenic, and thousands of genes contribute to its risk. The 145 GWAS loci identified to date do not fully reveal SCZ genetic risk pathways. In this study, we explore a cost-effective strategy to increase power of inference of novel pathways, by expanding the analysis to include sub-threshold GWAS (subGWAS) loci. We identify 180 subGWAS loci (e.g., $5 \times 10^{-8} < P \leq 10^{-6}$) based on SCZ summary statistics of 40,675 cases and 64,643 controls from CLOZUK and PGC datasets, and show that subGWAS loci contain substantial true genetic association signals. We merge GWAS (sigGWAS) and subGWAS loci and identify in total 304 high-confidence risk genes (HRGs) by jointly modeling the expanded set of loci. We identify *dendrite development and morphogenesis (*DDM*, GO:0016358* and *GO:0048813)* as a novel category of biological processes implicated in SCZ genetic risk. SigGWAS loci fail to detect DDM, which is predominantly enriched in subGWAS loci. Further, DDM genes are significantly enriched for heritability of SCZ, as well as bipolar disorder and major depression. Genes in this functional process show cell type specificity in neurons in both fetal and adult brains, and their involvement in SCZ risk is further supported by eQTL analysis of SCZ risk alleles. We derived induced pluripotent stem cell (iPSC) lines from sporadic SCZ patients and normal controls and observe increased neurite lengths and soma sizes in patient-derived iPSC lines along multiple time points during neuronal development, further validating the genetic findings. We also find that the implicated genes are enriched in FDA-approved drug targets, suggesting a therapeutic potential for targeting the implicated biological processes for prevention and treatment. Our results showcase that expanding the analysis to include subGWAS loci is a valuable strategy for enhancing power of uncovering disease mechanisms, especially those of weak effect size, for SCZ and other complex diseases.

Introduction

Schizophrenia (SCZ), a life-spanning psychiatric disorder afflicting approximately 1% of the population, is a leading cause of disability and premature death (Vos *et al.*, 2017; Piotrowski *et al.*, 2017), incurring an immense individual and societal burden upwards of 155 billion dollars in United States annually (Charlson *et al.*, 2018; Cloutier *et al.*, 2016). Incomplete understanding of disease pathophysiology continues to hinder the development of new therapeutics beyond the advent of antipsychotic medications 50 years ago (Owen *et al.*, 2016). Understanding the cause of SCZ is a critical step to relieve the burden.

SCZ is highly heritable, with an estimated heritability of ~80% (Sullivan *et al.*, 2003; Hilker *et al.*, 2018). Identifying risk genes is key to facilitating drug development (Lencz and Malhotra, 2015). During the past decade, more than a hundred loci associated with SCZ have been identified from Genome-wide association studies (GWAS) (Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014; Pardiñas *et al.*, 2018). However, the discovered loci collectively explain only a small proportion of the heritability, leaving the vast majority of genetic loci missing, which is a common challenge referred to as the “missing heritability” conundrum (Manolio *et al.*, 2009) in most complex diseases. Risk genes have been implicated based on GWAS-detected loci (Wang *et al.*, 2019), however, discoveries are constrained by the limited number of loci, and detection of more loci to infer new pathways and mechanisms is

necessary given that thousands of genes are likely involved in SCZ risk (Nguyen *et al.*, 2017). Increasing sample size is an important approach to unveil the hidden part of heritability, but current GWAS sample sizes are already very large, so the return is limited. Alternative approaches to uncovering risk loci, parallel to increasing GWAS samples, are essential. Several studies have shown that GWAS variants slightly below the genome-wide significance threshold ($P \leq 5 \times 10^{-8}$) account for most of the "missing heritability" in GWAS (Shi *et al.*, 2016; Yang *et al.*, 2010); we refer to variants with 'sub-threshold' significance as subGWAS variants hereinafter. The widely accepted genome-wide significance threshold for GWAS is based on a conservative Bonferroni correction, and adoption of such a stringent threshold possibly neglects numerous subGWAS variants that may reach genome-wide significance by inclusion of additional data (Panagiotou *et al.*, 2012). Moreover, a study implicates subGWAS variants as harboring plausible functional, disease-relevant signals (Wang *et al.*, 2016). These findings suggest that exploration into subGWAS loci has real potential to make further discoveries, given a current wealth of data.

Despite the success in identifying risk loci through GWAS, pinpointing corresponding risk genes remains a big challenge to understanding underlying biological mechanisms of SCZ susceptibility (Henriksen *et al.*, 2017). Furthermore, since SCZ is a highly heterogeneous and polygenic disorder, identifying pathways or functional networks through which disease risk alleles exert their impact is crucial for making progress in the prevention and treatment of this condition. Several gene sets and pathways relevant to SCZ pathology have been proposed (Purcell *et al.*, 2014; Hertzberg *et al.*, 2015; Pardiñas *et al.*, 2018). For instance, the involvement of synaptic processes have been repeatedly reported (McGlashan and Hoffman, 2000; Glessner *et al.*, 2010; Lips *et al.*, 2012; Osimo *et al.*, 2019; Berdenis van Berlekom *et al.*, 2020), and this can explain a progressive cortical grey matter reduction observed in imaging studies (Turetsky *et al.*, 1995; Fornito *et al.*, 2009; Heckers and Konradi, 2010). Despite these findings, our knowledge of the disease pathophysiology remains limited. Therefore, the potential value and need for additional mechanistic insights compels exploration into subGWAS risk signals for *bona fide* risk loci.

In this study we demonstrate and utilized the value of subGWAS variants ($5 \times 10^{-8} < P \leq 1 \times 10^{-6}$) in SCZ from different perspectives and combine subGWAS with the genome-wide significant GWAS (sigGWAS) loci together to make robust mechanistic inference based on the expanded loci. Specifically, we use an integrative Risk Gene Selector (iRIGS), a Bayesian model selection framework previously developed to nominate risk genes at GWAS loci (Wang *et al.*, 2019), to jointly analyze both sigGWAS and subGWAS loci together to predict risk genes. We are particularly interested in identifying novel mechanisms involved in SCZ. Based on the expanded gene list, we identified a novel category of biological processes involved in neuronal *dendrite development* and *morphogenesis* (DDM*, GO:0016358* and *GO:0048813*), key functions that are essential for normal synapse formation and maturation. Of particular note is that these biological processes are only enriced in subGWAS signals while absent in sigGWAS loci, supporting the value of subGWAS in identifying pathways with weaker effect size that are otherwise hard to discover without explicitly analyzing subGWAS loci. We validated the role of DDM in SCZ from multiple perspectives, including through normal and SCZ patient-derived induced pluripotent stem cells (iPSC) experiments. Our analyses indicated that subGWAS loci are a valuable resource to gain new insights into

the underlying SCZ risk mechanisms, and the newly implicated DDM processes provide targets for therapeutic development.

Results

SCZ subGWAS loci in large-scale GWAS contain substantial true associated variants

To investigate whether subGWAS loci harbor valid signals, we obtained two SCZ GWAS data sets for analyses, and investigated whether subGWAS loci could convert into sigGWAS with increased sample sizes. The first, published by the Schizophrenia Working Group of the Psychiatric Genomics Consortium (PGC) had a sample size of 36,989 cases and 113,075 controls and identified 108 independent sigGWAS loci (Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014). The second is from a meta-analysis combining the CLOZUK and PGC samples with a total sample size of 40,675 cases and 64,643 controls (Pardiñas *et al.*, 2018), referred to as CLOZUK+PGC hereinafter. CLOZUK+PGC identified 145 sigGWAS loci, among which there are 52 new loci not identified by PGC. We aimed to evaluate whether these new loci overlap subGWAS variants from the PGC study. Briefly, for each index single nucleotide polymorphism (SNP) of the 52 loci, we extended it to a linkage disequilibrium (LD) region by including other SNPs tagged by the index SNP with an r^2 threshold. When requiring $r^2 \geq 0.9$, 38 index SNPs from the 52 loci can be extended to LD regions (with the rest having no available SNPs in LD), among which 21 LD regions (55%) overlap at least one subGWAS SNP identified in the PGC study. With a *loose* criterion of $r^2 \geq 0.2$, all 52 index SNPs can be extended to valid LD regions, and 28 (54%) overlap at least one subGWAS SNP identified in PGC study (SF 1). If we relaxed the criteria of subGWAS in PGC study as $5 \times 10^{-8} < P \leq 1 \times 10^{-5}$, 33 of 38 LD regions (87%) overlap ≥ 1 subGWAS SNP when requiring $r^2 \geq 0.9$ and the ratio is 44 of 52 (85%) when requiring $r^2 \geq 0.2$. We also investigated distributions of *P*-values in those 52 loci from both PGC and CLOZUK+PGC studies and observed similar overall patterns with improved significance in CLOZUK+PGC (SF 2, an example). All these results showed that subGWAS loci in SCZ GWAS with thousands of samples contain true genetic signals.

SCZ subGWAS loci are enriched in brain tissue enhancers

The majority of GWAS SNPs lie in non-coding regions, and they affect their target genes through regulatory effects (Maurano *et al.*, 2012). To evaluate whether subGWAS loci are potentially functional in the brain, we collected enhancers identified in dorsolateral prefrontal cortex (DLPFC) from the ROADMAP Epigenomics Project (Roadmap Epigenomics Consortium *et al.*, 2015) to assess whether subGWAS loci significantly overlap with these enhancers. To this end, we defined 180 independent subGWAS index SNPs based on CLOZUK+PGC data (Methods). For each index SNP, we adopted the $r^2 \geq 0.2$ threshold to define LD regions and counted how many regions thus defined overlapped with DLPFC enhancers. We observed 77.2% (132/171) of subGWAS LD regions overlapped with ≥ 1 DLPFC enhancers, which is significantly higher than background ($P < 1 \times 10^{-3}$, permutation test, SF 3A, Methods), indicating more subGWAS LD regions overlap with DLPFC enhancers. Moreover, we observed that each subGWAS LD region overlaps 6.54 enhancers on average, which is also significantly higher than background ($P = 5 \times$

10^{-3}, permutation test, SF 3B, Methods), indicating each subGWAS LD region overlap with more DLPFC enhancers. The significance also holds for sigGWAS LD regions ($P < 1 \times 10^{-3}$, SF 3C, and $P = 5 \times 10^{-3}$, SF 3D). Next, we compared across all tissue types and observed that subGWAS LD regions overlapped with more enhancers in brain tissues compared to non-brain tissues (SF 4A, B; Methods). This trend is significant in DLPFC when comparing DLPFC versus all non-brain tissues ($P_{sig} = 1.75 \times 10^{-5}$, SF 4A; $P_{sub} = 7.5 \times 10^{-5}$, SF 4B; Wilcoxon test), indicating the enrichment of subGWAS in enhancer regions show brain-specificity.

Risk genes inferred from SCZ subGWAS harbor genuine SCZ risk genes

We next applied iRIGS, a method developed previously to predict SCZ risk genes (Wang *et al.*, 2019), to identify a set of high-confidence risk genes (HRGs). Briefly, iRIGS is a Bayesian model selection framework that leverages multi-omics data and gene network information to predict risk genes from GWAS loci. Features used in iRIGS include differential expression in SZ cases and controls (Fromer *et al.*, 2016), DRE-promoter links (Andersson *et al.*, 2014; Won *et al.*, 2016; Mifsud *et al.*, 2015), *de novo* mutations (Turner *et al.*, 2017), and distance to index SNP (DTS). We combined subGWAS and sigGWAS together to boost the power of risk gene inference, and identified 304 HRGs in total, among which 135 and 177 were identified in sigGWAS and subGWAS loci, respectively (Methods) (ST 1). Accordingly, we termed the 135 genes sigHRGs, the 177 genes subHRGs, and the total 304 genes allHRGs. In addition to HRGs, iRIGS also predicted sets of local background genes (LBGs) as cleaner controls than whole genome background for comparison purposes (Methods); we referred to the corresponding LBGs as sigLBGs, subLBGs, and allLBGs, respectively.

SCZ subGWAS-derived risk genes explain significant SCZ heritability.

We then evaluated SCZ heritability explained by subHRGs using a stratified LD score regression (LDSC) method (Finucane *et al.*, 2015). We included SNPs located within a ±10 kb window centered at the transcription start site (TSS) of each gene for LDSC analysis. We observed a significant enrichment of heritability in subHRGs (enrichment = 8.97, $P = 1.51 \times 10^{-4}$) compared to subLBGs (enrichment = 3.38, $P = 2.28 \times 10^{-3}$, SF 5). We also set the window size as ±100 kb for LDSC and observed consistent results (SF 6). Note that DTS (distance of SNP to gene TSS) used in iRIGS (Methods) is a confounding factor for LDSC, since genes closer to index SNPs are more likely to have a high LDSC score. To avoid this confounding effect, we excluded DTS here in the application of iRIGS to predict HRGs for LDSC enrichment analysis.

SCZ subGWAS-derived risk genes show spatiotemporal expression patterns characteristic of SCZ risk genes.

Next, we explored the spatiotemporal expression pattern of HRGs using expression data from GTEx (GTEx Consortium, 2015) and Brainspan (Miller *et al.*, 2014) (Methods). We observed strong brain specificity for sigHRGs ($P = 8.21 \times 10^{-7}$; Wilcoxon test) and reduced but noticeable brain specificity for

subHRGs (P = 2.58 x 10^{-3}; Wilcoxon test), compared to an absence of brain specificity for LBGs (P = 0.707 for sigLBGs, P = 0.688 for subLBGs, SF 7; Wilcoxon test). In addition, we also observed higher expression levels of both sigHRGs (P = 4.37 x 10^{-4}, SF 8; Wilcoxon test) and subHRGs (P = 4.48 x 10^{-4}, SF 8; Wilcoxon test) at prenatal stages compared to postnatal stages. These results are consistent with the previously discovered spatiotemporal expression pattern of SCZ risk genes (Wang *et al.*, 2019; Sey *et al.*, 2020).

SCZ subHRGs inferred from subGWAS loci expand the knowledge discovered by sigHRGs

To investigate whether subGWAS signals capture knowledge of disease etiology complementary to that inferred only from sigGWAS, we performed gene set enrichment analysis (GSEA) upon 18 gene sets that have been widely and repeatedly implicated in SCZ (Methods). First, we observed that subGWAS loci enhance the significance of enrichment in the seven gene sets that are significant in sigGWAS, with the trend that P values for these sets become more extreme, including evolutionarily constrained genes (ECG, P_{all} = 8.76 X 10^{-17}, P_{sig} = 6.5 x 10^{-11}, P_{sub} = 1.65 x 10^{-6}) and miRNA-137 (miR-137) targets (P = 3.04 X 10^{-6}, P_{sig} = 2.11 x 10^{-3}, P_{sub} = 1.12 x 10^{-2}) (Table 1). Second, combining subHRGs and sigHRGs led to the identification of new significant gene sets that sigHRGs and subHRGs failed to identify separately. For example, sigHRGs and subHRGs are not significantly enriched in the presynaptic active zone (PRAZ) set (P_{sig} = 0.27, P_{sub} = 0.288), while allHRGs show significant enrichment in the same gene set (P_{all} = 6.22 X 10^{-3}, OR = 6.57) after inclusion of subHRGs. The improvement can also be observed for calcium channel and signaling (CCS, P_{all} = 3.55 X 10^{-2}, P_{sig} = 5.03 X 10^{-2} and P_{sub} = 1) and metabotropic glutamate receptor subtype 5 (mGluR5, P_{all} = 1.38 X 10^{-2}, P_{sig} = 7.03 X 10^{-2}, P_{sub} =1), respectively. Lastly, we especially note that subHRGs alone can lead to the identification of new gene sets that sigHRGs fail to detect. For the set of fragile X mental retardation protein (FMRP) targets defined by Ascano et al. (Ascano *et al.*, 2012), the enrichment of subHRGs is significant (P_{sub} = 7.39 X 10^{-3}, OR = 3.13), while the significance is absent for sigHRGs (P_{sig} = 0.23, OR = 2.28). All of these observations imply that subGWAS can enhance and expand the knowledge inferred only from sigGWAS signals.

Table 1. GSEA with previously implicated SCZ-related gene sets. Significant gene sets are marked yellow.

Gene set[a]	allHRGs		sigHRGs		subHRGs	
	$P_{correction}$	OR[b]	$P_{correction}$	OR	$P_{correction}$	OR
ARC (25)	1	5.01 (2/4)	1	Inf (1/1)	1	2.44 (1/3)
AutDB (781)	5.39×10^{-21}	8.77 (56/94)	4.19×10^{-13}	10.72 (31/50)	2.07×10^{-7}	5.98 (25/48)
CCS (73)	3.55×10^{-2}	12.64 (5/7)	5.03×10^{-2}	21.4 (4/5)	1	4.88 (1/2)
ECG (998)	8.76×10^{-17}	7.25 (50/90)	6.50×10^{-11}	7.77 (31/57)	1.65×10^{-6}	6.29 (21/39)
Essential genes (3910)	1.85×10^{-15}	3.16 (121/383)	9.12×10^{-7}	3.06 (54/180)	5.13×10^{-10}	3.4 (70/209)
FMRP-Ascano (939)	2.67×10^{-4}	2.81 (32/93)	0.234	2.28 (14/48)	7.39×10^{-3}	3.13 (18/48)
FMRP-Darnel (832)	1.21×10^{-10}	4.83 (43/93)	1.01×10^{-5}	4.78 (23/52)	1.76×10^{-4}	4.26 (20/45)
GABA (18)	1	5 (1/2)	1	0 (0/1)	1	Inf (1/1)
mGluR5 (37)	1.38×10^{-2}	Inf (4/4)	7.31×10^{-2}	Inf (3/3)	1	Inf (1/1)
miR-137 targets (281)	3.04×10^{-6}	6.66 (19/34)	2.11×10^{-3}	8.3 (9/15)	1.12×10^{-2}	5.09 (10/20)
NMDAR (59)	1	10.02 (2/3)	1	5.25 (1/2)	1	Inf (1/1)
PRAZ (209)	6.22×10^{-3}	6.57 (9/16)	0.27	7.13 (4/7)	0.288	4.97 (5/10)
PRP (336)	0.057	3.65 (10/24)	1	2.97 (5/14)	0.468	4.14 (5/11)
PSD (1444)	7.39×10^{-5}	2.74 (38/113)	9.06×10^{-3}	2.96 (19/56)	4.13×10^{-2}	2.42 (20/63)
PSD-95 (107)	0.558	5.04 (4/8)	0.252	15.94 (3/4)	1	3.27 (2/5)
RBFOX1 (556)	2.27×10^{-6}	5.31 (23/46)	5.1×10^{-4}	6.77 (12/22)	1.52×10^{-2}	4.32 (11/24)
SYV (107)	1	2 (2/7])	1	0 (0/3)	1	4.9 (2/4)
TADA (179)	0.198	3.39 (8/20])	0.234	5.38 (5/10)	1	2.47 (4/12)

[a] The numbers of genes in the corresponding gene sets are in parentheses. [b] The numbers of genes in parentheses stand for the overlap gene numbers in HRG/(HRG+LPG). One-sided Fisher's exact test and Bonferroni correction were used for enrichment analyses. Please refer to the Methods for details of gene set abbreviations. OR: odds ratio;

We also investigated phenotypic manifestations of gene knockouts in mouse models to ask whether mutations in mouse genes orthologous to HRGs exhibit phenotypes related to SCZ (Methods). Among the 1721 terms (> 50 genes) in the Mammalian Phenotype Ontology (MPO) collected from the Mouse Genome Informatics (MGI) database, we observed 201 terms significantly enriched in allHRGs after Bonferroni correction (ST 2). We visualized the significant results of the three HRG lists in directed acyclic graphs (DAGs) to show the underlying hierarchical structure of ontology. The *nervous system phenotype* branch (*MP:0003631, P* = 6.2 X 10^{-21}) stands out from these enriched terms (*SF 9*), consistent with prior knowledge that SCZ reflects perturbations of neurodevelopmental processes (Gulsuner *et al.*, 2013). Some terms, however, are independently identified by subHRGs (Fig. 1A), e.g., *abnormal hindbrain morphology* (*MP:0000841,* P_{sub} = 4.82 x 10^{-2}, P_{sig} = 1), *abnormal hippocampus morphology* (*MP:0000807*, P_{sub} = 1.51 x 10^{-2}, P_{sig} = 0.799), *abnormal brain development* (*MP:0000913,* P_{sub} = 1.13 x 10^{-3}, P_{sig} = 1) and *abnormal forebrain development* (*MP:0003232,* P_{sub} = 3.51 x 10^{-2}, P_{sig} = 1). We also observed that subHRGs improve the power to identify terms that sigHRGs and subHRGs failed to identify individually (Fig. 1B), including *abnormal long-term potentiation* (*MP:0002207,* P_{all} = 3.17 x 10^{-2}, P_{sub} = 0.236, P_{sig} = 1), *abnormal synaptic depression* (*MP:0002915,* P_{all} = 4.03 x 10^{-2}, P_{sub} = 1, P_{sig} = 1), *abnormal neuron differentiation* (*MP:0009937,* P_{all} = 2.46 x 10^{-2}, P_{sub} = 1, P_{sig} = 1) and *abnormal neural tube morphology* (*MP:0002151,* P_{all} = 7.74 x 10^{-3}, P_{sub} = 1, P_{sig} = 0.625). These findings are consistent with previous observations from GSEA of gene sets in SCZ, supporting that subGWAS loci are potentially valuable to expanding our understanding of disease etiology.

SCZ subHRGs lead to novel discovery of dendrite development and morphogenesis implicated in SCZ genetic risk that sigHRGs fail to detect

To gain more specific functional insights underlying SCZ pathophysiology, we collected Gene Ontology (GO) terms under the 'biological process' category and performed GSEA (Methods). We observed significantly enriched terms across different layers of the hierarchical GO structure (SF 10 and ST 3). We are particularly interested in learning more specific and interpretable terms, we restricted further analysis to terms of $\leq$ 500 genes. We visualized enriched terms in Fig. 2, which illustrates three major clusters. In GSEA, sigHRGs make the major contribution to the identification of both clusters I and III (Fig. 2, pink dots), in which the enriched GO terms are related to well-known implicated functions, including synapse related processes (Osimo *et al.*, 2019) in cluster I (SF 11) and ion transportation channels (Giegling *et al.*, 2010; Berridge, 2013; Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014) in cluster III (SF 12).

Genes in dendrite development and morphogenesis terms explain heritability of SCZ and other relevant psychiatric disorders In cluster II, we observed a novel functional module involved in dendrite development and morphogenesis. We note that the major contribution to this is from subHRGs (Fig. 2, orange dots), not from sigHRGs. Specifically, enriched GO terms include *dendrite development* (DD, *GO:001635*, P_{sub} = 2.03 x 10^{-3}, P_{sig} = 0.24) and *dendrite morphogenesis* (DM, *GO:0048813*, P_{sub} = 1.67 x 10^{-3}, P_{sig} = 0.24), *regulation of dendrite development* (RDD*, GO:0050773*, P_{sub} = 4.86 x 10^{-4}, P_{sig} = 0.4), and *regulation of dendrite morphogenesis* (RDM, *GO:0048814*, P_{sub} = 7.61 x 10^{-4}, P_{sig} = 0.54) (Fig. 3 large orange circles). We found that the DD term includes ten genes overlapping subHRGs, and four genes overlapping with sigHRGs (ST 1), explaining why subHRGs are the main driver of identification of this enriched module (Fig. 3, and ST 1). We also found that the enrichment mainly originated from overlaps between terms involving regulation function in subHRGs, i.e., nine of ten genes overlapping between DD and subHRGs derive from RDD (ST 1). A comprehensive literature search revealed various levels of evidence supporting the involvement of DDM in SCZ (listed in ST 4). There are a limited number of terms that sigGWAS are enriched in this cluster (Fig. 3, pink circles), e.g. *negative regulation of nervous system development* (*GO:0051961*, P_{sig} = 6.68 x 10^{-3}), *negative regulation of neurogenesis* (*GO:0050768*, P_{sig} = 4.73 x 10^{-3}) and *negative regulation of axonogenesis* (*GO:0050771*, P_{sig} = 1.97 x 10^{-2}). While synapse dysfunction has been consistently implicated in SCZ pathology (Schijven *et al.*, 2018), our analyses advance the synaptic knowledge of disease pathology from a different perspective, (i.e., DDM) enabled by inferred risk genes at subGWAS loci. Other interesting terms identified by subGWAS include *neuron migration* (*GO:0001764*, P_{sub} = 2.96 x 10^{-3}, P_{sig} = 0.26), *glial cell migration* (*GO:0008347*, P_{sub} = 1.29 x 10^{-2}, P_{sig} = 0.40), *cerebral cortex radial glia guided migration* (*GO:0021801*, P_{sub} = 1.39 x 10^{-2}, P_{sig} = 0.29) (SF 13) and *memory* (*GO:0007613*, P_{sub} = 1.17 x 10^{-2}, P_{sig} = 0.19, SF 14).

To validate that DDM terms are genetically involved in SCZ, we extracted genes from relevant dendrite GO terms and performed LDSC analysis. To avoid confounding effects due to the use of sigGWAS and subGWAS loci in the discovery of the terms, we excluded allHRGs from these GO term genes for LDSC analysis. We observed significant enrichment of heritability in DDM related terms (Fig. 4), e.g., P = 8.25 x 10^{-3}, OR = 6.09 ± 1.92 for DD and P = 5.34 x 10^{-3}, OR = 8.84 ± 2.79 for DM. To further evaluate whether the relevant terms are enriched in genetic risk of related psychiatric disorders, we repeated LDSC analysis using GWAS summary statistics of bipolar disorder (BD, 20,352 cases and 31,358 controls) (Stahl *et al.*, 2019) and major depressive disorder (MDD) (245,363 cases and 561,190 controls) (Howard *et al.*, 2019), which have been previously shown share genetic susceptibility with SCZ (O'Donovan and Owen, 2016; Brainstorm Consortium *et al.*, 2018). We observed similar enrichment of dendrite related terms in genetic risk of BD and MDD (Fig. 4).

Previous studies have indicated that gene-disrupting rare coding variants are more abundant in SCZ cases compared to controls (Genovese *et al.*, 2016), and risk genes identified from common variants (mostly noncoding) are also likely overlapped with genes targeted by rare variants (mostly coding) (Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014). We sought to probe

whether rare coding variants in genes in DD and DM are also associated with SCZ risk. We collected rare variant association summary statistics data at gene level from SCHEMA (https://schema.broadinstitute.org/) (Singh *et al.*, 2020) and observed that genes of DD and DM sets have significantly smaller *P*-values than random background genes (P_{DD} = 8.55 x 10^{-6}, P_{DM} = 1.62 x 10^{-4}, SF 15; Wilcoxon test). After removing allHRGs from GO terms, the *P* values of the remaining genes remain significant (P_{DD} = 7.03 x 10^{-5}, P_{DM} = 5.67 x 10^{-4}, SF 15; Wilcoxon test).**Rare coding variants in genes in dendrite development and morphogenesis are associated with SCZ**

Genes in dendrite development and morphogenesis terms show specificity in neurons at fetal and adult stages

A current neuropathology model of SCZ is a revised neurodevelopment hypothesis proposing the contribution of neurodevelopmental processes at both early (pre- or perinatal stages) and late (childhood to adolescence) stages (Gogtay *et al.*, 2011). The majority of evidence supporting this model comes from neuroimaging studies of adult brain (Pantelis *et al.*, 2003), e.g. reduction of grey matter areas in cortex and hippocampus (Honea *et al.*, 2005; Takahashi *et al.*, 2009; Levitt *et al.*, 2010). However, there is little direct evidence showing abnormalities in fetal brain. Here, we sought to explore whether dendrite related molecular machinery is likely to function in neurons at both fetal and adult stages at cell type level. We collected four sets of single-cell sequencing (scRNA-seq) data from previous studies, including two of fetal brain and two of adult brain (Li *et al.*, 2018; Nowakowski *et al.*, 2017; Lake *et al.*, 2016) , and calculated the cell type specificity of neuronal cells (Methods). We observed that DD genes shows high cell type specificity in both fetal and adult neurons compared to background (Fig 5), including both excitatory (ExN) and inhibitory (InN) neurons. These results imply the potential defects of dendrites in both types of neurons at fetal and adult stages, supporting the revised neurodevelopment hypothesis of SCZ at molecular level.

SCZ subGWAS risk alleles are associated with expression of dendritic genes

Next, we investigated genetic influences on expression of dendritic genes. We collected four brain expression quantitative trait loci (eQTL) studies among which one was performed in fetal human brain (O'Brien *et al.*, 2018), and the other three were performed in adult human brain (Jaffe *et al.*, 2018; Hoffman *et al.*, 2019; GTEx Consortium, 2015).

LRP8 is one of our predicted subHRGs, and it was also previously suggested as a SCZ risk gene (Li *et al.*, 2016). The encoded protein plays a critical role in dendrite development through interacting with reelin signaling pathway (Niu *et al.*, 2004; Myant, 2010). We observed that the risk allele of one subGWAS SNP, rs5174 (risk allele: C, GWAS *P*-value: 1.45 x 10^{-6}), significantly upregulates the expression level of *LRP8* consistently in all four data sets ($P_{Obrien(fetal)}$ = 0.02, $P_{Jaffe(adult)}$ = 1.29 x 10^{-4}, $P_{Hoffman(adult)}$ = 1.34 x 10^{-7}, and $P_{GTEx(adult)} \leq$ 1.9 x 10^{-7} SF 16, ST 5; Wilcoxon test). A previous differential expression study (Fromer *et al.*, 2016) has shown that *LRP8* is over-expressed in SCZ patients (*P* = 1.22 x 10^{-3}), consistent with the eQTL effect here. Another interesting subHRG is *ASAP1*, which regulates the formation of dendritic spines

in dentate gyrus neurons (Yadav *et al.*, 2017; Jain *et al.*, 2012). In the corresponding subGWAS locus, we found rs17212137 (risk allele: T, GWAS *P*-value: 8.74×10^{-7}) is an eQTL down regulating *ASAP1* expression consistently in all the three adult brain studies ($P_{Jaffe(adult)} = 6.19 \times 10^{-5}$, $P_{Hoffman(adult)} = 6.25 \times 10^{-6}$, and $P_{GTEx(adult)} \leq 8.6 \times 10^{-7}$, SF 16, ST5; Wilcoxon test). We note that the observed the association between genetic variant and *LRP8* expression in *both* prenatal and postnatal stages, while the eQTL support of *ASAP1* was *only* significant in adult human brain. We speculate that the genetic influence on risk gene expression could happen at a specific stage or span a long period of time, reflecting the temporal expression pattern of dendritic genes observed in the section above.

Cortical neurons differentiated from SCZ patient-derived iPSC lines show abnormal dendrite development and morphogenesis along multiple time points

As direct examination of neuroanatomical changes related to SCZ is impractical in fetal brain of patients, we employed an induced pluripotent stem cell (iPSC) model that mimics human neuronal development. We generated iPSC cell lines from 4 sporadic SCZ patients and 4 healthy controls and differentiated the iPSC lines into cortical neurons (Methods). We focused on early stages of neuronal development and examined neuronal morphology at 1-4 weeks post differentiation (Methods). We observed significantly longer neurite (dendrite) lengths in developing SCZ neurons compared to controls across all four time points ($P = 3.29 \times 10^{-7}$, 1.45×10^{-11}, 4.35×10^{-10} and 9.33×10^{-4} at 1, 2, 3 and 4 weeks, respectively; Wilcoxon test) (Fig. 6). Moreover, the soma size is significantly larger in SCZ patient-derived iPSC lines than that in controls ($P = 9.72 \times 10^{-10}$, 2.90×10^{-11}, 1.65×10^{-6} and 5.95×10^{-4} at 1, 2, 3 and 4 weeks, respectively; Wilcoxon test). Fig. 6 left panel shows an image of dendrites of neurons from a typical normal control iPSC line (top) and a typical SCZ patient-derived iPSC line (bottom) at 2 weeks, which illustrates the increased dendrite density and soma size.

Genes in dendrite development and morphogenesis terms are likely to be potential drug targets.

Given the important role of DDM in SCZ pathology, we were interested in exploring whether genes corresponding to DD and DM terms have the potential for repositioning existing drugs for SCZ treatment. We utilized a list of 3,078 confirmed druggable targets curated in a previous study (Gaspar and Breen, 2017) for drug target enrichment analyses. Among the 233 genes from DD and DM terms, 70 of them (30%) are drug targets, including two subHRGs, *FYN* and *CDK5R1*. The overlap ratio is significantly higher than random ($P = 1.67 \times 10^{-13}$, OR = 3.18; Fisher's exact test). In particular, 8 dendritic genes (*CHRNA3, CHRNA7, CHRNB2, GRIN3A, DTNBP1, GSK3B, HDAC2,* and *TRPC5*) are targets of nervous system drugs, corresponding to a significant enrichment ($P = 0.02$, OR = 2.60; Fisher's exact test). The protein encoded by *CDK5R1* (p35) is a neuron-specific activator of cyclin-dependent kinase 5 (CDK5), and the decreased expression of p35 observed in SCZ was suggested to affect synaptic protein expression (Gaspar and Breen, 2017). The three genes *CHRNA3, CHRNA7,* and *CHRNB2* are subunits of the neuronal nicotinic acetylcholine receptor family (McKay *et al.*, 2007). *GRIN3A* encodes a subunit of the *N*-methyl-D-aspartate receptor (NMDAR), a key player in controlling synaptic plasticity (Paoletti *et al.*, 2013), while *FYN* tightly regulates NMDAR function (Tezuka *et al.*, 1999; Trepanier *et al.*, 2012). Interestingly, *FYN* is

implicated in SCZ risk in different populations (Wu *et al.*, 2013; Tsavou and Curtis, 2019). *DTNBP1* has been linked to SCZ susceptibility in multiple studies (Narr *et al.*, 2009) and proposed as a potential therapeutic target for SCZ treatment (Wang *et al.*, 2017). The overexpression of another dendritic gene, *HDAC2*, was previously shown to reduce dendritic spine density, synaptic plasticity, and synapse number, while its downregulation led to increased synapse number (Guan *et al.*, 2009). These results exemplify the important roles of DD and DM genes in SCZ etiology, synaptic development and plasticity, supporting the idea of drug repositioning of these genes for SCZ therapy.

Discussion

Although GWAS have successfully identified more than one hundred loci associated with SCZ, they explain only a small proportion of disease heritability. Increasing sample size is a straightforward solution to improve the power of GWAS; however, collection of additional samples that make meaningful contribution to discovery of novel loci is extremely effort consuming, and the current sample size is already exceptionally large. On the other hand, existing SCZ GWAS with a sample size of tens of thousands of individuals already provides us a wealth of data, and gaining additional information by mining this extant resource will be tremendously cost-effective and advance our understanding of the underlying biology of SCZ. As a proof of concept, we revisited two sequential SCZ GWAS and found that a substantial portion of these 'sub-threshold' significant loci from the PGC study would become genome-wide significant after inclusion of additional samples (in the CLOZUK+PGC study). Moreover, we found that brain-related enhancers are enriched in subGWAS loci, consistent with the knowledge that GWAS loci exert their functions through regulatory effects. These findings inspired us to explore the subGWAS loci to increase the discovery power beyond sigGWAS loci. As our primary interest is to make novel discoveries of biological processes involved in SCZ, inferring novel genes beyond GWAS loci is critical in achieving the goal.

Given the complexity of SCZ from multiple angles, inferring risk genes from associated loci is an established challenge. We employed a recently developed and validated computational framework for SCZ, iRGIS, to explore the value of subGWAS loci. When applying iRIGS to predict risk genes, we combined subGWAS and sigGWAS loci together as input, which has two benefits. On one hand, sigGWAS loci are more reliable signals and can improve the accuracy of prediction of subHRGs given that the input loci are jointly modeled. On the other hand, integration of subGWAS and sigGWAS loci increased the number of risk genes and therefore raised the discovery power of underlying biological mechanisms. As expected, some of the functional modules implicated in subHRGs are consistent with sigHRGs. In addition, we especially note that subGWAS loci can also implicate biological functions that are achieved only in subGWAS. Specifically, it is the analysis of subHRGs that led to the discovery of novel biological processes involved in DDM, which sigHRGs fail to detect. Interestingly, dendrite development is a subprocess of synapse development in neurogenesis, while the dysregulation of synapse function has been repeatedly implicated in SCZ (McGlashan and Hoffman, 2000; Glessner *et al.*, 2010; Lips *et al.*, 2012; Osimo *et al.*, 2019; Berdenis van Berlekom *et al.*, 2020). We reason that this is due to the heterogeneity of genetic architecture of SCZ, and to a certain extent, the heterogeneity can be reflected by effect sizes in

the association study. The sigGWAS loci implicating synapse development harbor variants with large effect size, while subGWAS loci implicating DDM harbor variants with relatively weaker effect sizes. The heterogeneity between loci with different effect sizes lead to the discovery of mutually exclusive functional modules even though dendrite development and synapse development are molecularly correlated. Given that functional gene modules may confer different effect sizes, modules with higher effect sizes are more likely to be detected first from GWAS. Most human complex diseases are highly heterogeneous and identified GWAS loci are far from complete due to the relatively limited sample sizes for most. Therefore, many gene modules with weaker effect sizes remain to be discovered. Our strategy presented here demonstrates that focusing on subGWAS signals is a powerful and valuable approach to identify this category of functional gene sets.

Dendritic alterations have long been observed in SCZ and other psychiatric conditions (Konopaske *et al.*, 2014). However, little is known about the link between genetic risk and dendritic dysfunction. In this study, we discovered dendrite dysmorphogenesis in SCZ based largely from GWAS data and validated our findings from multiple orthogonal angles, including from genetic evidence and iPSC models. The observed increase in dendritic length and density in SCZ patient-derived iPSC models compared to control samples is consistent with previous studies (Dimitrion *et al.*, 2017; Forrest *et al.*, 2017), providing further support of the involvement of dendrite development in SCZ genetic risk. For instance, the deletion of 15q11.2 is associated with SCZ, and in a study of hiPSC-derived neurons bearing 15q11.2 deletion, the authors observed an approximately 7.5-fold increase in dendritic density compared with neurons without the deletion (Dimitrion *et al.*, 2017). In another study, the authors found that one SCZ risk allele located near the miR-137 locus can significantly increase dendritic length and arborization in human neurons derived from hiPSCs (Forrest *et al.*, 2017). Reduced dendritic spine density in postmortem brain tissue from SCZ patients, density has also been reported in multiple studies (Konopaske *et al.*, 2014; Moyer *et al.*, 2015; Glausier and Lewis, 2013), seemingly in contradiction to the increased dendritic density observed in hiPSC-derived neurons here. We speculate that SCZ genetic risk leads to increased dendritic density during the prenatal life and early childhood, while neuronal and specifically dendritic dysfunction subsequently triggers aberrant pruning of synapses during adolescence, finally resulting in the reduced density observed in SCZ postmortem brains. Future studies are necessary to accurately understand the underlying mechanisms.

Synaptic dysfunction is long implicated in SCZ (Glantz and Lewis, 2000; Roberts *et al.*, 1996; Berdenis van Berlekom *et al.*, 2020; Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014; Fromer *et al.*, 2014; Osimo *et al.*, 2019; Mirnics *et al.*, 2001). Dendrites are the major sites for synaptic connections (Cline, 2001). Dendritic spines specialized from the shaft contain postsynaptic receptors for the vast majority of excitatory glutamatergic synapses (Sekino *et al.*, 2007). Thus, dendrite development and morphogenesis are critical determinants of synaptic functions. Conversely stabilization of dendrites also requires productive synaptogenesis and active synaptic input (Wu and Cline, 1998; Niell *et al.*, 2004; Sekino *et al.*, 2007). Therefore, dendrites and synapses contribute to the development and function of each. Synaptic loss and dendritic atrophy have been reported for neurodevelopmental and neurodegenerative disorders like SCZ, MDD and Alzheimer disease (AD) (Lin and Koleske, 2010),

suggesting a possible therapeutic strategy of maintaining normal dendritic architecture and restoring its related functions (e.g., synapse development). Perhaps preventing synaptic over-pruning during childhood and adolescence could be beneficial (Sakai, 2020). Supporting evidence from an earlier study showed that people chronically exposed to minocycline, an antibiotic with anti-inflammatory effects that reduces pruning, are at significant lower risk of incident psychosis (Sellgren *et al.*, 2019). Alternatively, restoration of DDM in adulthood may provide a path to treatment. We note that second-generation antipsychotic drugs such as olanzapine and clozapine can enhance neurite outgrowth and increase spine numbers (Critchlow *et al.*, 2006; Lu and Dwyer, 2005). In addition, another study demonstrated an exciting strategy of designing a synthetic synaptic organizer protein to restore dendritic spine numbers, which finally led to the improvement of hippocampus-dependent learning in AD (Suzuki *et al.*, 2020). Moreover, they anticipated that other extracellular scaffold proteins could be used to restore synaptic connectivity in other neurodevelopment and neurodegenerative diseases. Given the fact that therapeutics with direct genetic evidence are more likely to be successful (King *et al.*, 2019; Nelson *et al.*, 2015), such therapies to restore and stabilize dendrite morphogenesis are an important area of further exploration.

Methods

Replication of PGC subGWAS variants in CLOZUK+PGC study

We extracted the 52 new sigGWAS loci from CLOZUK+PGC study (Pardiñas *et al.*, 2018). For each index SNP of the 52 loci, the corresponding LD segment (with relevant PGC study SNPs) was retrieved from HaploReg V4 (Ward and Kellis, 2016). The lowest *P* value SNP in each region was treated as the tag SNP for the region. We also considered the LD segments containing ≥1 subGWAS SNP ($5 \times 10^{-8} < P \leq 1 \times 10^{-6}$) identified in the PGC study.

Defining independent subGWAS index SNPs for CLOZUK+PGC data as input to iRIGS

We collected the GWAS summary statistics from the CLOZUK+PGC study and utilized an iterative strategy to obtain independent subGWAS index SNPs:

In step 1, we removed all the SNPs within ± 500 kb intervals centered on sigGWAS index SNPs to avoid potential overlaps.

In step 2, we selected the SNP with the lowest *P* value from the remaining SNPs, and then removed all the other SNPs within its ± 500 kb region.

In step 3, we iterated step 2 until no further SNPs were left.

In step 4, we retained the selected SNPs with $5 \times 10^{-8} < P \leq 1 \times 10^{-6}$ and defined these SNPs as subGWAS index SNPs.

In total, we obtained 180 subGWAS index SNPs from the CLOZUK+PGC data.

Enhancers in GWAS LD regions

For each index SNP either from sigGWAS or subGWAS, we adopted a threshold of r^2 to define a LD region. Then we collected DLPFC enhancer data from the ROADMAP Epigenomics Project (Roadmap Epigenomics Consortium *et al.*, 2015). To evaluate whether sigGWAS or subGWAS LD regions significantly overlap with DLPFC enhancers, we followed the strategy proposed in a previous study to construct a set of background SNPs used for a permutation test (Wang *et al.*, 2016). Specifically, for either sigGWAS or subGWAS, we picked a matched background SNP list requiring that it has the same number of SNPs as sigGWAS/subGWAS and further each SNP in the background SNP list has similar features to a given index SNP in sigGWAS/subGWAS: matched minor allele frequency (±0.1), gene density in 1 Mb region centered at the selected SNP (±3), location of the picked SNP (whether located in a gene body), distance to TSS of the closest protein-coding gene (±25 kb) and if the LD region of the selected SNP contains a similar number of related SNPs under the same criterion, i.e., $r^2 \geq 0.2$ (±5). We randomly picked 1000 matched background SNP lists.

Three statistics are presented in the related section above:

First, we counted the number of LD regions overlapping ≥ 1 DFPLC enhancers in sigGWAS/subGWAS and 1000 random background SNP lists, respectively. *P* values were calculated using the observed numbers in sigGWAS/subGWAS compared to the background of random lists.

Second, we also calculated the mean number of enhancers overlapping each LD region in sigGWAS/subGWAS and 1000 random lists. *P* values were calculated based on the observed mean and background means.

Third, we collected enhancers from other tissues in addition to DFPLC from ROADMAP Epigenomics Project. To obtain clean results on tissues other than brain, we removed brain enhancers from all non-brain tissues. Then for a given tissue, we calculated the z score for the *i*th SNP in sigGWAS/subGWAS using the formula:

$$Z_i = \frac{C_i - \mathrm{mean}(C_i')}{var(C_i')}$$

Where C_i is the number of enhancers overlapped by the *i*th SNP LD region in sigGWAS/subGWA, C_i' is the number of enhancers overlapped by the corresponding LD regions for matched SNPs sampled. Z score > 0 means a given test SNP overlaps more enhancers than its random background samples and vice versa.

Applying iRIGS to identify HRGs for SCZ

We identified HRGs for SCZ by applying iRIGS (Wang *et al.*, 2019), a Bayesian framework integrates multiple features and a gene network to predict risk genes affected by GWAS variants affect. We adopted the same gene network and features from the original study in our analyses here. Integrated features

include the distance from target gene to index SNP (DTS), DNMs identified in SCZ (Turner *et al.*, 2017, differential expression between SCZ patients and controls (Fromer et al., 2016) , and distal regulatory elements (DREs)-promoter links (Andersson et al., 2014; Won et al., 2016; Mifsud et al., 2015). Refer to the original study for details on how to generate these features and the gene network.

Gene set enrichment analysis

Gene sets previously implicated in SCZ were collected as in (Wang *et al.*, 2019). The gene sets include: FMRP targets (Ascano *et al.*, 2012; Darnell *et al.*, 2011); postsynaptic density (PSD) proteins (Kirov *et al.*, 2012; Bayés *et al.*, 2011); genes related to presynaptic proteins (PRP), PRAZ, and synaptic vesicles (SYV) (Pirooznia *et al.*, 2012); the $GABA_A$ receptor complex (Pocklington *et al.*, 2015); CCS genes (Cross-Disorder Group of the Psychiatric Genomics Consortium, 2013) ; and targets of miR-137 (Schizophrenia Psychiatric Genome-Wide Association Study (GWAS) Consortium, 2011); genes from the database AutDB (Basu *et al.*, 2009); evolutionarily constrained genes (ECG) (Samocha *et al.*, 2014); essential genes (Ji *et al.*, 2016); ASD genes implicated by transmission and de novo association (TADA) test analysis (Sanders *et al.*, 2015); activity-regulated cytoskeleton-associated proteins (ARC), mGluR5 genes, and NMDAR (Kirov *et al.*, 2012); and targets of RNA binding protein, fox-1 homolog 1 (RBFOX1, (Weyn-Vanhentenryck *et al.*, 2014)), a brain- and muscle-specific splicing factor. *P* values were computed by ahypergeometric test and corrected using the Bonferroni method.

The terms of MPO (Smith and Eppig, 2009) were collected from MGI (http://www.informatics.jax.org/) and the corresponding gene sets were assembled as previously described (Wang *et al.*, 2019). In that study, only central nervous system terms were analyzed, and we modified the method here by including all terms to get a more comprehensive phenotypic profile. Briefly, we determined gene list of a specific term by including all the genes in itself and its descendant terms. In this study, we compiled 12490 terms in total and selected 1721 terms with > 50 genes for our analysis. *P* values were computed by hypergeometric tests and corrected for multiple testing using the Bonferroni method.

The GO terms were downloaded from Gene Ontology database (Ashburner *et al.*, 2000) and terms were grouped following the relationships of “is_a”, “part_of”, and “regulates”, i.e., one term is a descendant term of a parent term if it has any of the three kinds of relationships with its parent term. In our hierarchical structure, genes of one term include all the genes residing in its descendant terms. *P* values were calculated using clusterProfiler (Yu *et al.*, 2012) and the false discovery rate (FDR) method was used for multiple testing correction.

Gene expression analysis

We downloaded the transcript per million (TPM) values of genes from the GTEx portal (https://www.gtexportal.org/), and tissue specificity was measured by Preferential Expression Measure (PEM) (Huminiecki *et al.*, 2003; Kryuchkova-Mostacci and Robinson-Rechavi, 2017). Accounting for the fact that tissues from the same organ may be correlated with each other and underlie the same disease, we calculated the tissue specificity in a grouping manner. For a specific tissue, we computed its PEM

versus all tissues in other organs, e.g., we calculated the specificity of frontal cortex by comparing it with all other non-brain tissues but not include brain tissues.

To calculate the developmental-stage specificity, we downloaded the data from BrainSpan (www.brainspan.org). Similarly, we grouped the stages into two major ones: prenatal (S2-S7) and postnatal stages (S8-S14) and calculated PEM for each stage following the same grouping manner, i.e., each stage in a prenatal category was computed comparing only with postnatal stages, and vice versa.

To calculate the specificity of neuron cell types, we collected four sets of single-cell sequencing (scRNA-seq) data from previous studies, including two of fetal brain (Li *et al.*, 2018; Nowakowski *et al.*, 2017) and two of adult brain (Li *et al.*, 2018; Lake *et al.*, 2018). Cell types were annotated in the original studies, including excitatory neuron, inhibitory neuron, astrocyte, microglia, and others. We calculated the specificity of each gene in neuron following the same strategy used in a previous study (Skene *et al.*, 2018), i.e., the mean expression in neuronal cells divided by the sum of mean expression in all cell types. Note that in these studies, sub-clusters (e.g., excitatory neurons in prefrontal cortex and inhibitory neurons in hippocampus) were annotated for excitatory/inhibitory neuron, so we first calculated the specificity for these sub-clusters and then grouped them into two large categories, i.e., excitatory and inhibitory neurons.

Generation and characterization of human iPSC lines

Blood samples (8-10 ml) were collected from all participants (4 sporadic SCZ patients and 4 age- and sex- matched healthy controls; ST 6) used for isolation of peripheral blood mononuclear cells (PBMCs). PBMCs were isolated using Leucosep™ tubes (Greiner Bio-One) according to manufacturer's instructions, and cultured for 9 to 12 days in an erythroblast enrichment medium consisting of 50% Iscove's modified Dulbecco's medium and 50% Ham's F12 medium (Thermo Fisher Scientific), 1X synthetic lipids, and insulin-transferrin-selenium supplement (Thermo Fisher Scientific), 0.5% bovine serum albumin (BSA; Sigma), 50 µg/ml of L-ascorbic acid (Sigma), 200 µM 1-thioglycerol (Sigma), 1X Glutamax (Thermo Fisher Scientific), 50 ng/ml human stem cell factor (SCF) (R&D Systems), 10 ng/ml interleukin (IL)-3 (PeproTech), 100 µg/ml human holo-transferrin (R&D Systems), 40 ng/ml insulin-like growth factor 1 (Peprotech), and 1 µM dexamethasone (Sigma-Aldrich) (Chou *et al.*, 2015) to expand the erythroblast population. All studies followed institutional IRB, ISCRO approved by Emory University School of Medicine. Informed consents were obtained from all individuals.

iPSCs were generated using Cytotune-iPSC 2.0^{TM} Sendai viruses (Thermo Fisher Scientific) according to manufacturer's instructions. After 3 days, transduced cells were plated on irradiated mouse embryonic fibroblast (MEFs) and grown for 2 to 3 weeks in hiPSC medium consisting of D-MEM/F12 (Thermo Fisher Scientific), 20% Knockout Serum Replacement (KSR, Thermo Fisher Scientific), 1X Glutamax (Thermo Fisher Scientific), 1X MEM NEAA (Thermo Fisher Scientific), 100 µM 2-mercaptoethanol (Thermo Fisher Scientific), and 10 ng/ml human basic FGF (bFGF, PeproTech) until the emergence of individual colonies. Live hiPSC clones were manually picked and grown on MEFs using hiPSC medium for further expansion and characterization. Media were changed daily, and iPSC lines were passaged by collagenase (Thermo

Fisher Scientific, 1 mg/ml in D-MEM/F12 for 30 min at 37□C). Karyotyping analysis by standard G-banding technique was carried out by the Line Genetics Inc (Madison, WI). Results were interpreted by clinical laboratory specialists of Cell Line Genetics. Expression of pluripotency-associated markers (NANOG and SSEA4) are validated using immunostaining (SF 17).

Differentiation of iPSCs into forebrain-specific neural progenitors and cortical neurons

Cortical neuron differentiation from human iPSCs was adapted from previously established protocol (Wen *et al.*, 2014; Tang *et al.*, 2016; Bentea *et al.*, 2019). Briefly, hiPSCs colonies were detached from the feeder layer with 1 mg/ml collagenase (Thermo Fisher Scientific) treatment for 30 min and suspended in embryonic body (EB) medium, consisting of bFGF-free iPSC medium supplemented with 2 μM Dorsomorphin (Tocris) and 2 μM A-83 (Tocris), in non-treated polystyrene plates for 4 days with a daily medium change. After 4 days, EB medium was replaced by neural induction medium (NPC medium) consisting of DMEM/F12 (Thermo Fisher Scientific), 1X N2 supplement (Thermo Fisher Scientific), 1X MEM NEAA (Thermo Fisher Scientific), 2 μg/ml heparin (Sigma) and 2 μM cyclopamine (Tocris). The floating EBs were then transferred to Matrigel (Corning)-coated 6-well plates at day 7 to form neural tube-like rosettes. The attached rosettes were kept for 15 days with NPC medium change every other day. On day 22, the rosettes were picked mechanically and transferred to low attachment plates (Corning) to form neurospheres in NPC medium containing 1X B27 (Thermo Fisher Scientific). The neurospheres were then dissociated with Accutase (Thermo Fisher Scientific) and placed onto Poly-D-Lysine/laminin (Sigma)-coated coverslips in the neuronal culture medium, consisting of Neurobasal medium (Thermo Fisher Scientific) supplemented with 1X Glutamax (Thermo Fisher Scientific), 1X B27 (Thermo Fisher Scientific), 1 μM cAMP (Sigma), 200 ng/ml L-Ascorbic Acid (Sigma), 10 ng/ml BDNF (PeproTech) and 10 ng/ml GDNF (PeproTech).

Morphological analyses of hiPSC-derived cortical neurons

Human cortical neurons were fixed with 4% paraformaldehyde (Sigma) for 15 min at room temperature. Samples were permeabilized and blocked with 0.25% Triton X-100 (Sigma) and 10% donkey serum in PBS for 20 min as previously described protocol (Wen *et al.*, 2014). Samples were then incubated with anti-MAP2 primary antibody (Rabbit; 1:1000; Millipore AB5622) at 4□C overnight, followed by incubation with secondary antibody (anti-Rabbit Alexa Fluor 568; 1:1000; ThermoFisher A10037) for 1 hr at room temperature. Images were taken by a Nikon Eclipse Ti-E microscope and acquired from three or four independent cultures with identical settings for parallel cultures. In each culture, 25 fields (600 μm X 600 μm) of images were taken using a 20X objective, and the soma size and total dendritic length from each field were measured using ImageJ with NeurphologyJ plugin (NIH) according to developer's instructions.

Declarations

Acknowledgement

This study is supported by US NIH/NHGRI grants U01HG009086 (to R.C., Q. Wang, Q. Wei, Y.J., H.Y., X.Z., N.J.C., and B.L.), U24HG008956, and R01MH113362 (to J.S., N.J.C., and B.L.). The grant U01HG009086 supports the Vanderbilt Analysis Center for the Genome Sequencing Project (GSP) and U24 HG008956 supports the GSP Coordinating Center. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Author Contributions

R.C. and B.L. conceived the overall design of the study. R.C. and Q. Wang conducted most of the data analyses. C.X. and Z.W. performed the iPSC related experiments and analyses. Q. Wei, Y.J., H.Y., X.Z., and F.C. provided data integration and analyses. J.S.S., E.H.C., and N.J.C. contributed to the interpretation of the results. R.C., Q. Wang and B.L. wrote the manuscript, and all authors participated in the review and revision of the manuscript.

Competing interests

The authors declare no competing interests.

References

Andersson,R. *et al.* (2014) An atlas of active enhancers across human cell types and tissues. *Nature*, **507**, 455–461.

Ascano,M.,Jr *et al.* (2012) FMRP targets distinct mRNA sequence elements to regulate protein expression. *Nature*, **492**, 382–386.

Ashburner,M. *et al.* (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.*, **25**, 25–29.

Basu,S.N. *et al.* (2009) AutDB: a gene reference resource for autism research. *Nucleic Acids Res.*, **37**, D832-6.

Bayés,A. *et al.* (2011) Characterization of the proteome, diseases and evolution of the human postsynaptic density. *Nat. Neurosci.*, **14**, 19–21.

Bentea,E. *et al.* (2019) Kinase network dysregulation in a human induced pluripotent stem cell model of DISC1 schizophrenia. *Mol. Omics*, **15**, 173–188.

Berdenis van Berlekom,A. *et al.* (2020) Synapse Pathology in Schizophrenia: A Meta-analysis of Postsynaptic Elements in Postmortem Brain Studies. *Schizophr. Bull.*, **46**, 374–386.

Brainstorm Consortium *et al.* (2018) Analysis of shared heritability in common disorders of the brain. *Science*, **360**.

Charlson,F.J. *et al.* (2018) Global Epidemiology and Burden of Schizophrenia: Findings From the Global Burden of Disease Study 2016. *Schizophr. Bull.*, **44**, 1195–1203.

Chou,B.-K. *et al.* (2015) A facile method to establish human induced pluripotent stem cells from adult blood cells under feeder-free and xeno-free culture conditions: a clinically compliant approach. *Stem Cells Transl. Med.*, **4**, 320–332.

Cline,H.T. (2001) Dendritic arbor development and synaptogenesis. *Curr. Opin. Neurobiol.*, **11**, 118–126.

Cloutier,M. *et al.* (2016) The Economic Burden of Schizophrenia in the United States in 2013. *J. Clin. Psychiatry*, **77**, 764–771.

Critchlow,H.M. *et al.* (2006) Clozapine and haloperidol differentially regulate dendritic spine formation and synaptogenesis in rat hippocampal neurons. *Mol. Cell. Neurosci.*, **32**, 356–365.

Cross-Disorder Group of the Psychiatric Genomics Consortium (2013) Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. *Lancet*, **381**, 1371–1379.

Darnell,J.C. *et al.* (2011) FMRP stalls ribosomal translocation on mRNAs linked to synaptic function and autism. *Cell*, **146**, 247–261.

Dimitrion,P. *et al.* (2017) Low-Density Neuronal Cultures from Human Induced Pluripotent Stem Cells. *Mol Neuropsychiatry*, **3**, 28–36.

Finucane,H.K. *et al.* (2015) Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nat. Genet.*, **47**, 1228–1235.

Fornito,A. *et al.* (2009) Mapping grey matter reductions in schizophrenia: an anatomical likelihood estimation analysis of voxel-based morphometry studies. *Schizophr. Res.*, **108**, 104–113.

Forrest,M.P. *et al.* (2017) Open Chromatin Profiling in hiPSC-Derived Neurons Prioritizes Functional Noncoding Psychiatric Risk Variants and Highlights Neurodevelopmental Loci. *Cell Stem Cell*, **21**, 305-318.e8.

Fromer,M. *et al.* (2014) De novo mutations in schizophrenia implicate synaptic networks. *Nature*, **506**, 179–184.

Fromer,M. *et al.* (2016) Gene expression elucidates functional impact of polygenic risk for schizophrenia. *Nat. Neurosci.*, **19**, 1442–1453.

Gaspar,H.A. and Breen,G. (2017) Drug enrichment and discovery from schizophrenia genome-wide association results: an analysis and visualisation approach. *Sci. Rep.*, **7**, 12460.

Genovese,G. *et al.* (2016) Increased burden of ultra-rare protein-altering variants among 4,877 individuals with schizophrenia. *Nat. Neurosci.*, **19**, 1433–1441.

Glantz,L.A. and Lewis,D.A. (2000) Decreased dendritic spine density on prefrontal cortical pyramidal neurons in schizophrenia. *Arch. Gen. Psychiatry*, **57**, 65–73.

Glausier,J.R. and Lewis,D.A. (2013) Dendritic spine pathology in schizophrenia. *Neuroscience*, **251**, 90–107.

Glessner,J.T. *et al.* (2010) Strong synaptic transmission impact by copy number variations in schizophrenia. *Proc. Natl. Acad. Sci. U. S. A.*, **107**, 10584–10589.

Gogtay,N. *et al.* (2011) Age of onset of schizophrenia: perspectives from structural neuroimaging studies. *Schizophr. Bull.*, **37**, 504–513.

GTEx Consortium (2015) Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science*, **348**, 648–660.

Guan,J.-S. *et al.* (2009) HDAC2 negatively regulates memory formation and synaptic plasticity. *Nature*, **459**, 55–60.

Gulsuner,S. *et al.* (2013) Spatial and temporal mapping of de novo mutations in schizophrenia to a fetal prefrontal cortical network. *Cell*, **154**, 518–529.

Heckers,S. and Konradi,C. (2010) Hippocampal pathology in schizophrenia. *Curr. Top. Behav. Neurosci.*, **4**, 529–553.

Henriksen,M.G. *et al.* (2017) Genetics of Schizophrenia: Overview of Methods, Findings and Limitations. *Front. Hum. Neurosci.*, **11**, 322.

Hertzberg,L. *et al.* (2015) Integration of gene expression and GWAS results supports involvement of calcium signaling in Schizophrenia. *Schizophr. Res.*, **164**, 92–99.

Hilker,R. *et al.* (2018) Heritability of Schizophrenia and Schizophrenia Spectrum Based on the Nationwide Danish Twin Register. *Biol. Psychiatry*, **83**, 492–498.

Hoffman,G.E. *et al.* (2019) CommonMind Consortium provides transcriptomic and epigenomic data for Schizophrenia and Bipolar Disorder. *Sci. Data*, **6**, 180.

Honea,R. *et al.* (2005) Regional deficits in brain volume in schizophrenia: a meta-analysis of voxel-based morphometry studies. *Am. J. Psychiatry*, **162**, 2233–2245.

Howard,D.M. *et al.* (2019) Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions. *Nat. Neurosci.*

Huminiecki,L. *et al.* (2003) Congruence of tissue expression profiles from Gene Expression Atlas, SAGEmap and TissueInfo databases. *BMC Genomics*, **4**, 31.

Jaffe,A.E. *et al.* (2018) Developmental and genetic regulation of the human cortex transcriptome illuminate schizophrenia pathogenesis. *Nat. Neurosci.*, **21**, 1117–1125.

Jain,S. *et al.* (2012) Arf4 determines dentate gyrus-mediated pattern separation by regulating dendritic spine development. *PLoS One*, **7**, e46340.

Ji,X. *et al.* (2016) Increased burden of deleterious variants in essential genes in autism spectrum disorder. *Proc. Natl. Acad. Sci. U. S. A.*, 201613195.

King,E.A. *et al.* (2019) Are drug targets with genetic support twice as likely to be approved? Revised estimates of the impact of genetic support for drug mechanisms on the probability of drug approval. *PLoS Genet.*, **15**, e1008489.

Kirov,G. *et al.* (2012) De novo CNV analysis implicates specific abnormalities of postsynaptic signalling complexes in the pathogenesis of schizophrenia. *Mol. Psychiatry*, **17**, 142–153.

Konopaske,G.T. *et al.* (2014) Prefrontal cortical dendritic spine pathology in schizophrenia and bipolar disorder. *JAMA Psychiatry*, **71**, 1323–1331.

Kryuchkova-Mostacci,N. and Robinson-Rechavi,M. (2017) A benchmark of gene expression tissue-specificity metrics. *Brief. Bioinform.*, **18**, 205–214.

Lake,B.B. *et al.* (2018) Integrative single-cell analysis of transcriptional and epigenetic states in the human adult brain. *Nat. Biotechnol.*, **36**, 70–80.

Lake,B.B. *et al.* (2016) Neuronal subtypes and diversity revealed by single-nucleus RNA sequencing of the human brain. *Science*, **352**, 1586–1590.

Lencz,T. and Malhotra,A.K. (2015) Targeting the schizophrenia genome: a fast track strategy from GWAS to clinic. *Mol. Psychiatry*, **20**, 820–826.

Levitt,J.J. *et al.* (2010) A selective review of volumetric and morphometric imaging in schizophrenia. *Curr. Top. Behav. Neurosci.*, **4**, 243–281.

Li,M. *et al.* (2016) Convergent lines of evidence support LRP8 as a susceptibility gene for psychosis. *Mol. Neurobiol.*, **53**, 6608–6619.

Li,M. *et al.* (2018) Integrative functional genomic analysis of human brain development and neuropsychiatric risks. *Science*, **362**, eaat7615.

Lin,Y.-C. and Koleske,A.J. (2010) Mechanisms of synapse and dendrite maintenance and their disruption in psychiatric and neurodegenerative disorders. *Annu. Rev. Neurosci.*, **33**, 349–378.

Lips,E.S. *et al.* (2012) Functional gene group analysis identifies synaptic gene groups as risk factor for schizophrenia. *Mol. Psychiatry*, **17**, 996–1006.

Lu,X.-H. and Dwyer,D.S. (2005) Second-generation antipsychotic drugs, olanzapine, quetiapine, and clozapine enhance neurite outgrowth in PC12 cells via PI3K/AKT, ERK, and pertussis toxin-sensitive pathways. *J. Mol. Neurosci.*, **27**, 43–64.

Manolio,T.A. *et al.* (2009) Finding the missing heritability of complex diseases. *Nature*, **461**, 747–753.

Maurano,M.T. *et al.* (2012) Systematic localization of common disease-associated variation in regulatory DNA. *Science*, **337**, 1190–1195.

McGlashan,T.H. and Hoffman,R.E. (2000) Schizophrenia as a disorder of developmentally reduced synaptic connectivity. *Arch. Gen. Psychiatry*, **57**, 637–648.

McKay,B.E. *et al.* (2007) Regulation of synaptic transmission and plasticity by neuronal nicotinic acetylcholine receptors. *Biochem. Pharmacol.*, **74**, 1120–1133.

Mifsud,B. *et al.* (2015) Mapping long-range promoter contacts in human cells with high-resolution capture Hi-C. *Nat. Genet.*, **47**, 598–606.

Miller,J.A. *et al.* (2014) Transcriptional landscape of the prenatal human brain. *Nature*, **508**, 199–206.

Mirnics,K. *et al.* (2001) Analysis of complex brain disorders with gene expression microarrays: schizophrenia as a disease of the synapse. *Trends Neurosci.*, **24**, 479–486.

Moyer,C.E. *et al.* (2015) Dendritic spine alterations in schizophrenia. *Neurosci. Lett.*, **601**, 46–53.

Myant,N.B. (2010) Reelin and apolipoprotein E receptor 2 in the embryonic and mature brain: effects of an evolutionary change in the apoER2 gene. *Proc. Biol. Sci.*, **277**, 345–351.

Narr,K.L. *et al.* (2009) DTNBP1 is associated with imaging phenotypes in schizophrenia. *Hum. Brain Mapp.*, **30**, 3783–3794.

Nelson,M.R. *et al.* (2015) The support of human genetic evidence for approved drug indications. *Nat. Genet.*, **47**, 856–860.

Nguyen,H.T. *et al.* (2017) Integrated Bayesian analysis of rare exonic variants to identify risk genes for schizophrenia and neurodevelopmental disorders. *Genome Med.*, **9**, 114.

Niell,C.M. *et al.* (2004) In vivo imaging of synapse formation on a growing dendritic arbor. *Nat. Neurosci.*, **7**, 254–260.

Niu,S. *et al.* (2004) Reelin promotes hippocampal dendrite development through the VLDLR/ApoER2-Dab1 pathway. *Neuron*, **41**, 71–84.

Nowakowski,T.J. *et al.* (2017) Spatiotemporal gene expression trajectories reveal developmental hierarchies of the human cortex. *Science*, **358**, 1318–1323.

O'Brien,H.E. *et al.* (2018) Expression quantitative trait loci in the developing human brain and their enrichment in neuropsychiatric disorders. *Genome Biol.*, **19**, 194.

O'Donovan,M.C. and Owen,M.J. (2016) The implications of the shared genetics of psychiatric disorders. *Nat. Med.*, **22**, 1214–1219.

Osimo,E.F. *et al.* (2019) Synaptic loss in schizophrenia: a meta-analysis and systematic review of synaptic protein and mRNA measures. *Mol. Psychiatry*, **24**, 549–561.

Owen,M.J. *et al.* (2016) Schizophrenia. *Lancet*, **388**, 86–97.

Panagiotou,O.A. *et al.* (2012) What should the genome-wide significance threshold be? Empirical replication of borderline genetic associations. *Int. J. Epidemiol.*, **41**, 273–286.

Pantelis,C. *et al.* (2003) Early and late neurodevelopmental disturbances in schizophrenia and their functional consequences. *Aust. N. Z. J. Psychiatry*, **37**, 399–406.

Paoletti,P. *et al.* (2013) NMDA receptor subunit diversity: impact on receptor properties, synaptic plasticity and disease. *Nat. Rev. Neurosci.*, **14**, 383–400.

Pardiñas,A.F. *et al.* (2018) Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. *Nat. Genet.*

Piotrowski,P. *et al.* (2017) Causes of mortality in schizophrenia: An updated review of European studies. *Psychiatr. Danub.*, **29**, 108–120.

Pirooznia,M. *et al.* (2012) SynaptomeDB: an ontology-based knowledgebase for synaptic genes. *Bioinformatics*, **28**, 897–899.

Pocklington,A.J. *et al.* (2015) Novel findings from CNVs implicate inhibitory and excitatory signaling complexes in schizophrenia. *Neuron*, **86**, 1203–1214.

Purcell,S.M. *et al.* (2014) A polygenic burden of rare disruptive mutations in schizophrenia. *Nature*, **506**, 185–190.

Roadmap Epigenomics Consortium *et al.* (2015) Integrative analysis of 111 reference human epigenomes. *Nature*, **518**, 317–330.

Roberts,R.C. *et al.* (1996) Reduced striatal spine size in schizophrenia. *Neuroreport*, **7**, 1214–1218.

Sakai,J. (2020) Core Concept: How synaptic pruning shapes neural wiring during development and, possibly, in disease. *Proc. Natl. Acad. Sci. U. S. A.*, **117**, 16096–16099.

Samocha,K.E. *et al.* (2014) A framework for the interpretation of de novo mutation in human disease. *Nat. Genet.*, **46**, 944–950.

Sanders,S.J. *et al.* (2015) Insights into Autism Spectrum Disorder Genomic Architecture and Biology from 71 Risk Loci. *Neuron*, **87**, 1215–1233.

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

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

Sekino,Y. *et al.* (2007) Role of actin cytoskeleton in dendritic spine morphogenesis. *Neurochem. Int.*, **51**, 92–104.

Sellgren,C.M. *et al.* (2019) Increased synapse elimination by microglia in schizophrenia patient-derived models of synaptic pruning. *Nat. Neurosci.*, **22**, 374–385.

Shi,H. *et al.* (2016) Contrasting the genetic architecture of 30 complex traits from summary association data. *Am. J. Hum. Genet.*, **99**, 139–153.

Singh,T. *et al.* (2020) Exome sequencing identifies rare coding variants in 10 genes which confer substantial risk for schizophrenia. *medRxiv*.

Skene,N.G. *et al.* (2018) Genetic identification of brain cell types underlying schizophrenia. *Nat. Genet.*, **50**, 825–833.

Smith,C.L. and Eppig,J.T. (2009) The mammalian phenotype ontology: enabling robust annotation and comparative analysis. *Wiley Interdiscip. Rev. Syst. Biol. Med.*, **1**, 390–399.

Stahl,E.A. *et al.* (2019) Genome-wide association study identifies 30 loci associated with bipolar disorder. *Nat. Genet.*, **51**, 793–803.

Sullivan,P.F. *et al.* (2003) Schizophrenia as a complex trait: evidence from a meta-analysis of twin studies. *Arch. Gen. Psychiatry*, **60**, 1187–1192.

Suzuki,K. *et al.* (2020) A synthetic synaptic organizer protein restores glutamatergic neuronal circuits. *Science*, **369**.

Takahashi,T. *et al.* (2009) Progressive gray matter reduction of the superior temporal gyrus during transition to psychosis. *Arch. Gen. Psychiatry*, **66**, 366–376.

Tang,H. *et al.* (2016) Zika virus infects human cortical neural progenitors and attenuates their growth. *Cell Stem Cell*, **18**, 587–590.

Tezuka,T. *et al.* (1999) PSD-95 promotes Fyn-mediated tyrosine phosphorylation of the N-methyl-D-aspartate receptor subunit NR2A. *Proc. Natl. Acad. Sci. U. S. A.*, **96**, 435–440.

Trepanier,C.H. *et al.* (2012) Regulation of NMDA receptors by the tyrosine kinase Fyn. *FEBS J.*, **279**, 12–19.

Tsavou,A. and Curtis,D. (2019) In-silico investigation of coding variants potentially affecting the functioning of the glutamatergic N-methyl-D-aspartate receptor in schizophrenia. *Psychiatr. Genet.*, **29**, 44–50.

Turetsky,B. *et al.* (1995) Frontal and temporal lobe brain volumes in schizophrenia. Relationship to symptoms and clinical subtype. *Arch. Gen. Psychiatry*, **52**, 1061–1070.

Turner,T.N. *et al.* (2017) denovo-db: a compendium of human de novo variants. *Nucleic Acids Res.*, **45**, D804–D811.

Vos,T. *et al.* (2017) Global, regional, and national incidence, prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet*, **390**, 1211–1259.

Wang,H. *et al.* (2017) Dysbindin-1 involvement in the etiology of schizophrenia. *Int. J. Mol. Sci.*, **18**.

Wang,Q. *et al.* (2019) A Bayesian framework that integrates multi-omics data and gene networks predicts risk genes from schizophrenia GWAS data. *Nat. Neurosci.*, **22**, 691–699.

Wang,X. *et al.* (2016) Discovery and validation of sub-threshold genome-wide association study loci using epigenomic signatures. *eLife Sciences*, **5**, e10557.

Ward,L.D. and Kellis,M. (2016) HaploReg v4: systematic mining of putative causal variants, cell types, regulators and target genes for human complex traits and disease. *Nucleic Acids Res.*, **44**, D877-81.

Wen,Z. *et al.* (2014) Synaptic dysregulation in a human iPS cell model of mental disorders. *Nature*, **515**, 414–418.

Weyn-Vanhentenryck,S.M. *et al.* (2014) HITS-CLIP and integrative modeling define the Rbfox splicing-regulatory network linked to brain development and autism. *Cell Rep.*, **6**, 1139–1152.

Won,H. *et al.* (2016) Chromosome conformation elucidates regulatory relationships in developing human brain. *Nature*, **538**, 523–527.

Wu,G.Y. and Cline,H.T. (1998) Stabilization of dendritic arbor structure in vivo by CaMKII. *Science*, **279**, 222–226.

Wu,L. *et al.* (2013) Association study of the Fyn gene with schizophrenia in the Chinese-Han population. *Psychiatr. Genet.*, **23**, 39–40.

Yadav,S. *et al.* (2017) TAOK2 kinase mediates PSD95 stability and dendritic spine maturation through Septin7 phosphorylation. *Neuron*, **93**, 379–393.

Yang,J. *et al.* (2010) Common SNPs explain a large proportion of the heritability for human height. *Nat. Genet.*, **42**, 565.

Yu,G. *et al.* (2012) clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS*, **16**, 284–287.

Figures

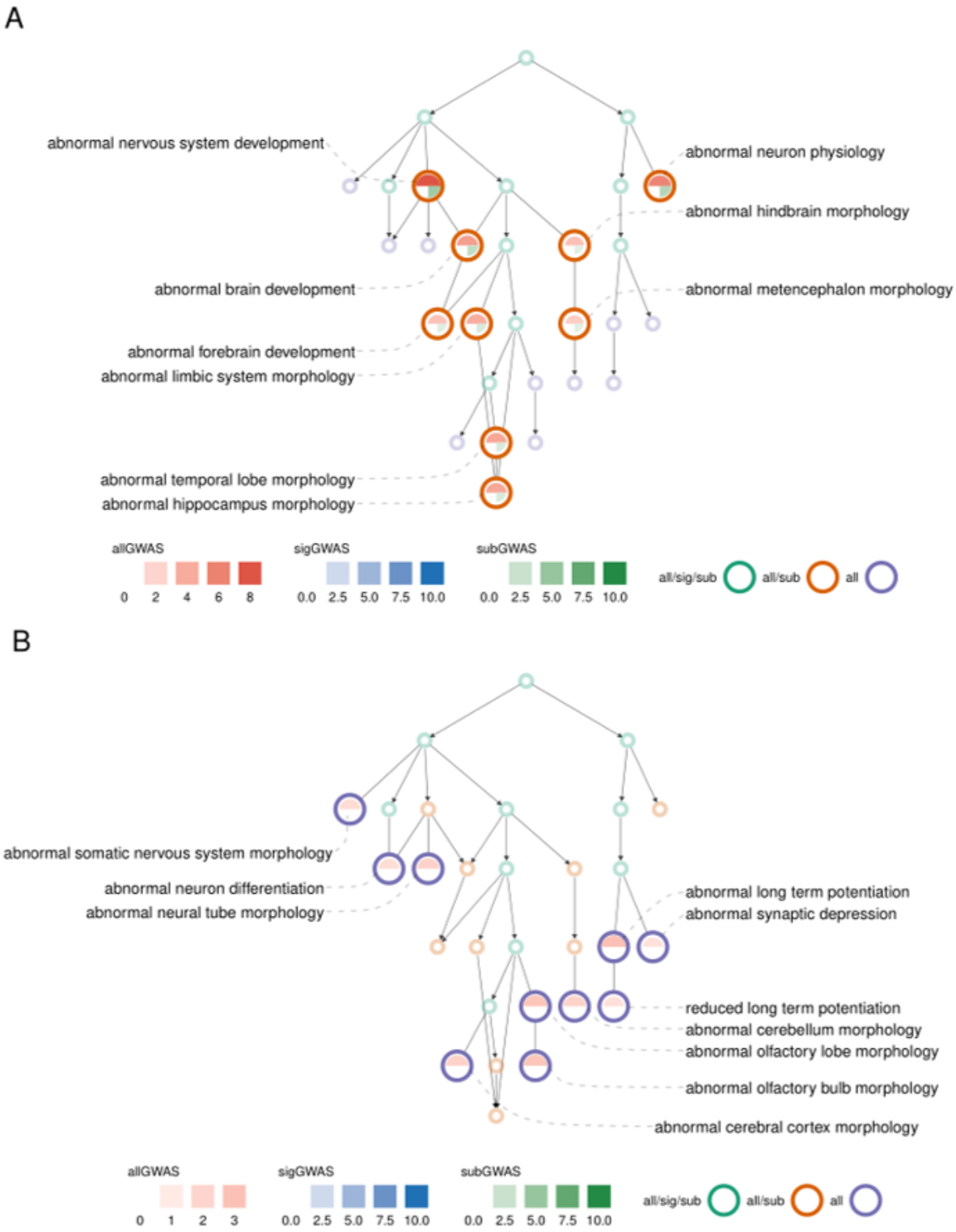

Figure 1

DAG of MPO terms under nervous system phenotype (MP:0003631). (A) terms identified by subGWAS independently. (B) Terms identified only by combination of subGWAS and sigGWAS, indicating the boosted power of subGWAS. all/sig/sub (green), significant in allHRGs, sigHRGs and subHRGs; all/sub (orange), significant in allHRGs and subHRGs; all (blue), significant in allHRGs. The terminology is adopted by other DAGs as well.

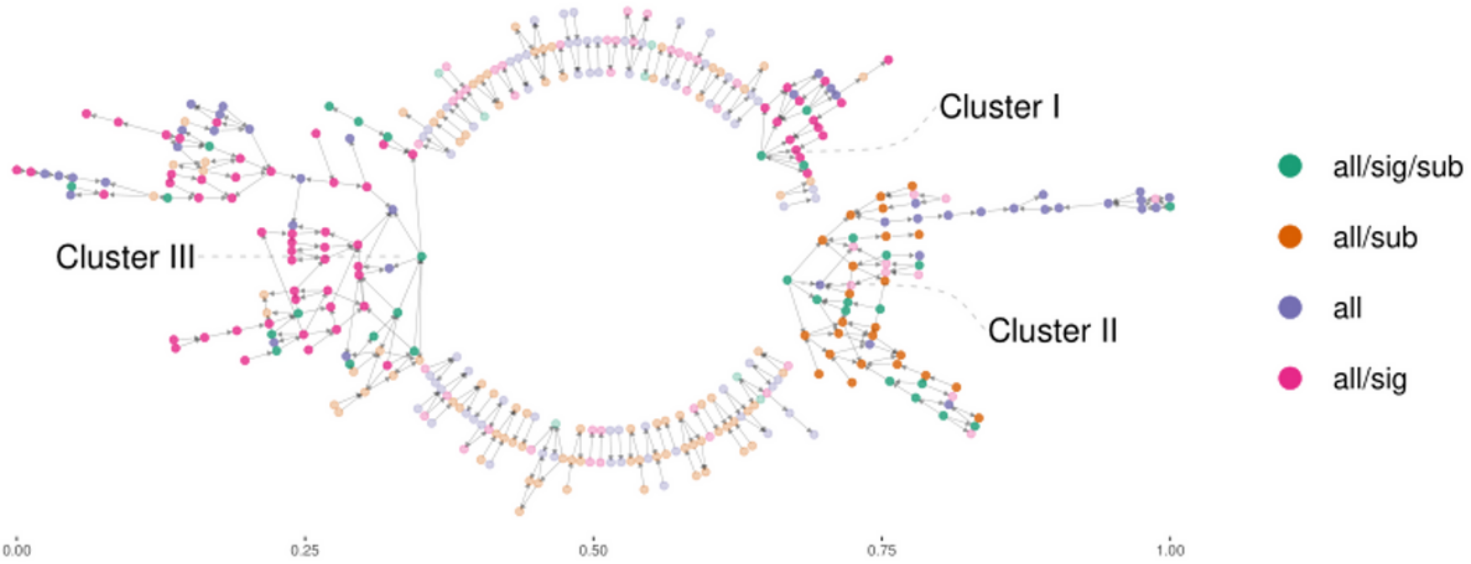

Figure 2

DAG of GSEA enrichment in GO terms with genes ≤ 500. Three clusters of terms stand out. Cluster I and III are mostly contributed by all/sig and all/sig/sub and Cluster II is mostly contributed by all/sub and all/sig/sub.

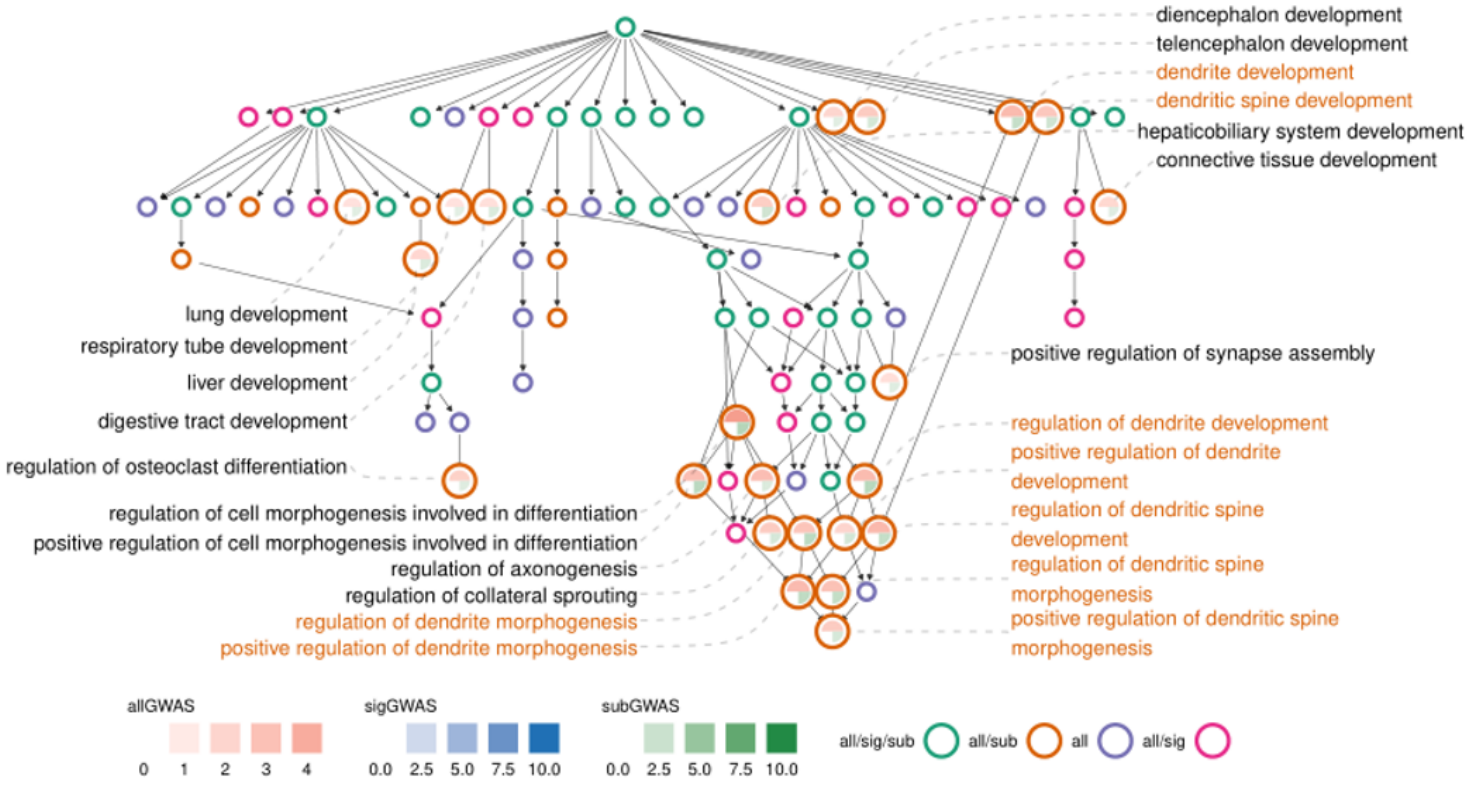

Figure 3

DAG of dendrite related terms under GO term of anatomical structure development (GO:0048856) identified by subGWAS. Dendrite terms are marked in orange text.

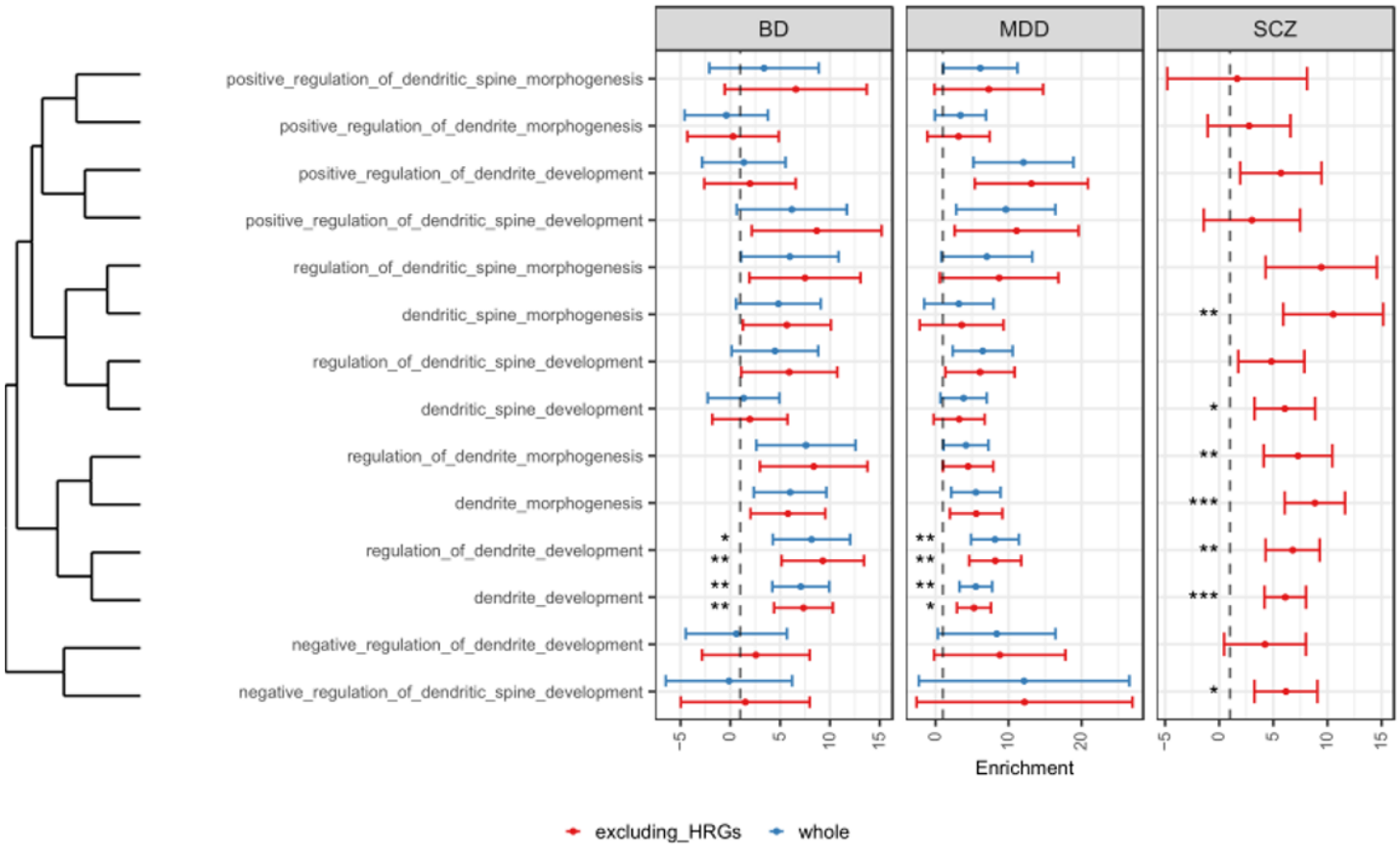

Figure 4

LDSC results of dendrite-related terms across psychiatric disorders. *: P ≤ 0.1; **: P ≤ 0.05; ***: P ≤ 0.01. The similarities of GO terms are computed using Jaccard index method.

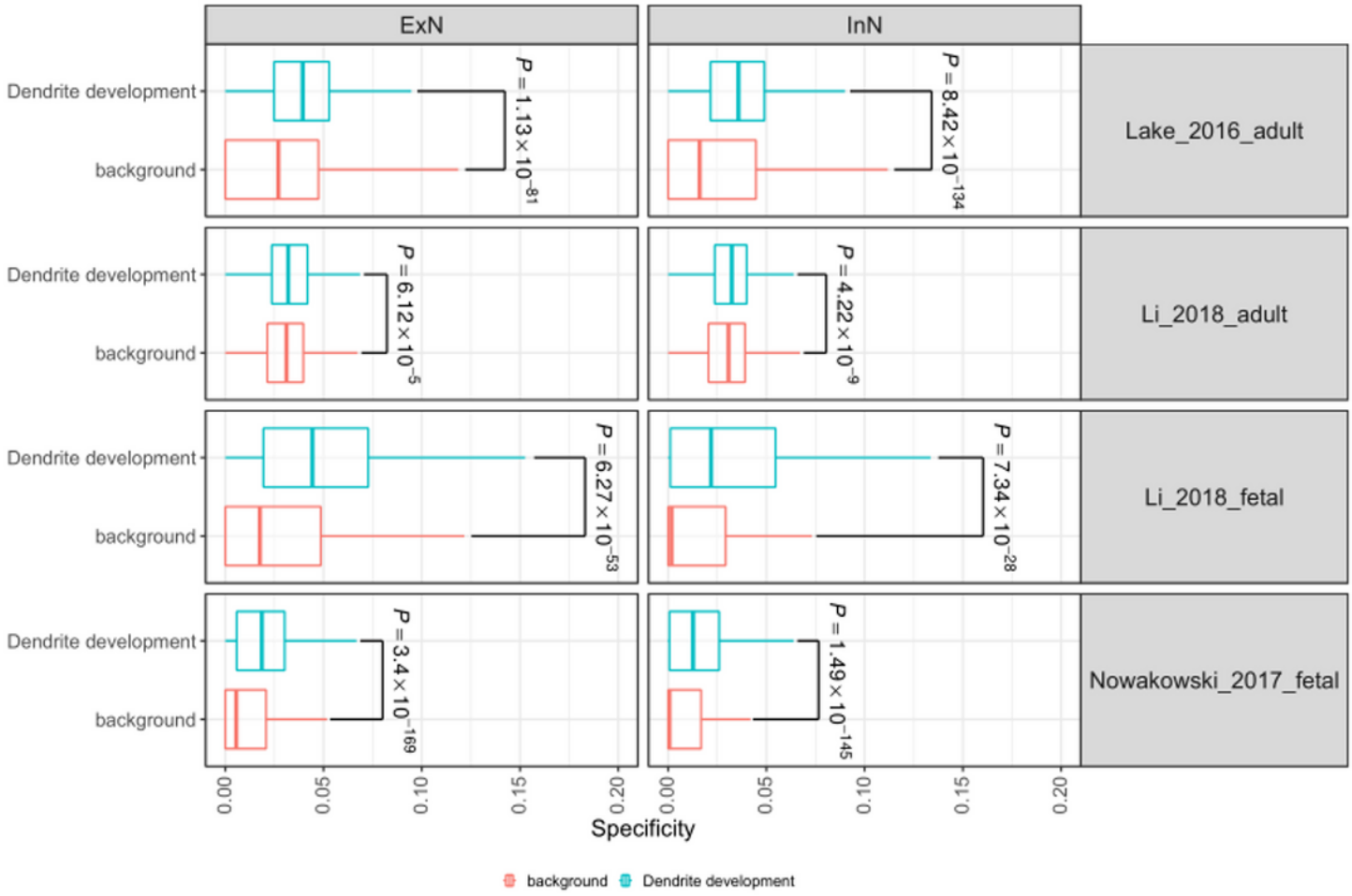

Figure 5

Expressions of dendrite development (GO:0016358) genes in neurons of fetal and adult brain.

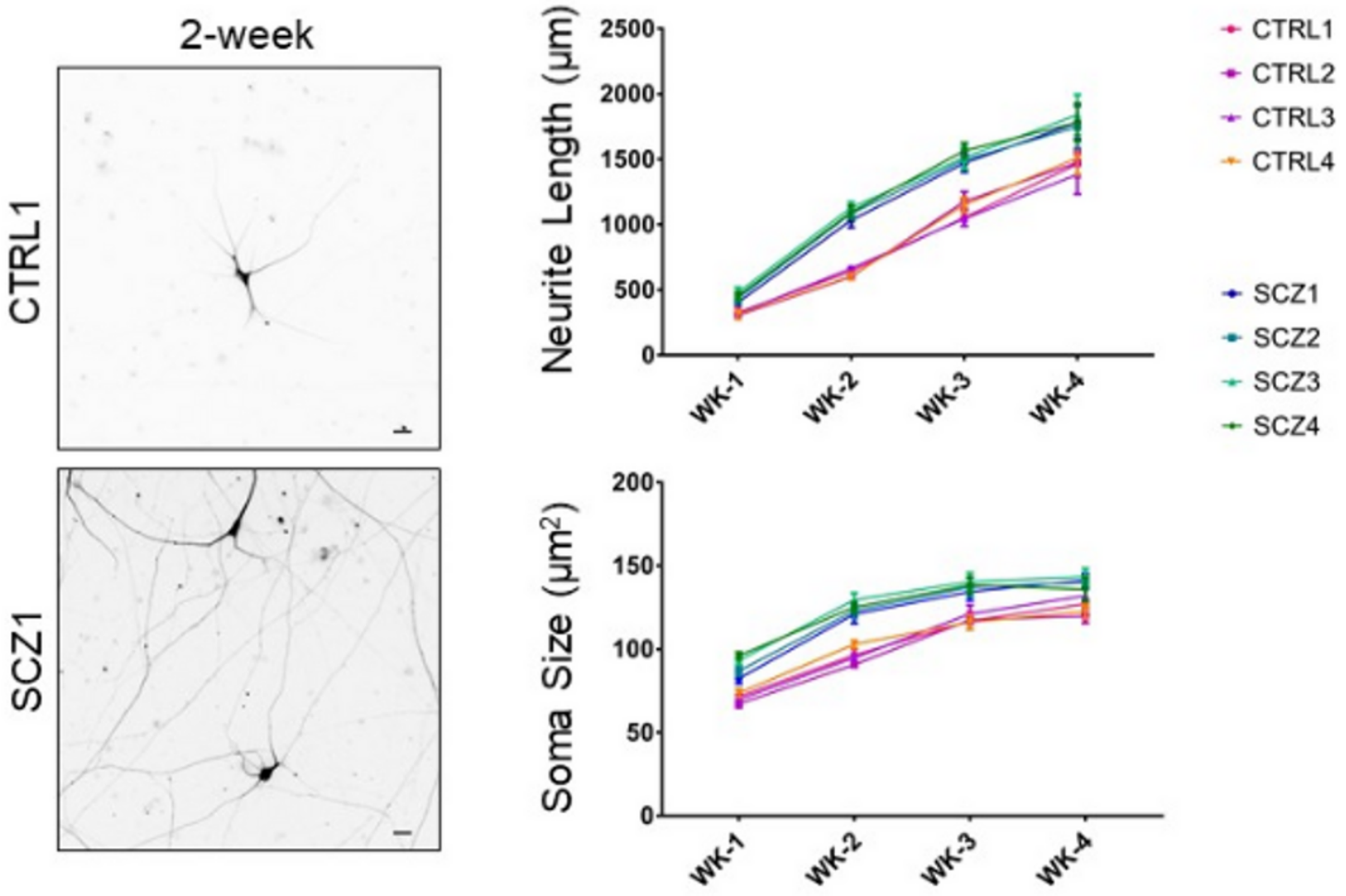

Figure 6

iPSC induced neurons show different dendrite development and morphology in SCZ patients. Data of 4 control and 4 SCZ patients at 4 time points of week 1/2/3/4 are collected along with the development of neuron.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SF.docx
- ST.xlsx