Regulatory variants at 2q33.1 confer schizophrenia risk by modulating distal gene TYW5 expression

Shiwu Li,1,2,† Jiao Li,1,2,† Jiewei Liu,1,† Junyang Wang,1,2 Xiaoyan Li,1,2 Yongxia Huo,1 Yifan Li,1,2 Yixing Liu,1,2 Ming Li,1,2,3 Xiao Xiao1 and Xiong-Jian Luo1,2,3,4

†These authors contributed equally to this work.

Abstract

Genome-wide association studies showed that genetic variants at 2q33.1 were strongly associated with schizophrenia. However, potential causal variants in this locus and their roles in schizophrenia remain unknown. Here we identified two functional variants (rs796364 and rs281759) that disrupt CTCF, RAD21 and FOXP2 binding at 2q33.1. We systematically investigated the regulatory mechanisms of these two variants with serial experiments, including reporter gene assays and electrophoretic mobility shift assay (EMSA). Intriguingly, these two SNPs physically interacted with TYW5 and showed the most significant associations with TYW5 expression in human brain. Consistently, CRISPR-Cas9-mediated genome editing also confirmed the regulatory effect of these two SNPs on TYW5 expression. Additionally, expression analysis indicated that TYW5 was significantly up-regulated in brains of schizophrenia cases compared with controls, suggesting that rs796364 and rs281759 might confer schizophrenia risk by modulating TYW5 expression. We over-expressed TYW5 in mouse neural stem cells (NSCs) and rat primary neurons to mimic its up-regulation in schizophrenia cases, and found significant alterations in proliferation and differentiation of NSCs, as well as dendritic spine density following TYW5 overexpression, indicating its important roles in neurodevelopment and spine morphogenesis. Furthermore, we independently confirmed the association between rs796364 and schizophrenia in a Chinese cohort of 8,202 subjects. Finally, transcriptome analysis revealed that TYW5 affected schizophrenia-associated pathways. These lines of evidence consistently revealed that rs796364 and rs281759 might contribute to schizophrenia risk through regulating expression of TYW5, a gene whose expression
dysregulation affects two important schizophrenia pathophysiologic processes (i.e.,
neurodevelopment and dendritic spine formation).

Author affiliations:

1 Key Laboratory of Animal Models and Human Disease Mechanisms of the Chinese Academy of Sciences, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, Yunnan, 650223, China
2 Kunming College of Life Science, University of Chinese Academy of Sciences, Kunming, Yunnan, 650204, China
3 KIZ-CUHK Joint Laboratory of Bioresources and Molecular Research in Common Diseases, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, Yunnan, 650223, China
4 Center for Excellence in Animal Evolution and Genetics, Chinese Academy of Sciences, Kunming, Yunnan, 650223, China

Correspondence to: Xiong-Jian Luo
Key Laboratory of Animal Models and Human Disease Mechanisms, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, Yunnan 650204, China
E-mail: luoxiongjian@mail.kiz.ac.cn.

Running title: TYW5 is a novel schizophrenia risk gene

Keywords: schizophrenia; genome-wide association studies; functional genomics; rs796364; rs281759; TYW5

Abbreviations: GWASs = genome-wide association studies; PGC = Psychiatric Genomics Consortium; SNP = single nucleotide polymorphism; LD = linkage disequilibrium; ChIP-Seq = chromatin immunoprecipitation sequencing; PWM = position weight matrix; eQTL = expression quantitative trait loci; EMSA = electrophoretic mobility shift assay; NSCs = neural stem cells; TWAS = transcriptome-wide association study; DEGs = differentially expressed
Introduction

Schizophrenia is a severe and chronic psychiatric disorder with about 1% lifetime prevalence. It is characterized by the presence of delusions, hallucinations, lack of motivation, alogia, avolition and cognitive impairments.\textsuperscript{1,2} Twin studies estimated that schizophrenia has a heritability (broad-sense heritability, total phenotypic variance explained by genetic variation or genes) around 79-81\%,\textsuperscript{3,4} indicating the dominant role of genetic factors in schizophrenia pathogenesis. Genetic studies, especially recent genome-wide association studies (GWASs), have made considerable advances in identifying risk loci and dissecting the genetic architecture of schizophrenia.\textsuperscript{5-12} For example, the Psychiatric Genomics Consortium (PGC) and other research groups have carried out large-scale GWASs and reported multiple robustly associated loci for schizophrenia in the past decade.\textsuperscript{5-10} In addition, Lam et al. conducted a meta-analysis (56,418 schizophrenia cases and 78,818 controls) by combining results from populations of European and East Asian ancestries and identified 176 schizophrenia-associated loci.\textsuperscript{10} Despite the fact that these GWASs have provided important insights into genetic etiology of schizophrenia, the genetic and pathogenic mechanisms of the most reported risk loci in schizophrenia remain elusive. Considering that vast majority of schizophrenia risk variants identified by GWASs are located in non-coding regions, it is likely that these variants contribute to schizophrenia susceptibility through regulating gene expression rather than altering protein structure or function.\textsuperscript{6} Therefore, identifying the functional (or potential causal) variants at the reported schizophrenia risk loci and elucidating their regulatory mechanisms will provide important insights into the genetic mechanisms of schizophrenia.

Genetic variants at 2q33.1 (the index variant is chr2_200825237_I, a single base insertion that was firstly reported by PGC2 in 2014.\textsuperscript{6} For each locus identified by GWAS, a representative variant or single nucleotide polymorphism (SNP) (in most cases, the variant with the smallest \( P \) value in the locus) was selected as the index variant) showed robust associations with schizophrenia in previous GWASs (Supplementary Table 1).\textsuperscript{6,8,10} indicating that this
locus may harbor authentic risk variants for schizophrenia. However, due to the complexity of linkage disequilibrium (LD, the non-random association between genetic variants at two or more loci) (the SNP with the smallest association $P$ value is not necessarily the causal variant) and gene regulation (genetic variants may regulate distal genes by affecting 3D chromosomal structure), pinpointing causal variants in risk loci remains a major challenge in the post-GWAS era. Functional genomics analysis is believed to promote identification of functional variants. Through integrating chromatin immunoprecipitation sequencing (ChIP-Seq) and position weight matrix (PWM) data, we identified 132 functional SNPs from the reported schizophrenia risk loci in our previous study, including two functional SNPs (rs796364 and rs281759) at the 2q33.1 locus (Fig. 1). Of note, among the 132 TF binding-disrupting SNPs (identified in our previous study), rs796364 ($P = 9.41 \times 10^{-17}$) and rs281759 ($P = 9.25 \times 10^{-16}$) showed the second and third most associations with schizophrenia in a recent GWAS (Fig. 1 and Fig. 2A), strongly suggesting that two SNPs are true risk variants for schizophrenia. Besides, we explored the associations between these two SNPs and schizophrenia in East Asian populations, and found that these two SNPs were also associated with schizophrenia in East Asian populations (rs796364, $P = 5.66 \times 10^{-6}$; rs281759, $P = 2.88 \times 10^{-6}$). Finally, we previous showed that rs796364 and rs281759 were significantly associated with expression of $TYW5$ and $FTCDNL1$ in multiple human brain tissues, suggesting the potential functional consequences of these two SNPs. These lines of evidence suggest that rs796364 and rs281759 are potential causal variants for schizophrenia, we therefore focused on these two SNPs in this study.

Of note, our functional genomics showed that rs796364 disrupted CTCF and RAD21 binding, while rs281759 disrupted FOXP2 binding (Fig. 2B-C). The distance between rs796364 (located in the promoter of $FTCDNL1$ gene) and rs281759 (located in the intron 1 of $C2orf69$ gene) is 71,600 bp, and these two functional SNPs were in high LD ($r^2 = 0.97$). ChIP-Seq, DNase-Seq and histone modifications data (from ENCODE) indicated that these two functional SNPs were located in actively transcribed genomic regions (i.e., regulatory elements) in human neuroblastoma cell lines, with CTCF, RAD21 and FOXP2 binding (Fig. 2D-E). It is clear that these functional SNPs could provide pivotal genetic insights into etiology of schizophrenia. Nevertheless, the regulatory mechanisms of these two SNPs and the corresponding biological
implications of this locus in schizophrenia remain unknown.

To elucidate how these two SNPs exerted their regulatory effects, we conducted serial experiments in this study. We firstly validated the regulatory effects of these two SNPs with reporter gene assays, and our EMSA showed that these two SNPs affected binding affinity of TFs. We then performed expression quantitative trait loci (eQTL) analysis to identify the potential target gene (or genes) regulated by these two SNPs, and found that they showed the most significant associations with TYW5 expression in human brain. Further integrative analyses through integrating GWAS associations and brain eQTL data indicated that TYW5 was a schizophrenia risk gene. Consistent with eQTL and integrative analyses, we found that TYW5 was significantly dysregulated in brains of schizophrenia cases compared with controls. Therefore, these two SNPs may confer risk of schizophrenia by modulating TYW5 expression.

We also mimicked the effects of TYW5 dysregulation in schizophrenia and found that TYW5 affected proliferation and differentiation of neural stem cells (NSCs), indicating pivotal roles of TYW5 in neurodevelopment. Moreover, we found that TYW5 affected the density and ratio of dendritic spines, further supporting its vital roles in spine morphogenesis and maturity. Finally, we showed that rs796364 was significantly associated with schizophrenia in Chinese populations (with the same risk allele as in Europeans), suggesting that these two SNPs are likely authentic risk variants for schizophrenia. In summary, our study demonstrates that functional genetic variants rs796364 and rs281759 (at 2q33.1) may confer risk of schizophrenia by affecting binding of TFs and regulating expression of TYW5, a novel schizophrenia risk gene with important roles in neurodevelopment and dendritic morphogenesis.

Materials and methods

Functional genomics analysis

Details of the functional genomics approach have been described in our previous study. Briefly, by integrating ChIP-Seq and PWM data, we identified 132 functional SNPs that disrupted binding of 21 distinct TFs. In this study, we focused on the 2q33.1 risk locus (the functional SNPs in this locus are rs796364 and rs281759). The flowchart of the functional
genomics for the 2q33.1 locus is provided in Fig. 1.

**Linkage disequilibrium (LD) analysis**

To perform LD analysis, we used a well-established LD analysis database SNiPA, which calculates LD values using the data from the 1000 Genomes Project. The main parameters used were as follows: GRCh37 genome build, 1000 genome phase3 v5, and European populations. For further detailed information, please refer to the original SNiPA publication, and the SNiPA official website (https://snipa.helmholtz-muenchen.de/snipa3/).

**eQTL analysis**

To explore the potential target genes of the identified functional SNPs, we utilized two human brain eQTL datasets, CMC (CommonMind Consortium) and LIBD2 (Lieber Institute for Brain Development, Phase II) to perform eQTL analysis. Both eQTL datasets used tissues from human dorsolateral prefrontal cortex (DLPFC). The CMC dataset contains RNA sequencing (RNA-seq) data of 258 schizophrenia cases and 279 controls. The detailed information about the sample collection, RNA-seq data processing and eQTL calculation is provided in the original paper. The LIBD2 dataset performed brain eQTL analysis using 412 samples (175 schizophrenia cases and 237 healthy controls). Gene expression levels were quantified with RNA-seq and detailed information can be found in the original paper.

**Transcriptome-wide association study (TWAS)**

To identify genes whose expression change were associated with schizophrenia, we performed a TWAS analysis by using the FUSION software. Four predictive models, including LASSO, GBLUP, Elastic Net and BSLM were used. genome-wide association data and two SNP-expression weights (from the DLPFC) (i.e., CMC and LIBD2) were used for TWAS analysis. Bonferroni correction was applied to correct the TWAS P values. Details about the TWAS analysis was provided in the original paper, and FUSION website (http://gusevlab.org/projects/fusion/).
**TYW5 expression analysis in brains of schizophrenia cases and controls**

We used expression data from the PsychEncode (http://resource.psychencode.org/) to explore if TYW5 expression was dysregulated in the brains of schizophrenia cases compared with controls. Briefly, expression data (from the brain tissues) of 559 schizophrenia cases and 936 controls were measured through RNA-seq. The detailed information about PsychEncode was described in the original study.

**Cell culture**

Cell lines (HEK293T, SH-SY5Y and SK-N-SH) used in this study were kindly provided by the lab of Dr. Ming Li (Kunming institute of zoology, Chinese academy of sciences) (all cells were originally from ATCC) and cultured as previously described. Procedures for isolation and culture of mouse NSCs (Embryonic day 14) and rat cortical neurons (Embryonic day 18) have also been previously described, and detailed information about cell culture is listed in the Supplementary material. All cells were cultured at 37 °C with 95% air and 5% CO₂. Mycoplasma tests were performed periodically using specific PCR primers, and no mycoplasma contamination was detected in cells used in this study.

**Reporter gene assays (enhancer assay)**

Reporter gene assays were conducted as previously described. Briefly, about 500 bp genomic DNA containing rs796364 or rs281759 were amplified with PCR and inserted into pGL3-promoter (Promega, E1761) vector using specific primers (Supplementary Table 2). The pGL3-promoter (which contains promoter sequences but lacks of enhancer sequences, thus luciferase expression is moderate level) vector can be used to test the enhancer (or repressor) effect of target DNA sequence. If the target DNA sequence was inserted into pGL3-promoter vector and the inserted sequences has enhancer activity, the luciferase activity will significantly increase compared with pGL3-promoter vector (Supplementary Fig. 1). The constructed vectors were then used to transform DH5α competent cells. Vectors containing the alterative
allele were generated by PCR-mediated point mutation, with the using of the Golden Star T6 Super PCR Mix (TSINGKE, TSE101), corresponding mutation primers (Supplementary Table 2) and DpnI (NEB, R0176S). All constructed vectors were verified by Sanger sequencing. Detailed information about reporter gene assays were provided in the Supplementary methods.

**EMSA**

EMSA was used to investigate the interaction between DNA sequences (containing the test SNP) and binding proteins.\(^{28,29}\) We conducted EMSA assay as previously described.\(^{28}\) Briefly, 36 bp DNA probes containing the test SNPs were synthesized (Supplementary Table 3) and labeled with biotin (using Biotin Labeling Kit, Beyotime, GS008). The labeled probes were then incubated with nuclear protein extracts (NPEs) (from SH-SY5Y cells) (using Nuclear and Cytoplasmic Protein Extraction Kit, Beyotime, P0028). Chemiluminescent EMSA Kit (Beyotime, GS009) was used to detect if the probes interacted with nuclear extracts. The results of EMSA were quantified with image J software (https://imagej.nih.gov/ij/) by scanning gray value of each target band. And the relative gray value is defined as following: (gray value of target band) / (total gray value of the same lane).

**Chromatin interaction analysis**

We explored the long-range chromatin interaction between two regulatory variants (rs794364 and rs281759) and \(TYW5\) using the chromatin interaction data from 3DIV (A 3D-genome Interaction Viewer and database).\(^{30}\) Interaction data from human brain tissues and neuroblastoma cell lines were selected and visualized with 3DIV.

**Knockout of genomic regions containing rs796364 and rs281759**

We knocked out about 300 bp genomic DNA containing rs796364 or rs281759 using CRISPR-cas9 system as previously described.\(^{17}\) Briefly, two sgRNAs (one sgRNA located upstream while another sgRNA was downstream of the target SNP) were designed using the CRISPR sgRNA design tool (https://zlab.bio/guide-design-resources) (Supplementary Table 4). These
sgRNAs were then inserted into PX458M and EZ-guide XH vectors to assemble a knockout vector (PX458M-sgRNA1/2, with two sgRNAs insertion) by using double restriction enzymes XhoI (Thermo Scientific, FD0694) and HindIII (Thermo Scientific, FD0504). For HEK293T cells knockout assay, HEK293T cells were placed into 12-well plates (2.0 × 10^5/mL) containing 1 mL culture medium, and transfected with 1.5 µg knockout vectors (PX458M-sgRNA1/2) or control plasmids (PX458M) by using PEI. 72 hours post-transfection, the cells were collected for total RNA and DNA extraction. For knockout of rs796364 and rs281759 in SH-SY5Y and SK-N-SH cell lines, we inserted sgRNAs into lentiCRISPR v2 (Addgene, 52961), and the lentiviruses (expressing CAS9 and sgRNA1/2) packaged with pMD2G (Addgene, 12259) and psPAX2 (Addgene, 12260) were used to co-infect the SH-SY5Y and SK-N-SH cells. The infected cells were cultured with 1μg/mL puromycin for 12 days to select the stably infected cells. Genomic DNA was extracted for knockout efficiency verification (using PCR), and total RNA was used to detect TYW5 gene expression (using qPCR). As suggested by Vandesompele et al., different internal control genes were used for HEK293T (ACTB and GAPDH), SK-N-SH and SH-SY5Y (ACTB and B2M) cell lines. The geometric averages of the two internal control genes were used to normalize the qPCR data, and expression level was determined by using the2^-ΔΔCt approach. The primers used for PCR and qPCR in this study are provided in Supplementary Table 4 and 5.

Knockdown of CTCF, RAD21 and FOXP2

In order to test whether TFs (CTCF, RAD21 and FOXP2) regulated TYW5 gene expression, we knocked down these TFs using shRNAs designed using ThermoFisher shRNA design tool (http://rnaidesigner.thermofisher.com/rnaexpress/). The targeting sequences were as follows: CTCF-shRNA, 5'-GCGAAAGCAGCATTCCTATAT-3', RAD21-shRNA, 5'-GCCATTACTTATACCTGGAAGA-3', FOXP2-shRNA, 5'-GCAAACAAGTGGATTGAAATC-3'. The synthesized shRNA oligos (Supplementary Table 6) were inserted into the pLKO.1-EGFP-Puro vector at double restriction sites (AgeI and EcoRI), and the recombinant plasmids were verified by Sanger sequencing. Lentiviral viruses produced by control plasmids and target shRNA vectors were used to infect HEK293T,
SH-SY5Y and SK-N-SH cell lines. After 72-hour infection, 1-2 µg/mL puromycin was used to kill uninfected cells. qPCR was used to quantify expression of target genes (i.e., CTCF, RAD21, FOXP2, TYW5) in control and TFs knockdown cells. The qPCR primers used are listed in Supplementary Table 5.

**TYW5 overexpression in mouse NSCs**

Considering that TYW5 was up-regulated in schizophrenia cases, we evaluated the effect of TYW5 overexpression on neurodevelopment using mouse NSCs. Briefly, the lentiviruses (produced by PCDH-TYW5-6×his and control vectors) were used to infect mouse NSC and puromycin (1-2 µg/mL, increased gradually for a week) was used to select NSCs stably expressing human TYW5 gene. RT-PCR and Western blot were used to validate overexpression of TYW5. Primers used for RT-PCR are listed in Supplementary Table 5. Antibodies used for Western blot were as follows: primary antibodies were mouse anti-his (1:500, Proteintech, 66005-1-Ig) and mouse anti-beta actin antibody (1:1000, Proteintech, 66009-1-Ig), and secondary antibodies was HRP-labeled goat anti-mouse (1:1000, Beyotime, A0216).

**Proliferation assays**

To investigate if TYW5 overexpression affected proliferation of NSCs, we used the Cell Counting Kit-8 (CCK-8) and performed EdU incorporation assays to test proliferation capabilities. For CCK-8 assay, $1 \times 10^4$ cells were plated into 96-well plates (pre-coated with 5 µg/mL laminin). After culturing for 8, 42 and 66 hours, the absorbance at 450 nm wavelength was measured according to the manufacturer’s instructions. For EdU assay, $1 \times 10^5$ NSCs were plated into 48-well plates (pre-coated with 5 µg/mL laminin). After culturing for 48 h, 20 µM EdU (RIBO-Bio, C00054) was added to label dividing cells. After incubation for 45 min, Cell-Light EdU Apollo567 In Vitro Kit (RIBO-Bio, C10310-1) was used to detect EdU labeled cells. Briefly, cells were washed by PBS, fixed with 4% PFA for 15min, permeabilized with 0.3% PBST for 20 minutes, and then incubated with 100µL 1×Apollo® Staining reaction solution for 1 hour. Laser scanning confocal microscope (OLYMPUS, FV1000) was used to acquire images. Twenty-five images in each group (five images in each well) were used to
analyze the ratio of EdU positive cells (scaled to DAPI) using the Image J software.

**Differentiation of NSCs**

To investigate if *TYW5* overexpression affected differentiation of NSCs, we conducted differentiation assay as previously described. In brief, 3 × 10^5 cells (per well) in 24 well plates and 2 × 10^6 cells (per well) in 6-well plates pre-coated with 5 µg/mL laminin were cultured in proliferation mediums for 12 hours. The proliferation mediums were then replaced with differentiation mediums (DMEM/F12, containing 1% penicillin/streptavidin, 1% N2 supplement, 1% B27 supplement, and 2 µg/mL heparin). Three days post differentiation, immunofluorescence staining was carried out to measure the proportions of astrocytes (GFAP positive) and neurons (MAP2 positive) in all cells, and qPCR were performed to detect expression of *Gfap*, *Map2*, *Tubb3* (expressing TUJ1 protein, a marker for newly generated immature post-mitotic neurons) and *Cldn11* (expressing O4 protein, a marker for oligodendroglia cells). The primary antibodies used were rabbit anti-GFAP (1:1000, Sigma, G9269) and rabbit anti-MAP2 (1:200, Millipore, AB5622). The secondary antibodies were goat anti-rabbit 555 (1:500, Invitrogen, A32732). The images were captured by laser scanning confocal microscopes (Olympus, FV1000) with the same scanning parameters. Fifteen images in each group (five images from each biological replicate) were used to analyze the ratio of GFAP and MAP2 positive cells (scale to DAPI) using Image J software. Primers used for qPCR were provided in Supplementary Table 5. Besides, Western-blot also used to detect the GFAP and MAP2 protein level at three different time points (differentiation of mouse NSCs for 24, 48 and 72 hours). Primary antibodies used were rabbit anti-GFAP (1:1000, Sigma, G9269) and rabbit anti-MAP2 (1:500, Millipore, AB5622) and mouse anti-GAPDH (1:2000, Proteintech, 60004-1-Ig). Secondary antibodies were HRP-labeled goat anti-rabbit (1:1000, Beyotime, A0208) and HRP-labeled goat anti-mouse (1:1000, Beyotime, A0216).

**Morphological analysis of rat cortical neurons**

Previous studies have shown dysregulation of dendritic spines in schizophrenia. To further investigate if *TYW5* overexpression affected dendritic spines, we carried out morphological...
analysis of dendritic spines as previously described. Briefly, after culturing the isolated rat primary neurons for 14-15 days in vitro, 3μg pCAG-TYW5-1×flag vectors and 1μg Venus plasmids (expressing GFP protein) were co-transfected into rat cortical neurons using Lipofectamine 3000 (Invitrogen, L3000015). 72 hours post-transfection, immunofluorescence staining was performed. Primary antibodies were rabbit anti-flag antibody (1:150, Sigma, F7425-2MG) and chicken anti-GFP antibody (1:1000, Abcam, ab13970). Secondary antibodies were goat anti-rabbit 555 (1:500, Invitrogen, A32732) and donkey anti-chicken Cy™2 (1:200, Jackson Immuno Research, 703-225-155). Laser scanning confocal microscope (Leica, TCS SP8 X) was used to take images for neurons co-expressing GFP and TYW5-flag fusion protein, with the use of same scanning parameters. Images of at least 20 independent neurons from each group were captured for morphological analyses. Two secondary or tertiary dendrites of each neuron (at least total length of 100 μm) were selected to quantify spine numbers and classify spines (thin, long thin, stubby, mushroom, filopodia and branched). The Image J software was used for the analysis. The obtained data were compared between two different groups using two-tailed Student’s t-test.

**Transcriptome analysis**

To explore genes and biological processes regulated by TYW5, we conducted RNA-seq in control mouse NSCs and mouse NSCs stably overexpressing TYW5. RNA-seq was performed by ANOROAD Company (Beijing), with the using of HiSeq/X-ten Illumina platform. Transcriptome analysis was performed as follows: First, clean reads were mapped to the mouse genome (GRCm38) by using bowtie2 (v2.3.5) with default parameters. The obtained sam files were then converted to bam and sorted by SAMtools (v1.9). And HTSeq (v0.12.4) was used to quantify read counts for each bam files. Differentially expressed genes (DEGs) were identified using the DEseq2 package in R (v3.5.1), and P values were corrected using the Benjamini-Hochberg method. The thresholds of DEGs were set as $| \log_2 \text{(foldchange)} | > 0.5$ and $P_{\text{adj}} < 0.01$. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses were performed with DAVID (v6.8) (https://david.ncifcrf.gov).
Data availability statement

The data that support the findings of this study are available in Schizophrenia Working Group of the Psychiatric Genomics Consortium (PGC), GTEx, LIBD, CMC, ENCODE, 3DIV, PsychEncode, PROMO, AliBaba2.1, Brennand Lab RNA-seq, CoDex Viewer, UCSC Cell Browser and GEO (GSE106589). These data were derived from the following resources available in the public domain: PGC, https://www.med.unc.edu/pgc/; GTEx, https://www.gtexportal.org/; LIBD, http://eqtl.brainseq.org/; ENCODE, https://www.encodeproject.org/; 3DIV, https://www.kobic.kr/3div/; PsychEncode, http://resource.psychencode.org/; PROMO software website, http://alggen.lsi.upc.es/cgi-bin/promo_v3/promo/promoinit.cgi?dirDB=TF_8.3; AliBaba2.1 website, http://gene-regulation.com/pub/programs/alibaba2/index.html; Brennand Lab RNA-seq website, https://schroden.shinyapps.io/BrennandLab-ExpressionApp-limited/; CoDex Viewer website, http://solo.bmap.ucla.edu/shiny/webapp/; UCSC Cell Browser, https://cells.ucsc.edu/; GEO (GSE106589), https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE106589.

Results

Reporter gene assays validated the regulatory effects of rs796364 and rs281759

Dual-luciferase reporter gene assays can be used to evaluate the regulatory effect of target DNA sequences.45 By inserting target DNA sequences into regulatory region (promoter or enhancer) of the reporter gene (firefly luciferase) vector, the regulatory effect of target DNA sequences can be detected by measuring the activity of firefly luciferase activity. Renilla luciferase was used to control the transfection efficiency and the normalized ratio (i.e., firefly luciferase/renilla luciferase) reflects the regulatory effect of target DNA sequences (Supplementary Fig. 1). To validate the regulatory effects of rs796364 and rs281759, we
conducted reporter gene assays in neuroblastoma cell lines from human brain (SK-N-SH and SH-SY5Y) and HEK293T (a cell line from human kidney). For rs796364, the A and C allele of rs796364 did not show significant difference in luciferase activity in HEK293T (Fig. 3A). However, the A allele conferred significant higher luciferase activity than the C allele in SK-N-SH and SH-SY5Y cells (Fig. 3B, C), suggesting the cell-type specific regulatory effect of rs796364 (i.e., rs796364 has regulatory effect in neuroblastoma cell lines, but not in HEK293T). For rs281759, the T allele exhibited significant higher luciferase activity than C allele in all three tested cell lines (Fig. 3D-F). These results confirmed the regulatory effects of rs796364 and rs281759, and suggesting that these two functional SNPs regulate expression level of their target gene (or genes) by modulating activity of enhancer elements.

**EMSA showed that rs796364 and rs281759 affected binding affinities of TFs**

EMSA is widely used to evaluate the interaction between target DNA sequences and transcription factors.\(^{28,29}\) If target DNA sequence binds specific transcription factor(s), the protein-probe complexes migrate more slowly than the corresponding free probes in electrophoresis. To investigate if rs796364 and rs281759 affect binding of TFs, we conducted EMSA. EMSA results showed that these two SNPs significantly altered binding affinities to nuclear proteins extracted from SH-SY5Y cells (which contained abundant proteins and TFs) (Fig. 3G-J). The A allele of rs796364 showed higher affinity to TFs compared with the C allele (Fig. 3G, H), and the T allele of rs281759 showed higher affinity to TFs compared with the C allele (Fig. 3I, J). Further bioinformatic analysis suggested that different alleles of these two SNPs affect binding of several TFs (Supplementary Fig. 2 and Supplementary Table 7). These results indicated that both of these two functional SNPs (rs796364 and rs281759) may affect TFs binding. Interestingly, results of reporter gene assay and EMAS showed that the alleles associated with higher affinity (rs796364-A and rs281759-T) to nuclear protein extractions exhibited higher luciferase activity.
Rs796364 and rs281759 showed the most significant associations with *TYW5* expression in human brain

Our reporter gene assays and EMSA verified the regulatory effects of rs796364 and rs281759, indicating that these two SNPs might confer schizophrenia risk by regulating gene expression. To further identify the potential target gene (or genes) regulated by these two SNPs, we examined their correlations with expression of genes in the CMC and LIBD2 brain eQTL datasets.\(^{22,23}\) We found that these two SNPs were significantly associated with mRNA levels of 4 genes (within 1 Mb distance) in CMC datasets, including *C2orf47*, *FTCDNL1*, *SATB2* and *TYW5* (Supplementary Fig. 3A, B). Intriguingly, these two SNPs showed the most significant associations with *TYW5* expression in both CMC (\(P_{\text{rs796364}} = 9.25 \times 10^{-5}, P_{\text{rs281759}} = 1.47 \times 10^{-4}\)) and LIBD2 (\(P_{\text{rs796364}} = 4.51 \times 10^{-23}, P_{\text{rs281759}} = 3.83 \times 10^{-25}\)) datasets (Fig. 4A-D), implicating that they likely conferred schizophrenia risk through regulating *TYW5* expression.

CRISPR-Cas9-mediated genome editing validated the regulatory effects of rs796364 and rs281759 on *TYW5*

As our eQTL analysis implicated that these two functional SNPs regulate *TYW5* expression, we sought to validate if they indeed exerted regulatory effects on this gene. Considering that both SNPs located downstream of *TYW5* (Fig. 2A) and reporter gene assays revealed strong enhancer activity of the genomic sequences containing rs796364 in neuroblastoma cells (Fig. 3B, C), it is likely that rs796364 and rs281759 regulate *TYW5* through affecting enhancer activity. We thus knocked out about 300 bp genomic sequences surrounding these two SNPs in three cell lines (HEK293T, SH-SY5Y and SK-N-SH), and found that deletion of these two SNPs resulted in significant down-regulation of *TYW5* in all three cell lines (\(P < 0.05\)) (Fig. 4E, F and Supplementary Fig. 4). These data indicated that these two SNPs are located in the enhancer elements of *TYW5* and it is likely that they regulate *TYW5* expression by altering enhancer activity. Interestingly, chromatin interaction analyses in human neuroblastoma cell line SK-N-DZ and brain tissues (using 3DIV) showed that these two SNPs physically interact with *TYW5* gene in the DLPFC, hippocampus and SK-N-DZ cell line (Supplementary Fig. 5-
suggesting rs796364 and rs281759 may regulate *TYW5* expression by affecting chromatin interactions.

**Knock-down of corresponding TFs affected *TYW5* expression**

Our functional genomics analysis revealed that rs796364 and rs281759 disrupted binding of CTCF, RAD21 and FOXP2, providing molecular explanations for their regulatory impact on *TYW5* expression. To further examine the regulatory effects of these TFs on *TYW5* expression, we knocked down *CTCF, RAD21* and *FOXP2* expression in HEK293T, SH-SY5Y and SK-N-SH cells (using shRNAs). qPCR results showed that RAD21 knockdown resulted in significant down-regulation of *TYW5* (*P* < 0.05), while CTCF knockdown led to significant up-regulation of *TYW5* (*P* < 0.05) in all three tested cell lines (Fig. 4G-P). As *FOXP2* expression level in SH-SY5Y and SK-N-SH cell lines was not detected by RT-PCR (Supplementary Fig. 9), we did not knocked-down *FOXP2* in SH-SY5Y and SK-N-SH cell lines. Taken together, these results indicated that rs796364 and rs281759 regulate *TYW5* expression through affecting TFs binding affinity.

**Dysregulation of *TYW5* in schizophrenia cases**

We provided convergent lines of evidence (including functional genomics, eQTL analysis, Cas9-mediated genome editing, TFs knockdown) suggesting that functional SNPs rs796364 and rs281759 might confer schizophrenia risk through regulating *TYW5* expression. It is therefore necessary to examine whether expression perturbation of *TYW5* was involved in schizophrenia. We thus compared *TYW5* expression in brains of schizophrenia cases *versus* that in controls using expression data from the PsychEncode.\(^{25}\) Consistent with our above hypothesis, we found that *TYW5* expression was significantly up-regulated in brains of schizophrenia cases compared with controls (*P* = 8.20 × 10\(^{-4}\)) (Fig. 4Q), indicating a potentially pivotal role of *TYW5* dysregulation in schizophrenia.

**Overexpression of *TYW5* inhibits proliferation of mouse NSCs**

Appropriate proliferation and differentiation of NSCs are necessary for neurodevelopment,\(^{46}\)
and previous studies have shown that dysfunction of neurodevelopment might play important roles in the pathogenesis of schizophrenia.\textsuperscript{47,48} To further characterize the roles of *TYW5* in schizophrenia, we utilized the NSCs model (a model that was widely used to investigate the role of schizophrenia risk genes in neurodevelopment).\textsuperscript{49,50} Of note, the neurodevelopmental hypothesis of schizophrenia posits that schizophrenia risk genes may confer risk of this disorder by affecting brain development.\textsuperscript{47,48} As *TYW5* was significantly up-regulated in schizophrenia patients, we conducted overexpression experiments to mimic the effects of *TYW5* up-regulation. Cells from the ventricular zone of embryonic mouse brains (E14) were isolated and the identity of these cells were validated with three well-characterized NSC markers (SOX2, NESTIN and PAX6) (Supplementary Fig. 10). The NSCs were then infected with the packaged viral particles and puromycin (1 μg/mL) was added to select the stably infected cells for two weeks. RT-PCR and Western-blot confirmed stable overexpression of human *TYW5* in the infected mouse NSCs (Fig. 5A, B). We then conducted EdU incorporation and CCK-8 assays to assess the effect of *TYW5* overexpression on proliferation of these cells. Both assays showed that *TYW5* overexpression significantly inhibited proliferation of mouse NSCs (Fig. 5C-E). Taken together, *TYW5* overexpression affected proliferation of NSCs and thereby modulated neurodevelopment.

**Overexpression of *TYW5* affects differentiation of mouse NSCs**

In addition to proliferation, differentiation is also important for neurodevelopment.\textsuperscript{46} We thus sought to investigate if *TYW5* had a role in neuronal differentiation. NSCs can differentiate into glia cells and neurons, two major cell types in the human brain. GFAP is a marker for astrocytes, MAP2 is a marker for mature neurons, TUJ1 (encoded by *Tubb3*) is a marker for the immature post-mitotic neurons, and the oligodendrocytes can be labeled with O4 (encoded by *Cldn11*). These cell types are generated during the differentiation process of the NSCs, and these markers can be used to investigate the differentiation process of NSCs.\textsuperscript{51,52} After differentiation of NSCs for three days, GFAP and MAP2 positive cells were counted. We found that the ratio of GFAP (a marker for glia cells) (but not MAP2, a marker for neurons) positive cells was significantly decreased in cells overexpressing *TYW5* compared with control cells (*P* < 0.05) (Fig. 6A-D),
indicating that \textit{TYW5} overexpression impaired differentiation of glia cells. In addition, we also measured the mRNA level of \textit{Gfap, Map2, Tubb3} and \textit{Cldn11} in \textit{TYW5} overexpression and control cells, with the use of \textit{Actb} and \textit{Ywhaz} as endogenous reference genes (Supplementary Table 5). We found that expression of \textit{Tubb3} and \textit{Cldn11} (but not \textit{Gfap, Map2}) were significantly down-regulated in \textit{TYW5} overexpression cells compared with controls ($P < 0.05$) (Supplementary Fig. 11). In order to further validate the effect of \textit{TYW5} overexpressing on differentiation of NSCs, we used western-blot to detect the GFAP and MAP2 protein expression after differentiation of NSCs for 24, 48, and 72 hours. After differentiation of NSCs for 24 hours, GFAP was significantly down-regulated in \textit{TYW5} overexpressing groups compared with controls. After differentiation of NSCs for 48 hours, MAP2 was significantly upregulated in \textit{TYW5} overexpressing groups compared with controls. However, GFAP and MAP2 did not show significant differences after differentiation of NSCs for 72 hours (Fig. 6E-H). A possible explanation for this is the degradation of GFAP and MAP2 proteins (Fig. 6E). These data demonstrated that \textit{TYW5} overexpression may affect differentiation of mouse NSCs.

**\textit{TYW5} regulates schizophrenia-related pathways**

Our above data showed that \textit{TYW5} overexpression affected proliferation and differentiation of NSCs. Nevertheless, the underlying biological processes and signaling pathways remain unclear. We thus used the NSCs that stably overexpressing \textit{TYW5} and control NSCs (cultured in proliferation medium) to perform transcriptome analysis through RNA-seq. A total of 304 DEGs (183 up-regulated, 121 down-regulated) were identified (Fig. 7A, B and Supplementary Table 8). GO analysis showed that the DEGs were significantly enriched in multiple essential physiological processes including DNA replication, cell adhesion and cell cycle ($P_{\text{Benjamini}} < 0.05$) (Fig. 7C). Further KEGG analysis showed that the DEGs were significantly enriched in schizophrenia-associated signaling pathways, including ECM-receptor interaction, focal adhesion and PI3K-Akt ($P_{\text{Benjamini}} < 0.05$) (Fig. 7D). Of note, both GO and KEGG analyses indicated that the DEGs showed the most significant enrichment in DNA replication related pathways. Collectively, these results suggested that \textit{TYW5} may contribute to schizophrenia susceptibility through regulating ECM-receptor interaction, focal
adhesion and PI3K-Akt pathways.\textsuperscript{53-55}

**Overexpression of *TYW5* affects the morphology and density of dendritic spines**

In addition to the abnormality of neurodevelopment, density of dendritic spines was also significantly reduced in schizophrenia cases,\textsuperscript{33-35} indicating that formation of dendritic spine may also participate in schizophrenia pathogenesis. We thus investigated the effects of *TYW5* overexpression on dendritic spine formation using rat primary cortical neurons. We found that *TYW5* overexpression affected the dendritic spine density (Fig. 8A-C). The density and ratio of stubby spines were significantly decreased, while the density and ratio of thin spines were significantly increased in *TYW5* overexpressing neurons compared with controls ($P < 0.05$) (Fig. 8A-C). However, the density and ratio of mushroom spines, others spines (including long thin spines, filopodia spines and branched spines) and total spines were not affected ($P > 0.05$). We further analyzed the long thin spines, filopodia spines and branched spines and found that the density and ratio of branched spines were significantly decreased in *TYW5* overexpressing neurons compared with controls ($P < 0.05$) (Supplementary Fig. 12E, F). Taken together, these results showed that overexpression of *TYW5* could affect the formation of rat neurons dendritic spine, proving potentially important roles of *TYW5* in formation of dendritic spines.

**Discussion**

Identifying the functional (or causal) risk variants at the reported risk loci and elucidating their roles in schizophrenia is an important and necessary step to understand the genetic mechanisms and pathogenesis of schizophrenia in the post-GWAS era. Though recent studies have demonstrated potential regulatory mechanisms of some schizophrenia risk variants,\textsuperscript{16,25,56,57} the functional variants and their roles in schizophrenia for most of the GWAS risk loci remain largely unknown. In this study, we systematically characterized the regulatory mechanisms of risk variants rs796364 and rs281759, two functional SNPs identified in our previous functional genomics study.\textsuperscript{16} We have provided convergent lines of evidence to support the regulatory
effects of rs796364 and rs281759. First, we validated the regulatory effects of these two functional SNPs with reporter gene assays. Second, our EMSA results showed that they could affect binding affinity of TFs. Third, PAINTOR fine-mapping also suggesting that rs796364 is a potential causal variant (Posterior probability of causality = 0.66) (Supplementary Table 9). Fourth, we conducted the eQTL analysis and have found the potential target gene regulated by these two SNPs, TYW5, whose mRNA expression was significantly associated with these two SNPs in human brain. Fifth, our TWAS integrative analyses showed that TYW5 is a schizophrenia risk gene whose expression change might have a role in this disorder (TWAS P(CMC) = 1.34 × 10^{-11}, TWAS P(LIBD2) = 1.50 × 10^{-12}) (GWAS associations and two brain eQTL datasets (CMC and LIBD2) were used for TWAS integrative analyses) (Supplementary Table 10). We also further confirmed the regulatory effect of rs796364 and rs281759 on TYW5 expression using CRISPR-Cas9-mediated genome editing, and defined the transcription factors facilitating such regulatory effect. Therefore, rs796364 and rs281759 likely confer schizophrenia risk by affecting TYW5 expression. Consistent with this, we found that these two regulatory variants physically interacted with TYW5 in the DLPFC, hippocampus and neuroblastoma cells of the human brain. These lines data indicate that these two functional SNPs regulate TYW5 expression through altering transcription factors binding and long-range chromatin interaction. Intriguingly, expression analysis showed that TYW5 was significantly up-regulated in brains of schizophrenia cases (Fig. 4Q), highlighting its pivotal roles in schizophrenia pathogenesis.

In addition, we also independently confirmed that rs796364 was significantly associated with schizophrenia in non-Europeans (i.e., Chinese population) (P = 9.68 × 10^{-4}) (data was from a recent study of our lab), with the same risk allele as in Europeans. Meta-analysis (a total of 59,911 cases and 83,527 controls) using published data indicated that rs796364 was strongly associated with schizophrenia (P = 1.14 × 10^{-19}). This successful genetic validation in independent population provided robust evidence for the involvement of these two regulatory variants in schizophrenia, and further supporting the potential functional consequences of this SNP. Considering that cognitive impairment is a core symptom of schizophrenia and previous studies have showed that schizophrenia risk variants were also associated with cognitive performance, if rs796364 represents an authentic risk variant for schizophrenia, it may
also be associated with cognitive performance. Examination of association between rs796364 and cognitive performance using published data showed that rs796364 was also associated with cognitive performances \( P_{\text{Intelligence}} = 3.91 \times 10^{-4}, P_{RT} = 1.42 \times 10^{-3} \) and \( P_{VNR} = 9.00 \times 10^{-3} \), and the risk allele predicted poorer cognitive performance. The successful validation in genetic independent populations and its association with cognitive performance provides extra support for the involvement of rs796364 in schizophrenia.

We noted that rs796364 is located in the promoter of \( \text{FTCDNL1} \) gene (Fig. 2). We thus explored \( \text{FTCDNL1} \) expression in different cell types of the human brain. \( \text{FTCDNL1} \) expression in different neural cells is much lower than \( \text{TYW5} \) (Supplementary Fig. 13 A-B). We also examined \( \text{FTCDNL1} \) and \( \text{TYW5} \) expression in Cortical Development Expression Viewer and UCSC Cell Browser, and observed similar results (Supplementary Fig. 14 and Supplementary Fig. 15 A, B). Intriguingly, \( \text{TYW5} \) has a higher expression level in neural progenitor cells (NPCs) and 6-week-old forebrain neurons (derived from human induced pluripotent stem cells (hiPSC) of normal subject) (Supplementary Fig. 16). Of note, our RT-PCR showed no obvious expression of \( \text{FTCDNL1} \) in HEK293T, SK-N-SH and SH-SY5Y cell lines (Supplementary Fig. 9). Finally, spatio-temporal expression data showed that \( \text{TYW5} \) is stably and highly expressed in different regions of the human brain across whole life (Supplementary Fig. 17), implying the pivotal role of \( \text{TYW5} \) in human brain. These expression data suggest that rs796364 and rs281759 might confer schizophrenia risk by modulating \( \text{TYW5} \) (i.e., compared with \( \text{FTCDNL1} \), it is more likely that \( \text{TYW5} \) represents the potential risk gene at this locus).

In addition to elucidating the regulatory mechanisms of rs796364 and rs281759, we also provided further evidence to support the involvement of \( \text{TYW5} \) in schizophrenia pathogenesis. \( \text{TYW5} \) was up-regulated in brains of schizophrenia cases and overexpression of \( \text{TYW5} \) affected proliferation and differentiation of NSCs, indicating that \( \text{TYW5} \) has a role in neurodevelopment. Intriguingly, we also found that \( \text{TYW5} \) regulated the density of dendritic spines, an important neuronal structure that was frequently found to be dysregulated in schizophrenics. We also checked the associations between the DEGs (between \( \text{TYW5} \)-overexpression and control NSCs) and schizophrenia (detailed information were provided in Supplementary Material). Our DEGs (including \text{Snap91, Ndrg4, Ier3, Ip6k3}) showed significant associations with
schizophrenia, suggesting that TYW5 may confer risk of schizophrenia by regulating these genes. Besides, our transcriptome analysis showed that TYW5 regulated schizophrenia-associated pathways, including ECM-receptor interaction, Focal adhesion, PI3K-Akt signaling pathway.\textsuperscript{53-55} Interestingly, a previous TWAS (which used fetal brain eQTL) also suggested that TYW5 was a schizophrenia risk gene (TWAS \( P = 4.16 \times 10^{-9} \)) whose expression alteration in the early brain developmental stage might have a role in schizophrenia.\textsuperscript{69} Finally, association exploration between TYW5 and other psychiatric disorders (including bipolar disorder, depression and attention deficit and hyperactivity) showed suggestive association between TYW5 and bipolar disorder (rs73066802, \( P = 2.61 \times 10^{-5} \)) (Supplementary Fig. 18).\textsuperscript{70-72} These consistent and convergent lines of evidence demonstrated that TYW5 might confer schizophrenia risk through affecting neurodevelopment and schizophrenia-associated pathways.

TYW5 encodes TRNA-YW Synthesizing Protein 5 (a Jumonji C (JmjC)-domain-containing protein) (JMJD), a tRNA modification enzyme which forms the OHyW nucleoside by carbon hydroxylation by using 2-oxoglutarate (2-OG) and Fe\(^{2+}\) ion as cofactors.\textsuperscript{73,74} JMJD proteins usually act as epigenetic regulators. However, TYW5 acts as RNA hydroxylases.\textsuperscript{73,74} TYW5 is involved in Wybutosine biosynthesis pathway, with six enzymes (including TRMT5 and TYW1-5) work sequentially to transfer the original nucleoside in position 37 of tRNAPhe into hydroxy wybutosine (OHyW).\textsuperscript{73,75} By homo-dimerization, a large and positively charged patch is formed for tRNA binding. TYW5 expression is significantly decreased under iron depletion.\textsuperscript{76} Interestingly, the levels of trace element iron in blood of schizophrenia cases were lower than that of normal people.\textsuperscript{77} To date, the function of TYW5 remains largely unknown and there is no study to investigate its role in brain development and schizophrenia. Interestingly, we found that TYW5 protein is mainly located in the cytoplasm and nucleus of five different cells (HEK293T, SH-SY5Y, SK-N-SH, mouse NSCs and rat neurons) (Supplementary Fig. 19-22), and for the first time we showed its pivotal role in neurodevelopment and spine morphogenesis (i.e., dysregulation of TYW5 affected proliferation and differentiation of NSCs, and spine density of neurons). Though the pathophysiology of schizophrenia remains elusive, accumulating data supports the involvement of aberrant neurodevelopment.\textsuperscript{47,48} Our findings highlight TYW5 as an important gene in appropriate neurodevelopment and schizophrenia
susceptibility.

Despite the above implications of this study for the genetic basis of schizophrenia, there are several limitations to be acknowledged. First, though we showed that rs796364 and rs281759 could regulate TYW5 expression (likely through affecting binding of CTCF, RAD21 and FOXP2), the exact mechanisms still need to be elucidated, probably through precise single-base editing (using CRISPR-Cas9 system) of these SNPs. Second, in addition to TYW5, chromatin interaction analysis also showed interactions between rs796364 and rs281759 and other genes (including C2orf69 and MAIP1). And gene expression analysis also showed significant down-regulation of FTCDNL1 in brains of schizophrenia cases compared with controls ($P = 5.92 \times 10^{-4}$, FDR = $6.31 \times 10^{-3}$) in PsychENCODE. Therefore, we could not exclude the possibility that these two SNPs may confer schizophrenia risk by modulating other genes (rather than TYW5) completely. Third, though PAINTOR fine-mapping supported the causality of rs796364, we noticed that another SNP rs2949006 also has high posterior probability (0.838) for causality (Supplementary Table 9). eQTL analysis showed that rs2949006 was significantly associated with FTCDNL1 ($P = 9.31 \times 10^{-14}$) and TYW5 expression ($P = 2.52 \times 10^{-8}$) in human brain. The strong association between rs2949006 and FTCDNL1 suggested that FTCDNL1 may also be a potential risk gene for schizophrenia. Furthermore, FTCDNL1 was significantly down-regulated in schizophrenia cases compared with controls, further supporting the potential role of FTCDNL1 in schizophrenia. These data suggest that risk variants in this region may confer risk of schizophrenia by modulating FTCDNL1. More work is needed to investigate if FTCDNL1 is an authentic risk gene for schizophrenia. Fourth, our results showed that TYW5 overexpression affected the proliferation and differentiation of NSCs, and the formation of dendritic spines in vitro, relevant physiological processes linking these alterations to schizophrenia pathogenesis remain to be illuminated using animal models. Fifth, we used the mouse NSC model to explore the role of TYW5 in neurodevelopment. The rationale behind the use of NSC are as follows: (1) The neurodevelopmental hypothesis of schizophrenia posits that schizophrenia risk genes may confer risk of this disorder by affecting brain development. Consistent with this hypothesis, previous studies also showed the pivotal roles of schizophrenia risk genes in early neurodevelopment. Thus, we think that mouse NSCs may represent a feasible model to explore the function of schizophrenia risk genes;
(2) Compared with human NSCs, culturing (and manipulating) of mouse NSCs is much easier, and growth rate of mouse NSCs is also faster; (3) Immature/mature neurons may also be a good model. Nevertheless, isolating and culturing of immature/mature neurons remain challenge currently. More work (including using of human NSCs and immature/mature neurons) is needed to elucidate the role of TYW5 in schizophrenia pathogenesis. Sixth, considering the polygenicity nature of schizophrenia, the clinical relevance of this study is limited due to the tiny influence of each individual risk locus on the risk of schizophrenia. Further work is needed to explore the potential clinical relevance of this study in schizophrenia. Finally, while our study provides insights into the biological relevance of TYW5 in schizophrenia, further characterization of its functionality and underlying mechanisms by which this gene participate in the disorder is needed.

In summary, we systematically characterized the regulatory mechanisms of rs796364 and rs281759. We showed that these two functional SNPs confer schizophrenia risk by regulating expression of TYW5, a previous uncharacterized gene with important roles in neurodevelopment and spine morphogenesis (Supplementary Fig. 23).

Acknowledgements

One of the Brain eQTL datasets used in this study were generated as part of the CommonMind Consortium supported by funding from Takeda Pharmaceuticals Company Limited, F. Hoffman-La Roche Ltd and NIH grants R01MH085542, R01MH093725, P50MH066392, P50MH080405, R01MH097276, RO1-MH-075916, P50M096891, P50MH084053S1, R37MH057881 and R37MH057881S1, HHSN271201300031C, AG02219, AG05138 and MH06692. Brain tissue for the study was obtained from the following brain bank collections: the Mount Sinai NIH Brain and Tissue Repository, the University of Pennsylvania Alzheimer’s Disease Core Center, the University of Pittsburgh NeuroBioBank and Brain and Tissue Repositories and the NIMH Human Brain Collection Core. CMC Leadership: Pamela Sklar, Joseph Buxbaum (Icahn School of Medicine at Mount Sinai), Bernie Devlin, David Lewis (University of Pittsburgh), Raquel Gur, Chang-Gyu Hahn (University of Pennsylvania), Keisuke Hirai, Hiroyoshi Toyoshiba (Takeda Pharmaceuticals Company...
Limited), Enrico Domenici, Laurent Essioux (F. Hoffman-La Roche Ltd), Lara Mangravite, Mette Peters (Sage Bionetworks), Thomas Lehner, Barbara Lipska (NIMH). The Genotype-Tissue Expression (GTEx) Project was supported by the Common Fund of the Office of the Director of the National Institutes of Health, and by NCI, NHGRI, NHLBI, NIDA, NIMH, and NINDS. We thank Miss. Qian Li for her technique support.

**Funding**

This study was equally supported by the Key Research Project of Yunnan Province (202101AS070055) and the Strategic Priority Research Program of the Chinese Academy of Sciences (XDPB17). Also was supported by the Innovative Research Team of Science and Technology department of Yunnan Province (2019HC004), the Distinguished Young Scientists grant of the Yunnan Province (202001AV070006) (to X-J.L.), the National Nature Science Foundation of China (31970561) (to X-J.L.), the Western Light Innovative Research Team of Chinese Academy of Sciences (to X-J.L.), and the Western Light Program of Chinese Academy of Sciences (to J.W.L.).

**Author contributions**

XJL conceived, designed and supervised the whole study. SWL and JL conducted all experiments and analyzed data under the supervision of XJL. JWL and YXH performed functional genomics study and transcriptome analysis. XYL performed eQTL and differential expression analyses. JYW isolated the rat cortical neurons and YFL conducted the reporter gene assays. YXL participated in proliferation and differentiation of NSCs. XX and ML provided critical comments on study design and manuscript writing. SWL interpreted the results and drafted the manuscript. XJL oversaw the project and finalized the manuscript. All authors revised the manuscript critically and approved the final version.

**Competing interests**

The authors declare no competing interests.
Supplementary material

Supplementary material is available at *Brain* online.

References

1. Mueser KT, McGurk SR. Schizophrenia. *Lancet*. 2004;363(9426):2063-2072.
2. Tandon R, Gaebel W, Barch DM, et al. Definition and description of schizophrenia in the DSM-5. *Schizophrenia Research*. 2013;150(1):3-10.
3. Sullivan PF, Kendler KS, Neale MC. Schizophrenia as a Complex Trait: Evidence From a Meta-analysis of Twin Studies. *Archives of General Psychiatry*. 2003;60(12):1187-1192.
4. Hilker R, Helenius D, Fagerlund B, et al. Heritability of Schizophrenia and Schizophrenia Spectrum Based on the Nationwide Danish Twin Register. *Biological Psychiatry*. 2018;83(6):492-498.
5. Ripke S, O'Dushlaine C, Chambert K, et al. Genome-wide association analysis identifies 13 new risk loci for schizophrenia. *Nature Genetics*. 2013;45(10):1150-1159.
6. Ripke S, Neale BM, Corvin A, et al. Biological insights from 108 schizophrenia-associated genetic loci. *Nature*. 2014;511(7510):421-427.
7. Li Z, Chen J, Yu H, et al. Genome-wide association analysis identifies 30 new susceptibility loci for schizophrenia. *Nature Genetics*. 2017;49(11):1576-1583.
8. Pardiñas AF, Holmans P, Pocklington AJ, et al. Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. *Nature genetics*. 2018;50(3):381-389.
9. Ripke S, Sanders A, Kendler K, et al. Schizophrenia Psychiatric Genome-Wide Association Study (GWAS) Consortium. Genome-wide association study identifies five new schizophrenia loci. *Nature genetics*. 2011;43:969-976.
10. Lam M, Chen C-Y, Li Z, et al. Comparative genetic architectures of schizophrenia in East Asian and European populations. *Nature Genetics*. 2019;51(12):1670-1678.
11. Yue WH, Wang HF, Sun LD, et al. Genome-wide association study identifies a susceptibility locus for schizophrenia in Han Chinese at 11p11.2. *Nat Genet.* 2011;43(12):1228-1231.
12. Shi Y, Li Z, Xu Q, et al. Common variants on 8p12 and 1q24.2 confer risk of schizophrenia. *Nat Genet.* 2011;43(12):1224-1227.
13. Mah W, Won H. The three-dimensional landscape of the genome in human brain tissue unveils regulatory mechanisms leading to schizophrenia risk. *Schizophrenia Research*. 2020;217:17-25.
14. Won H, de la Torre-Ubieta L, Stein JL, et al. Chromosome conformation elucidates regulatory relationships in developing human brain. *Nature*. 2016;538(7626):523-527.
15. Whitington T, Gao P, Song W, et al. Gene regulatory mechanisms underpinning prostate cancer susceptibility. *Nature Genetics*. 2016;48(4):387-397.
16. Huo Y, Li S, Liu J, et al. Functional genomics reveal gene regulatory mechanisms underlying schizophrenia risk. *Nature Communications*. 2019;10(1):670.
17. Li S, Li Y, Li X, et al. Regulatory mechanisms of major depressive disorder risk variants. *Molecular Psychiatry.* 2020;25(9):1926-1945.
18. An integrated encyclopedia of DNA elements in the human genome. *Nature.* 2012;489(7414):57-74.
19. Davis CA, Hitz BC, Sloan CA, et al. The Encyclopedia of DNA elements (ENCODE): data portal update. *Nucleic acids research.* 2018;46(D1):D794-d801.
20. Auton A, Abecasis GR, Altshuler DM, et al. A global reference for human genetic variation. *Nature.* 2015;526(7571):68-74.
21. Arnold M, Raffler J, Pfeufer A, et al. SNiPA: an interactive, genetic variant-centered annotation browser. *Bioinformatics.* 2015;31(8):1334-1336.
22. Fromer M, Roussos P, Sieberts SK, et al. Gene expression elucidates functional impact of polygenic risk for schizophrenia. *Nature neuroscience.* 2016;19(11):1442-1453.
23. Collado-Torres L, Burke EE, Peterson A, et al. Regional Heterogeneity in Gene Expression, Regulation, and Coherence in the Frontal Cortex and Hippocampus across Development and Schizophrenia. *Neuron.* 2019;103(2):203-16.e8.
24. Gusev A, Ko A, Shi H, et al. Integrative approaches for large-scale transcriptome-wide association studies. *Nat Genet.* 2016;48(3):245-252.
25. Wang D, Liu S, Warrell J, et al. Comprehensive functional genomic resource and integrative model for the human brain. *Science.* 2018;362(6420):eaat8464.
26. Zhang Y, Li S, Li X, et al. Convergent lines of evidence support NOTCH4 as a schizophrenia risk gene. *Journal of Medical Genetics.* 2020;jmedgenet-2020-106830.
27. Yang Z, Zhou D, Li H, et al. The genome-wide risk alleles for psychiatric disorders at 3p21.1 show convergent effects on mRNA expression, cognitive function, and mushroom dendritic spine. *Molecular Psychiatry.* 2020;25(1):48-66.
28. Miller DE, Patel ZH, Lu X, et al. Screening for Functional Non-coding Genetic Variants Using Electrophoretic Mobility Shift Assay (EMSA) and DNA-affinity Precipitation Assay (DAPA). *Journal of visualized experiments : JoVE.* 2016(114):54093.
29. Hellman LM, Fried MG. Electrophoretic mobility shift assay (EMSA) for detecting protein-nucleic acid interactions. *Nature protocols.* 2007;2(8):1849-1861.
30. Yang D, Jang I, Choi J, et al. 3DIV: A 3D-genome Interaction Viewer and database. *Nucleic acids research.* 2018;46(D1):D52-d57.
31. Vandesompele J, De Preter K, Pattyn F, et al. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome biology.* 2002;3(7):Research0034.
32. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods.* 2001;25(4):402-408.
33. Glausier JR, Lewis DA. Dendritic spine pathology in schizophrenia. *Neuroscience.* 2013;251:90-107.
34. Garey LJ, Ong WY, Patel TS, et al. Reduced dendritic spine density on cerebral cortical pyramidal neurons in schizophrenia. *J Neurol Neurosurg Psychiatry.* 1998;65(4):446-453.
35. Glantz LA, Lewis DA. Decreased Dendritic Spine Density on Prefrontal Cortical Pyramidal Neurons in Schizophrenia. *Archives of General Psychiatry.* 2000;57(1):65-73.
36. Srivastava DP, Woolfrey KM, Penzes PJJoVE. Analysis of dendritic spine morphology in cultured CNS neurons. 2011(53):e2794.
37. Dumitriu D, Rodriguez A, Morrison J. High-throughput, detailed, cell-specific neuroanatomy of dendritic spines using microinjection and confocal microscopy. Nature protocols. 2011;6:1391-1411.

38. Zagrebelsky M, Holz A, Dechant G, et al. The p75 neurotrophin receptor negatively modulates dendrite complexity and spine density in hippocampal neurons. The Journal of neuroscience: the official journal of the Society for Neuroscience. 2005;25(43):9989-9999.

39. Langmead B, Wilks C, Antonescu V, et al. Scaling read aligners to hundreds of threads on general-purpose processors. Bioinformatics. 2018;35(3):421-432.

40. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nature methods. 2012;9(4):357-359.

41. Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009;25(16):2078-2079.

42. Anders S, Pyl PT, Huber W. HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics. 2015;31(2):166-169.

43. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome biology. 2014;15(12):550.

44. Huang da W, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic acids research. 2009;37(1):1-13.

45. Sherf, B., Navarro, S.L., Hannah, R., & Wood, K. Dual-Luciferase TM Reporter Assay: An Advanced Co-Reporter Technology Integrating Firefly and Renilla Luciferase Assays; Promega Notes. 1996; 57, 2–9.

46. Götz M, Huttner WB. The cell biology of neurogenesis. Nature Reviews Molecular Cell Biology. 2005;6(10):777-788.

47. Insel TR. Rethinking schizophrenia. Nature. 2010;468(7321):187-193.

48. Weinberger DR. Implications of normal brain development for the pathogenesis of schizophrenia. Arch Gen Psychiatry. 1987;44(7):660-669.

49. Mao Y, Ge X, Frank CL, et al. Disrupted in schizophrenia 1 regulates neuronal progenitor proliferation via modulation of GSK3beta/beta-catenin signaling. Cell. 2009;136(6):1017-1031.

50. Kim JY, Liu CY, Zhang F, et al. Interplay between DISC1 and GABA signaling regulates neurogenesis in mice and risk for schizophrenia. Cell. 2012;148(5):1051-1064.

51. Kirschenbaum B, Nedergaard M, Preuss A, et al. In vitro neuronal production and differentiation by precursor cells derived from the adult human forebrain. Cerebral cortex (New York, NY: 1991). 1994;4(4):576-589.

52. Ying Z, Gonzalez-Martinez J, Tilelli C, et al. Expression of neural stem cell surface marker CD133 in balloon cells of human focal cortical dysplasia. Epilepsia. 2005;46(11):1716-1723.

53. Pantazopoulos H, Markota M, Jaquet F, et al. Aggrecan and chondroitin-6-sulfate abnormalities in schizophrenia and bipolar disorder: a postmortem study on the amygdala. Transl Psychiatry. 2015;5(1):e496-e496.

54. Fan Y, Abrahamsen G, Mills R, et al. Focal adhesion dynamics are altered in schizophrenia. Biol Psychiatry. 2013;74(6):418-426.

55. Zheng W, Wang H, Zeng Z, et al. The possible role of the Akt signaling pathway in...
schizophrenia. *Brain research*. 2012;1470:145-158.

56. Schrode N, Ho SM, Yamamuro K, et al. Synergistic effects of common schizophrenia risk variants. *Nat Genet*. 2019;51(10):1475-1485.

57. Chang H, Cai X, Li H-J, et al. Functional genomics identify a regulatory risk variation rs4420550 in the 16p11.2 schizophrenia-associated locus. *Biological Psychiatry*. 2020; 89(3):246-255.

58. Kichaev G, Yang WY, Lindstrom S, et al. Integrating functional data to prioritize causal variants in statistical fine-mapping studies. *PLoS genetics*. 2014;10(10):e1004722.

59. Liu J, Li S, Li X, et al. Genome-wide association study followed by trans-ancestry meta-analysis identify 17 new risk loci for schizophrenia. *BMC Medicine*. 2021; DOI:10.1186/s12916-021-02039-9.

60. Sharma T, Antonova L. Cognitive function in schizophrenia. Deficits, functional consequences, and future treatment. *The Psychiatric clinics of North America*. 2003;26(1):25-40.

61. Kuperberg G, Heckers S. Schizophrenia and cognitive function. *Current opinion in neurobiology*. 2000;10(2):205-210.

62. Savage JE, Jansen PR, Stringer S, et al. Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. *Nat Genet*. 2018;50(7):912-919.

63. Davies G, Marioni RE, Liewald DC, et al. Genome-wide association study of cognitive functions and educational attainment in UK Biobank (N=112 151). *Mol Psychiatry*. 2016;21(6):758-767.

64. Polioudakis D, de la Torre-Ubieta L, Langerman J, et al. A Single-Cell Transcriptomic Atlas of Human Neocortical Development during Mid-gestation. *Neuron*. 2019;103(5):785-801.e8.

65. Speir ML, Bhaduri A, Markov NS, et al. UCSC Cell Browser: Visualize Your Single-Cell Data. *Bioinformatics*. 2021; Jul 9:btab503.

66. Nowakowski TJ, Bhaduri A, Pollen AA, et al. Spatiotemporal gene expression trajectories reveal developmental hierarchies of the human cortex. *Science*. 2017;358(6368):1318-1323.

67. Kang HJ, Kawasawa YI, Cheng F, et al. Spatio-temporal transcriptome of the human brain. *Nature*. 2011;478(7370):483-489.

68. Pletikos M, Sousa AM, Sedmak G, et al. Temporal specification and bilaterality of human neocortical topographic gene expression. *Neuron*. 2014;81(2):321-332.

69. Walker RL, Ramaswami G, Hartl C, et al. Genetic Control of Expression and Splicing in Developing Human Brain Informs Disease Mechanisms. *Cell*. 2019;179(3):750-71.e22.

70. Stahl EA, Breen G, Forstner AJ, et al. Genome-wide association study identifies 30 loci associated with bipolar disorder. *Nat Genet*. 2019;51(5):793-803.

71. Wray NR, Ripke S, Mattheisen M, et al. Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nat Genet*. 2018;50(5):668-681.

72. Demontis D, Walters RK, Martin J, et al. Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. *Nat Genet*. 2019;51(1):63-75.

73. Noma A, Ishitani R, Kato M, et al. Expanding role of the jumonji C domain as an RNA hydroxylase. *The Journal of biological chemistry*. 2010;285(45):34503-34507.
74. Kato M, Araiso Y, Noma A, et al. Crystal structure of a novel JmjC-domain-containing protein, TYW5, involved in tRNA modification. *Nucleic acids research.* 2011;39(4):1576-1585.
75. Noma A, Kirino Y, Ikeuchi Y, et al. Biosynthesis of wybutosine, a hyper-modified nucleoside in eukaryotic phenylalanine tRNA. *The EMBO journal.* 2006;25(10):2142-2154.
76. Key J, Sen NE, Arsović A, et al. Systematic Surveys of Iron Homeostasis Mechanisms Reveal Ferritin Superfamily and Nucleotide Surveillance Regulation to be Modified by PINK1 Absence. *Cells.* 2020;9(10):2229
77. Saghazadeh A, Mahmoudi M, Shahrokhi S, et al. Trace elements in schizophrenia: a systematic review and meta-analysis of 39 studies (N = 5151 participants). *Nutrition reviews.* 2020;78(4):278-303.
78. Purcell SM, Wray NR, Stone JL, et al. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature.* 2009;460(7256):748-752.
Figure legends

Figure 1. Flowchart of the functional genomics analysis of the 2q33.1 locus. The 2q33.1 locus (index variant Chr2_200825237_I) was ranked among the top 10 most significant risk loci for schizophrenia in PGC2 (2014) ($P = 5.65 \times 10^{-14}$). Our previous functional genomics identified two functional SNPs (rs796364 and rs281759) on the 2q33.1, and they ranked as the second ($P_{rs796364} = 9.41 \times 10^{-17}$) and third most significant associations ($P_{rs281759} = 9.25 \times 10^{-16}$) (with schizophrenia) in study of Pardinas et al. Briefly, SNPs in LD ($r^2 > 0.3$) with the index variant (Chr2_200825237_I) were extracted using genotype data (Europeans) from the 1000 Genomes Project. The DNA binding motifs of the included TFs were then derived and compared with the position weight matrices (PWMs) (from PWMs database (JASPAR, TRANSFC, Uniprobe and Hi-SELEX)) and the best matched PWMs were used for subsequent analysis. FIMO software was used to scan if the test SNP is located in binding motif and different alleles of this SNP disrupt TFs binding. Two functional SNPs (rs796364 and rs281759) were identified at the 2q33.1 locus and these two SNPs disrupt CTCF, RAD21 and FOXP2 binding. Furthermore, brain eQTL data (CMC and LIBD2) were used to identify the potential target genes of rs796364 and rs281759, and expression data from the CMC (279 controls and 258 schizophrenia cases) and PsychEncode (936 controls and 559 schizophrenia cases) were used to explore if the target gene of rs796364 and rs281759 was dysregulated in schizophrenia. These data suggested that rs796364 and rs281759 may confer schizophrenia risk by regulating TYW5 expression.

Figure 2 Functional genomics identified two functional SNPs (rs796364 and rs281759) at 2q33.1. (A) rs796364 ($P = 9.25 \times 10^{-16}$) and rs281759 ($P = 9.41 \times 10^{-17}$) were significantly associated with schizophrenia (data from PGC+CLOZUK). Of note, these two SNPs are in high linkage disequilibrium (LD) ($r^2 = 0.97$) and both SNPs are located in the downstream of TYW5. The distance between these two SNPs is 71,600 bp. (B-C) Functional genomics showed that rs796364 and rs281759 disrupts TFs binding, rs796364 disrupts binding of CTCF and RAD21 (B), and rs281759 disrupts binding of FOXP2 (C). (D-E) rs796364 and rs281759 are located in genomics regions with signals of DNase-Seq, ChIP-Seq and histone modifications.
in human neuroblastoma cell lines, indicating that they are located in actively transcribed (i.e., regulatory elements) regions (with corresponding TFs binding).

Figure 3 Reporter gene assays and EMSA validated the regulatory effects of rs796364 and rs281759. (A-F) Reporter gene assays (enhancer assay) showed that different alleles of rs796364 and rs281759 affected the luciferase activity significantly. Compared with the C allele, the A allele of rs796364 conferred significantly higher luciferase activity in SH-SY5Y and SK-N-SH cells. For rs281759, the T allele exhibited significantly higher activity compared with C allele in all three tested cell lines. As these two SNPs were cloned into pGL3-promoter, these assays reflect the regulatory (i.e., enhancer) effect of the inserted fragments on luciferase promoter. (G-J) rs796364 and rs281759 affected binding affinity of NPEs. The quantification data (gray values, arbitrary unit) of the probes/protein complexes were firstly quantified with Image J. The obtained values were then normalized to the total quantity of probes. The quantification data of (G) was showed in (H) and the quantification data of (I) was showed in (J). For rs796364, the A allele showed stronger binding compared with C allele (G-H). For rs281759, the T allele showed stronger binding compared with C allele (I-J). Data represent mean ± SD. n = 16. Two-tailed Student’s t-test was used to test if the difference was significant.

Figure 4 Rs796364 and rs281759 are associated with TYW5 expression in the human brain. (A-D) Brain eQTL analysis (CMC and LIBD2) showed that rs796364 and rs281759 were significantly associated with TYW5 expression in the human brain.22,23 (E-F) Knockout of genomic sequences (about 300 bp) containing rs796364 or rs281759 resulted in significant down-regulation of TYW5 expression, indicating that the genomic region surrounding rs796364 and rs281759 act as enhancers to regulate TYW5 expression. (G-P) TFs (CTCF, RAD21 and FOXP2) knockdown altered TYW5 gene expression, indicating that TYW5 gene expression was regulated by CTCF, RAD21 and FOXP2. CTCF, RAD21 and FOXP2 were significantly knocked-down in HEK293T (G-I), SH-SY5Y (K, L) and SK-N-SH (N, O). (J, M and P) qPCR showed that TYW5 expression was significantly altered in TFs knocked-down cells. (Q) TYW5 expression was significantly up-regulated in brains of schizophrenia patients compared with controls. Data represent mean ± SD. n = 3 for (F-P). ACTB and GAPDH were used as
endogenous reference genes for HEK293T, and ACTB and B2M were used as endogenous reference genes for SH-SY5Y and SK-N-SH. Two-tailed Student’s t-test was used to test if the difference was significant.

**Figure 5 TYW5 overexpression significantly inhibited proliferation of mouse NSCs.** (A, B) Validation of TYW5 overexpression in mouse NSCs with RT-PCR (B) and Western-blot (c). To distinguish the endogenous and the overexpressed TYW5 (both can express TYW5), we ligated a 6×His tag to TYW5 to detect if the constructed overexpression vector can express TYW5 successfully in NSC by using antibody that can recognize His tag. (C-E) EdU incorporation assays and CCK-8 assays showed that TYW5 overexpression significantly inhibited proliferation of mouse NSCs. (C-D) Compared with controls, the ratio of EdU positive cells were significantly decreased in TYW5 overexpression groups. (C) Representative immunofluorescence images for EdU staining, the statistical analysis result of (C) was showed in (D). (E) CCK-8 assay showed that the OD\textsubscript{450} value of TYW5 overexpression groups were significantly decreased compared with controls. Two-tailed Student’s t-test was used for statistical test. Data represent mean ± SD, n = 5 for (D), n = 8 for (E).

**Figure 6 TYW5 overexpression significantly affected differentiation of mouse NSCs.** (A-B) Representative immunofluorescence images for GFAP (a marker for astrocyte cells) and MAP2 (a marker for mature neurons) staining. (C) Compared with controls, the ratio of GFAP positive cells were significantly decreased in TYW5 overexpression groups, indicating TYW5 overexpression impaired differentiation of NSCs into glia cells. (D) By contrast, the ratio of MAP2 positive cells was not changed. (C) The quantification data for (A). (D) The quantification data for (B). (E-H) Western-blot results for GFAP and MAP2 expression in TYW5 overexpression and control groups. For MAP2, two bands (MAP2 A/B, molecular weight: 270-300 kDa; and MAP2 C/D, molecular weight: 70-75 kDa) were detected. (F-H) The quantification data of MAP2 A/B, MAP2 C/D and GFAP protein expression for (E). Two-tailed Student’s t-test was used for statistical test. Data represent mean ± SD, n = 3 for (C-D and F-H).
**Figure 7** *TYW5* regulates schizophrenia-associated pathways. (A) Heatmap for the 304 (183 up-regulated and 121 down-regulated) differentially expressed genes (|fold change| > 0.5 and \( P_{\text{adj}} < 0.01 \)) in *TYW5* overexpression group compared with control group. (B) Heatmap for the top 25 up or down-regulated genes. (C) GO analysis showed that the differentially expressed genes showed the most significant enrichment in DNA replication. (D) KEGG enrichment analysis showed that differentially expressed genes showed significant enrichment schizophrenia-associated pathways, including ECM-receptor interaction, Focal adhesion and PI3K-Akt signaling pathways.

**Figure 8** *TYW5* overexpression affected dendritic spine density. (A) Representative immunofluorescence images for GFP (visualizing neuronal morphologies with green) and *TYW5*-1×flag (indicating overexpression with red) in rat primary cortical neurons. The arrows with different colors pointed to different type of spines, stubby (green), thin (orange), long thin (blue) and mushroom (white) spines. (B-C) Compared with control the density and ratio of stubby spines were significantly decreased in *TYW5* overexpressing neurons. However, the density and ratio of thin spines were significantly increased in *TYW5* overexpressing neurons. (B) The quantification data for the density of stubby spines in (A). (C) The quantification data for the ratio of stubby spines in (A). Two-tailed *Student’s* \( t \)-test was used for statistical test. Data represent \( \text{mean} \pm \text{SD} \), the number of neurons used for this quantitative analysis in controls and experiment groups are \( n = 20 \) and \( n = 22 \).
108 schizophrenia-associated genetic loci (128 index SNPs) (PGC2, 2014)

Index variant chr2_200825237_I ($P=5.65 \times 10^{-14}$) at 2q33.1 ranked the top 10 most significant variants among the 128 index SNPs.

LD SNPs extraction ($r^2>0.3$)

101 SNPs in LD with chr2_200825237_I

**Functional genomics**

30 TFs ChIP-seq

Motifs of 30 TFs

FIMO

PWM data (JASPAR, TRANSFC, Uniprobe, HI-SELEX)

Identification of two functional SNPs, rs796364 ($P=9.41 \times 10^{-11}$, disrupts CTCF and RAD21 binding) & rs281759 ($P=9.25 \times 10^{-18}$, disrupts FOXP2 binding)

These two SNPs showed the second and third most significant associations with schizophrenia in study of Pardinas et al.

eQTL analysis

Differential expression analysis

**schizophrenia risk gene** **TYW5**

*Figure 1*
Figure 2
Figure 3
Figure 4
Figure 6
Figure 8