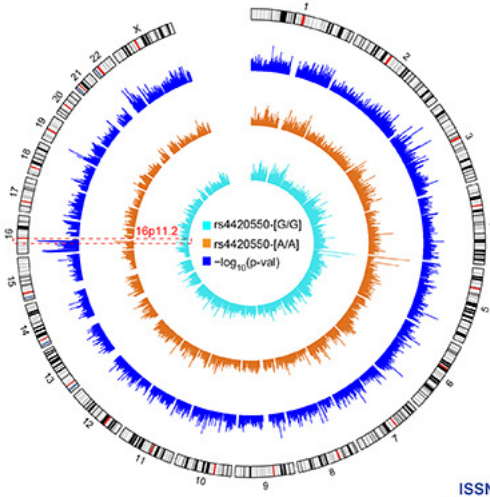
Biological Psychiatry

Volume 89, Number 3
February 1, 2021

Transdiagnostic Perspectives on Psychosis Risk



ISSN 0006-3223
www.sobp.org/journal
Official Journal of the Society of Biological Psychiatry

Biological Psychiatry

Archival Report

Functional Genomics Identify a Regulatory Risk Variation rs4420550 in the 16p11.2 Schizophrenia-Associated Locus

Hong Chang, Xin Cai, Hui-Juan Li, Wei-Peng Liu, Li-Juan Zhao, Chu-Yi Zhang, Jun-Yang Wang, Jie-Wei Liu, Xiao-Lei Ma, Lu Wang, Yong-Gang Yao, Xiong-Jian Luo, Ming Li, and Xiao Xiao

ABSTRACT

BACKGROUND: Genome-wide association studies (GWASs) have reported hundreds of genomic loci associated with schizophrenia, yet identifying the functional risk variations is a key step in elucidating the underlying mechanisms.

METHODS: We applied multiple bioinformatics and molecular approaches, including expression quantitative trait loci analyses, epigenome signature identification, luciferase reporter assay, chromatin conformation capture, homology-directed genome editing by CRISPR/Cas9 (clustered regularly interspaced short palindromic repeats/Cas9), RNA sequencing, and ATAC-Seq (assay for transposase-accessible chromatin using sequencing).

RESULTS: We found that the schizophrenia GWAS risk variations at 16p11.2 were significantly associated with messenger RNA levels of multiple genes in human brain, and one of the leading expression quantitative trait loci genes, *MAPK3*, is located ~200 kb away from these risk variations in the genome. Further analyses based on the epigenome marks in human brain and cell lines suggested that a noncoding single nucleotide polymorphism, rs4420550 ($p = 2.36 \times 10^{-9}$ in schizophrenia GWAS), was within a DNA enhancer region, which was validated via in vitro luciferase reporter assays. The chromatin conformation capture experiment showed that the rs4420550 region physically interacted with the *MAPK3* promoter and *TAOK2* promoter. Precise CRISPR/Cas9 editing of a single base pair in cells followed by RNA sequencing further confirmed the regulatory effects of rs4420550 on the transcription of 16p11.2 genes, and ATAC-Seq demonstrated that rs4420550 affected chromatin accessibility at the 16p11.2 region. The rs4420550-[A/A] cells showed significantly higher proliferation rates compared with rs4420550-[G/G] cells.

CONCLUSIONS: These results together suggest that rs4420550 is a functional risk variation, and this study illustrates an example of comprehensive functional characterization of schizophrenia GWAS risk loci.

https://doi.org/10.1016/j.biopsych.2020.09.016

Accumulating evidence suggests that schizophrenia is a highly heritable polygenic disorder, and genome-wide surveys have reported more than 100 risk loci containing common single nucleotide polymorphisms (SNPs) (1,2) or rare copy number variations (CNVs) (3–6). Among them, the chromosome 16p11.2 region has been consistently highlighted in genetic studies of schizophrenia. For example, rare microduplications at the recurrent ~600-kb 16p11.2 BP4-BP5 locus are significantly associated with schizophrenia (4,5,7). This 16p11.2 CNV region encompasses more than 25 genes. A previous study examined the effect of each individual gene on zebrafish head sizes and identified *KCTD13* as the major driver, which participated in head-size regulation by interacting with *MAPK3* and *MVP* (8). Besides, Blizinsky *et al.* (9) observed increased dendritic arborization in cortical pyramidal neurons isolated from mice carrying the heterozygous 16p11.2 duplication, which was rescued through pharmacological targeting of ERK-1 (extracellular signal-regulated kinase 1) (encoded by *MAPK3*). Other studies also reported regulatory effects of 16p11.2 CNV genes on dendritic spine development and synaptic function (10–12), which are believed to contribute to the pathogenesis of schizophrenia (13–18).

Common variations in the proxy 16p11.2 CNV region also showed genome-wide significant associations with schizophrenia (1,2,19). However, risk SNPs primarily locate in noncoding regions with undetermined physiological impact. It is presumed that such noncoding DNA variations should either be in linkage with unknown causal mutations or exert regulatory effects themselves (20–27), e.g., they might affect messenger RNA (mRNA) expression through modulating physical interaction between transcription factors and sequence elements (e.g., DNA enhancers) (28–34). In the present study, we performed brain expression quantitative trait loci (eQTLs) analyses and found that schizophrenia risk SNPs at 16p11.2 were significantly associated with mRNA expression of multiple genes. Combinatory analyses including the

 ISSN: 0006-3223

epigenome signature identification, luciferase reporter assay, chromatin conformation capture (3C), homology-directed genome editing by CRISPR/Cas9 (clustered regularly interspaced short palindromic repeats/Cas9), RNA sequencing (RNA-Seq), and ATAC-Seq (assay for transposase-accessible chromatin using sequencing) identified a functional noncoding SNP rs4420550 residing in an enhancer region at 16p11.2.

METHODS AND MATERIALS

The study protocol was approved by the Institutional Review Board of Kunming Institute of Zoology, Chinese Academy of Sciences. Detailed methods and materials are described in the Methods and Materials in Supplement 1. The primer sequences in real-time quantitative polymerase chain reaction (RT-qPCR) analysis are shown in Table S1 in Supplement 2, the primers used in the 3C assay are listed in Table S2 in Supplement 2, the single guide RNA oligo sequences targeting genes are shown in Table S3 in Supplement 2, and the nucleotide sequences of the double-stranded oligonucleotides in electrophoretic mobility shift assay are shown in Table S4 in Supplement 2.

The eQTL analyses were performed using datasets of the dorsolateral prefrontal cortex (DLPFC) tissues from BrainSeq Phase 1 (n = 412) (35), Brain xQTL (n = 494) (36), CommonMind Consortium (n = 467) (37), Genotype-Tissue Expression (GTEx) Project (n = 175) (38), and PsychENCODE Consortium (n = 1387) (39). We applied the summary data–based Mendelian randomization (SMR) analysis (40,41) to prioritize risk genes, which integrated the schizophrenia genome-wide association studies (GWASs) (40,675 cases and 64,643 controls) (2) with the aforementioned eQTL datasets to assess the pleiotropic effects of SNPs on diagnosis and mRNA expressions.

Linkage disequilibrium (LD) between pairwise SNPs was calculated using Haploview v4.1 (42) based on the European genotype data from 1000 Genomes Project (43). Histone marks (such as H3K4me1, H3K4me3, and H3K27ac) were queried in the available Roadmap Epigenomics projects and ENCODE (Encyclopedia of DNA Elements) datasets of brain tissues and cell lines (HEK293 [human embryonic kidney 293] and SK-N-SH) (44,45). RegulomeDB (http://www.regulomedb.org/) (46) and HaploReg v4.1 (https://pubs.broadinstitute.org/mammals/haploreg/haploreg.php) (47) were also used to determine the SNPs overlapped with open chromatin peaks and chromatin immunoprecipitation sequencing (ChIP-Seq) peaks of histone modifications in brain tissues.

Luciferase reporter assay was conducted in the HEK293T and SK-N-SH cells to examine the effects of SNPs on DNA enhancer activities. DNA fragments containing target SNPs were amplified, and site-directed mutagenesis was employed to generate sequences containing either alleles. The DNA fragments were then cloned into the pGL3-promoter vector, and equal amounts of each plasmid were transiently transfected into cells together with pRL-TK using Lipofectamine 2000 (Thermo Fisher Scientific, Waltham, MA). Twenty-four hours posttransfection, cells were lysed and luciferase activity was determined using the Dual-Luciferase Reporter Assay System (Promega, Madison, WI). The firefly luciferase activity was normalized to that of *Renilla* luciferase to control for transfection efficiency variations. All assays were performed in triplicate in at least three independent experiments. Two-tailed t test was performed for statistical analyses.

The 3C assay was performed following previous studies (48,49) to detect physical interactions between genomic regions. A total of 1×10^7 HEK293T cells in 10 mL of Dulbecco's Modified Eagle Medium culture medium were fixed using 1% formaldehyde at room temperature for 10 minutes, and the fixed cells were lysed in 5 mL lysis buffer for 15 minutes at 4°C with gentle pipetting. The mixture was centrifuged, and the nuclei pellet was then suspended in 500 μL of 1.2× restriction enzyme buffer. Following this, the pellets were digested with 400 U restriction enzyme EcoRI (New England Biolabs, Ipswich, MA), and T4 DNA ligase (100 U; Promega) was added to the diluted chromatin and incubated for 4 hours at 16°C. Proteinase K (500 μg) was then added and incubated at 65°C overnight to reverse crosslinks, and RNA was removed through RNase A (300 μg) treatment for 60 minutes at 37°C. TaqMan RT-qPCR (Thermo Fisher Scientific) was performed to determine the interaction frequencies between target sites of interest, and the PCR products were purified using a QIAGEN quick gel purification kit for Sanger sequencing of each chimeric DNA (QIAGEN, Hilden, Germany). To normalize primer efficiency, control PCR templates were generated by digestion and random ligation of bacterial artificial chromosomes covering the 16p11.2 region.

Homology-directed repair–mediated genome editing of rs4420550 by CRISPR/Cas9 was conducted in HEK293T cells following a previous study with modifications (50). The protospacer sequence of CRISPR/Cas9 targeting rs4420550 (5′-GATGTTGCTAGGAGCTGACC-3′) was designed using Optimized CRISPR Design (http://crispr.mit.edu), and annealed oligomers were cloned into the pL-CRISPR-E.EFS.GFP plasmid. A single-stranded oligodeoxynucleotide (5′-gttttttgtgaaattattttgctcaacaggacaccaccatggccgtgctgttggtgccaggCcagctcctagcaacatcaaggcttctggagtgagggtgcaaaccagctcccagctggcagc-3′) was synthesized as a homology-directed repair template for precise editing of the rs4420550 locus. The pL-CRISPR-E.EFS.GFP construct and the single-stranded oligodeoxynucleotide were transiently cotransfected into HEK293T cells on 100-mm plates using Lipofectamine 3000. Forty-eight hours after transfection, cells positive for green fluorescent protein were sorted with a flow cytometer. Colonies derived from single cells were then obtained from these cells, and four colonies were selected for each homozygous genotype ([A/A] or [G/G]) of rs4420550. Sanger sequencing was conducted to ensure that the colonies had the same DNA sequence, except for the rs4420550 locus.

These eight single-cell colonies were then subject to RT-qPCR, RNA-Seq, Western blotting, and ATAC-Seq analyses. The raw RNA-Seq and ATAC-Seq data in rs4420550-edited HEK293T cells are accessible in the Gene Expression Omnibus repository (accession number GSE152177). The Wald test in DESeq2 (51) was used to analyze the mRNA expression differences revealed by RNA-Seq and different chromatin accessibility peaks in ATAC-Seq between colonies carrying [A/A] or [G/G] genotypes at rs4420550. Relative mRNA expression assessed by RT-qPCR was presented as

means of $2^{-\Delta\Delta Ct}$; densities of bands in Western blotting were quantified using ImageJ (version 1.52a; National Institutes of Health, Bethesda, MD); and statistical differences between experimental groups in both assays were examined using 2-tailed *t* test.

Cell proliferation assays were performed using the CCK-8 (cell counting kit-8) reagent (Beyotime Biotechnology, Haimen, China). HEK293T cells with genotype rs4420550-[A/A] or -[G/G] were seeded into 96-well plates at a density of 3×10^3 cells/well and cultured for 24, 48, and 72 hours, and 10 μL of CCK-8 was added into each well and incubated with the cells at 37°C for 1 hour. Optical density at 450 nm was measured for each well using a multidetection microplate reader (Bio-Rad Laboratories, Hercules, CA). All assays were performed in triplicate in at least three independent experiments, and statistical differences of cell proliferation rates at each time point were examined through 2-tailed *t* test.

RESULTS

The Schizophrenia GWAS Risk SNP rs12691307 Is Associated With mRNA Levels of *MAPK3*, *INO80E*, and Other Genes at 16p11.2

To investigate potential impact of the schizophrenia risk SNPs at 16p11.2 on human brain transcriptome, we first examined the correlations between rs12691307 [the leading schizophrenia GWAS risk SNP at 16p11.2, $p = 1.30 \times 10^{-10}$ in 34,241 cases and 45,604 controls (1)] (odds ratio for A allele [i.e., the risk allele], 1.073) and brain transcriptomes in four independent DLPFC eQTL datasets [BrainSeq Phase 1 (n = 412) (35), Brain xQTL (n = 494) (36), CommonMind (n = 467) (37), and GTEx Project frontal cortex (n = 175) (38)], and we found that the risk SNP rs12691307 was consistently and significantly associated with mRNA levels of *MAPK3* ($p = 2.57 \times 10^{-11}$, 8.62×10^{-18}, 1.93×10^{-10}, and 9.70×10^{-11}, respectively) and *INO80E* ($p = 2.33 \times 10^{-10}$, 4.58×10^{-12}, 1.24×10^{-9}, and 3.60×10^{-5}, respectively) (Table S5 in Supplement 2). We also examined the eQTL associations of rs12691307 in a larger dataset from PsychENCODE (n = 1387) (39). Despite that some of these samples were overlapped with those of BrainSeq (35) and CommonMind (36), rs12691307 was again significantly associated with mRNA expression levels of *MAPK3* ($p = 1.31 \times 10^{-26}$) and *INO80E* ($p = 4.98 \times 10^{-19}$) (Table S5 in Supplement 2).

SMR Integrative Analysis of DLPFC eQTL Datasets and Differential Expression Analysis Confirm *MAPK3* and *INO80E* as Schizophrenia Susceptibility Genes

We then performed the SMR analysis (41) integrating schizophrenia GWAS and brain eQTL datasets to define 16p11.2 genes whose mRNA levels were affected by schizophrenia genetic risk. In the DLPFC eQTL datasets from BrainSeq Phase 1 (35), Brain xQTL (36), CommonMind (37), GTEx Project frontal cortex (38), and PsychENCODE (39), we found that mRNA levels of *MAPK3* and *INO80E* were again significantly associated with genetic risk of schizophrenia (*MAPK3*: $p_{\text{SMR-multi}} = 2.68 \times 10^{-5}$, 6.56×10^{-6}, 2.16×10^{-5}, 1.75×10^{-4}, and 2.80×10^{-6}, respectively; *INO80E*: $p_{\text{SMR-multi}} = 2.71 \times 10^{-7}$, 7.59×10^{-7}, 6.45×10^{-6}, 7.57×10^{-5}, and 7.22×10^{-7}, respectively) (Table S6 in Supplement 2). Genetic risk of schizophrenia consistently predicted higher mRNA levels of both *MAPK3* and *INO80E* in all these eQTL datasets.

Besides, the PsychENCODE Consortium has recently analyzed the differentially expressed genes in the DLPFC tissues from 559 cases with schizophrenia versus those from 936 controls (52). Using these data, we found that the mRNA expression of *MAPK3* was significantly higher in the DLPFC of patients with schizophrenia compared with unaffected controls ($p = 3.49 \times 10^{-4}$) (Table S7 in Supplement 2). However, the mRNA levels of *INO80E* did not differ between patients with schizophrenia and controls in their study (p = .429) (52).

LD Analysis of the Schizophrenia Risk SNPs at 16p11.2

We then sought to identify the functional variation(s) among the over 70 schizophrenia risk SNPs at the 16p11.2 GWAS locus ($p < 5.00 \times 10^{-8}$) (Figure 1A) (1). We first examined the LD structure of these SNPs using European genotype data from the 1000 Genomes Project (43). These SNPs primarily constitute two LD blocks (Figure S1 in Supplement 1). rs12691307 locates in LD block 1, and schizophrenia risk SNPs in this LD block generally show stronger association with *MAPK3* than with *INO80E*, and their association signals with *MAPK3* are also stronger than those of the SNPs in LD block 2 [taking results from GTEx Project dataset (38) as an example] (Table S8 in Supplement 2). However, the association signals between *INO80E* and LD block 2 SNPs are stronger than those between *INO80E* and LD block 1 SNPs. Further analyses revealed that the schizophrenia risk SNPs in LD block 2 were located either within or near *INO80E*. Nevertheless, the distance between LD block 1 SNPs and *MAPK3* on the genome is quite far (~200 kb), suggesting long-range regulatory effects of these SNPs. Given that dysregulation of *MAPK3* has been observed in patients with schizophrenia (52) and in behaviorally abnormal animals (53), further studies investigating whether and how the schizophrenia risk SNPs in LD block 1 regulate *MAPK3* mRNA expression is necessary.

Identification of a Potential Regulatory SNP rs4420550 in an Enhancer Element at 16p11.2

To prioritize the regulatory SNP(s) in LD block 1, we examined their functional annotations in human brain tissues and cells using data from the ENCODE (44) and Roadmap Epigenomics projects (45). By comparing the SNP locations with regions of open chromatin depicted by DNase I hypersensitivity, and with regions of active histone H3 lysine modifications (H3K4me1, H3K4me3 and H3K27ac), we found a "potential regulatory region" spanning the GWAS leading risk SNP rs12691307. This region exhibited the largest spatial overlap with regulatory marks in multiple human cell lines (HEK293 and SK-N-SH) and brain tissues (Figure 1A).

This "potential regulatory region" primarily resides in the 5′ flanking region of *KCTD13* and harbors 5 high-LD schizophrenia risk SNPs (rs12716972, rs12716973, rs4420550, rs4424923, and rs12691307) (Figure 1A). However, their eQTL associations with *KCTD13* expression were much weaker than

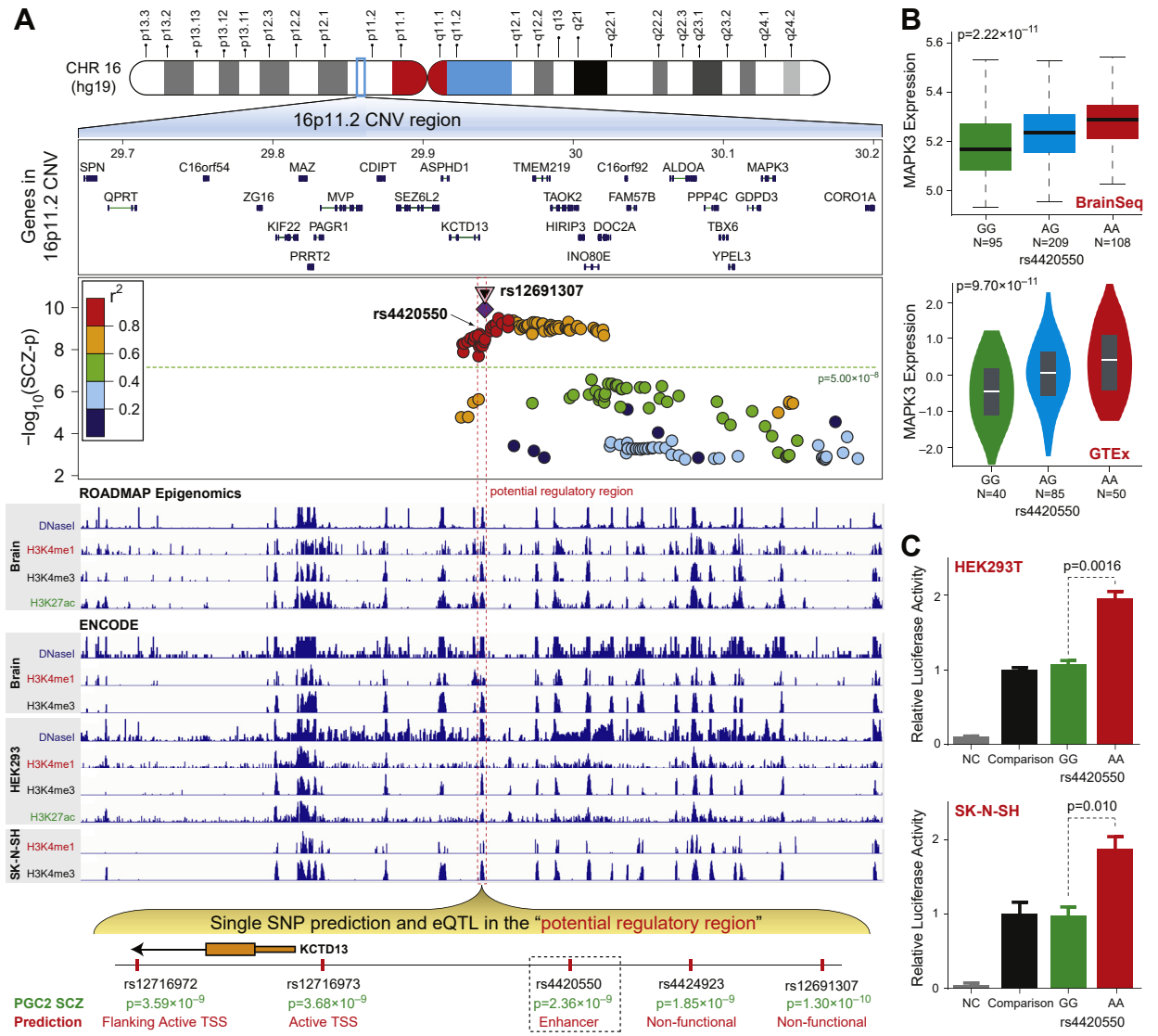


Figure 1. Molecular characterization of rs4420550. **(A)** Genetic associations of SNPs spanning the 16p11.2 CNV region with schizophrenia, and the epigenome signatures in human brains and cell types, lead to identification of a "potential regulatory region" and a regulatory SNP (rs4420550). A physical map of the region is given and depicts known genes within the region, and the LD is defined based on the SNP rs12691307. **(B)** eQTL analyses of rs4420550 with *MAPK3* mRNA expression in the BrainSeq Phase 1 and GTEx Project datasets. **(C)** Results of the reporter gene assay testing the regulatory activities of rs4420550. Effects of rs4420550 allele variation on pGL3-promoter activity are shown in the panels for HEK293T and SK-N-SH cells. The Comparison in the figure represents the empty pGL3 promoter vector, and NC represents the empty pGL3 basic vector. The y-axis values represent fold changes of luciferase activity relative to the empty pGL3-promoter vector. Error bars represent SEM of three biological replicates. Data are representative of three independent experiments with consistent results. CHR, chromosome; CNV, copy number variation; ENCODE, Encyclopedia of DNA Elements; eQTL, expression quantitative trait locus; GTEx, Genotype-Tissue Expression; LD, linkage disequilibrium; mRNA, messenger RNA; NC, negative control; SCZ, schizophrenia; SNP, single nucleotide polymorphism; TSS, transcription start site.

with *MAPK3* (Tables S9 and S10 in Supplement 2). We then used RegulomeDB (46) and HaploReg (47) to prioritize the most likely causative SNP(s) among these 5 SNPs. Intriguingly, analyses using the RegulomeDB dataset (46) suggested that rs4420550 was in an enhancer region in brain tissues (Figure S2 in Supplement 1), rs12716972 and rs12716973 were in a (flanking) active transcription start site (TSS) region, and rs4424923 and rs12691307 were unlikely to be functional (Figure S3 in Supplement 1). Consistently, rs4420550 also colocalized with enhancer histone marks in 16 human tissues, including the brain, in the HaploReg dataset (Figure S4 in Supplement 1) (47).

rs4420550 is in high LD with rs12691307 (r^2 = .99), and is also genome-wide significantly associated with schizophrenia ($p = 2.36 \times 10^{-9}$; odds ratio for A allele, 1.065) (1). The eQTL analyses suggested that rs4420550 was significantly associated with *MAPK3* expression in human brains (Figure 1B) and other organs (Figure S5 in Supplement 1). We then performed

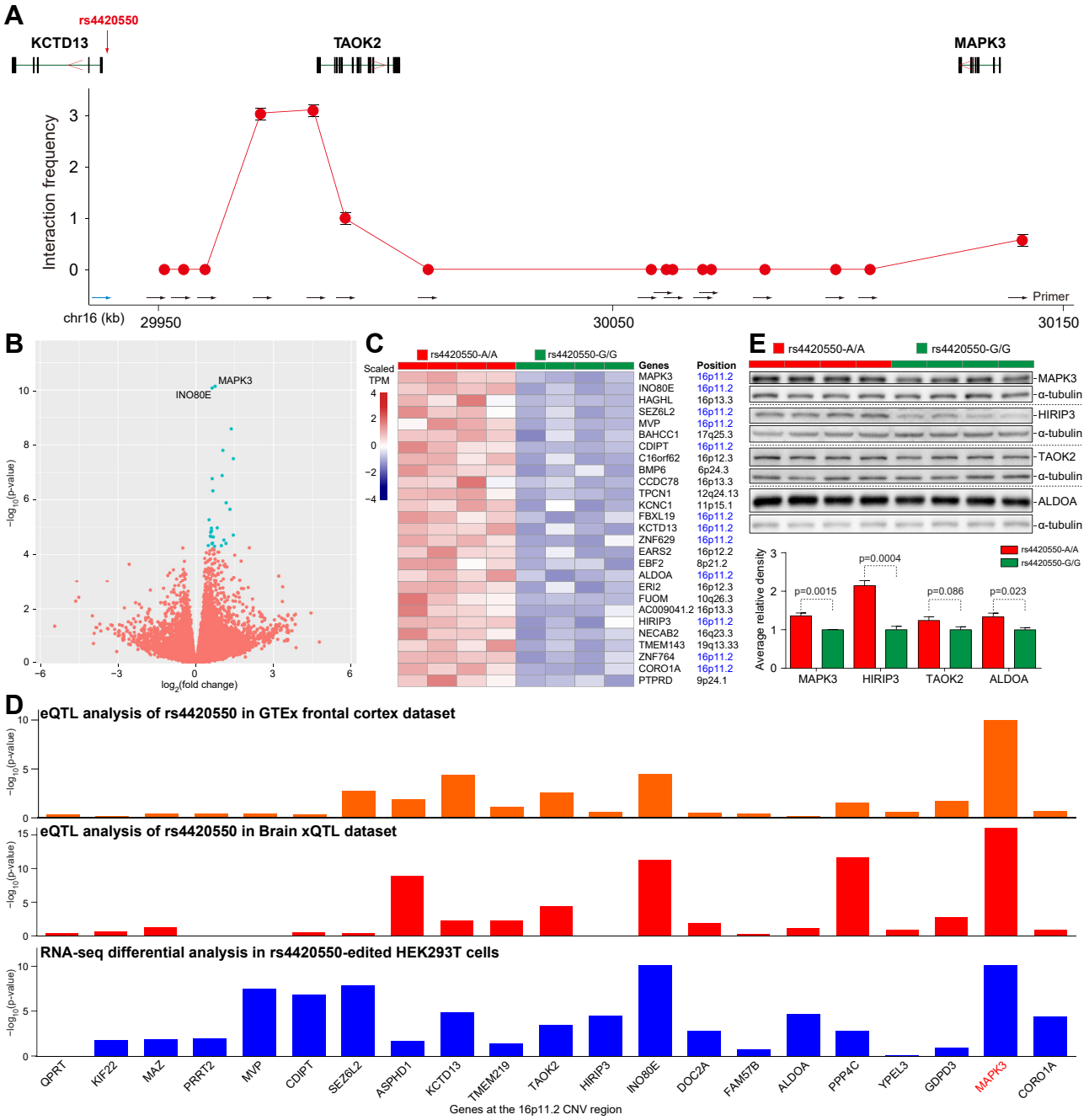


Figure 2. 3C analysis of 16p11.2 region and RNA-Seq analysis of the rs4420550-[A/A] and -[G/G] HEK293T cells. **(A)** 3C interaction profiles between the area (containing rs4420550) and the *MAPK3* promoter region. 3C libraries were generated with EcoRI, with the anchor point set at the rs4420550 locus. Graphs represent two biological replicates. Error bars indicate standard error. **(B)** Volcano plots showing effect size estimates by significance of each gene for the RNA-Seq analysis in the rs4420550-[A/A] and -[G/G] HEK293T cells. Effect sizes captured as $\log_2$(fold change) are shown on the x-axis, and significance levels measured as $-\log_{10}(p \text{ value})$ are shown on the y-axis. Each dot represents an individual gene. Blue dots represent significantly upregulated genes with $\log_2$(fold change) > 0.26 at false discovery rate $< .05$. **(C)** Heatmap of top dysregulated genes in the RNA-Seq analysis of rs4420550-[A/A] and -[G/G] HEK293T cells. **(D)** Comparisons of genes within the proxy 16p11.2 CNV region among GTEx Project dataset, Brain xQTL dataset, and rs4420550-[A/A] and -[G/G] HEK293T cells. **(E)** Western blotting of proteins encoded by the genes at 16p11.2 in rs4420550-[A/A] and -[G/G] HEK293T cells (top) and quantification (bottom). Error bars represent SEM of four biological replicates. Data are representative of three independent experiments with consistent results. Note: the antibodies against these 16p11.2 proteins have been verified through sgRNA knockdown plasmids (Figure S18 in Supplement 1), and the full-scaled gel plot of Western blotting is shown in Figure S19 in Supplement 1. 3C, chromatin conformation capture; CNV, copy number variation; eQTL, expression quantitative trait locus; GTEx, Genotype-Tissue Expression; RNA-Seq, RNA sequencing; sgRNA, single guide RNA; TPM, transcripts per million.

experimental verification of the regulatory effect of rs4420550 on gene expression. The 589-bp sequence surrounding rs4420550 was cloned into pGL3-promoter vectors, and the enhancer luciferase activities of vectors containing both alleles of rs4420550 were measured in different cell lines. Consistent with the eQTL associations, the homozygotes of risk allele [A/A] led to ~80% higher luciferase activities compared with the protective alleles [G/G] in both HEK293T cells (p = .0016) and SK-N-SH cells (p = .01) (Figure 1C). We also cloned the 477-bp region surrounding rs12691307 into pGL3-promoter vectors to examine its allelic effects on enhancer activities, but no difference was seen between different alleles (Figure S6 in Supplement 1).

The rs4420550 Locus Physically Interacts With *MAPK3* Promoter and *TAOK2* Promoter in the 3C Assay

Given the ~200-kb distance between rs4420550 and *MAPK3* in the genome, we hypothesized that the regulatory effect likely resulted from interaction between a distal enhancer and target gene promoters owing to chromatin looping. To investigate if the region surrounding rs4420550 acted as an enhancer element for *MAPK3* via this mechanism, we performed 3C assay. Briefly, a 16p11.2 regional-wide 3C in HEK293T cells was carried out using a series of PCR primers designed with sequences spanning from rs4420550 to the distal *MAPK3* promoter. Intriguingly, an increased interaction frequency was detected between the restriction fragment containing rs4420550 and that containing *MAPK3* promoter in our 3C library, demonstrating a direct physical interaction between rs4420550 and *MAPK3* promoter (Figure 2A; the RT-qPCR *C*t values are shown in Table S11 in Supplement 2). Moreover, in the 3C experiments, higher peaks of interaction frequencies were observed between the restriction fragment containing rs4420550 and that containing *TAOK2* promoter (Figure 2A), suggesting that rs4420550 likely affected expression of *TAOK2* through the chromosome looping mechanism.

We also examined the potential physical interactions within the 16p11.2 region using public Hi-C data in human tissues (30,54) via the 3D Genome Browser (55). In the DLPFC and developing brain cortical plates, the DNA sequences spanning rs4420550 and *MAPK3* (and other 16p11.2 genes such as *TAOK2*) were located in the same large topologically associated domain (Figure S7 in Supplement 1). In the nonbrain tissues (e.g., adrenal gland, bladder, lung, pancreas), rs4420550 and those genes were also in the same topologically associated domain, suggesting that their physical interactions were not restricted to brains. These Hi-C data are in line with the eQTL analyses in human tissues and the 3C results in HEK293T cells.

Precise CRISPR/Cas9 Editing at rs4420550 Confirms Its Regulatory Effect on *MAPK3* and Other 16p11.2 Genes

The wild-type HEK293T cells carry [A/A] genotype at rs4420550. To gain direct evidence of the regulatory effect of rs4420550 on *MAPK3* mRNA expression, we employed homology-directed repair–mediated genome editing by CRISPR/Cas9 and thereby obtained HEK293T cells carrying homozygotes [G/G] genotype at rs4420550. T7EN1 assay was performed to ensure that no off-target effects were introduced by CRISPR/Cas9 (Figure S8 in Supplement 1 and Table S12 in Supplement 2). Four single cell–derived colonies of each genotype ([A/A] and [G/G]) at rs4420550 were expanded, and their total mRNAs were isolated. RT-qPCR of *MAPK3* and other 16p11.2 genes in these eight colonies was conducted. As expected, the *MAPK3* mRNA levels in rs4420550-[A/A] cells were higher than those in the rs4420550-[G/G] cells (p < .0001), and other 16p11.2 genes (e.g., *INO80E*, *KCTD13*, *TAOK2*, *ALDOA*, *HIRIP3*) were also expressed differently between cells of different genotypes (Figure S9 in Supplement 1). RNA-Seq was then performed on these 8 samples, and 48 differentially expressed genes (false discovery rate [FDR] < .05) with clustering of samples according to the rs4420550 genotypes were identified (Table S13 in Supplement 2). The expression changes of *MAPK3* showed the most significant association with rs4420550 ($p = 7.33 \times 10^{-11}$, FDR = 5.47×10^{-7}) (Figure 2B), followed by *INO80E* ($p = 8.50 \times 10^{-10}$, FDR = 5.47×10^{-7}). Notably, 12 of the 48 genes were located in the 16p11.2 region (FDR < .05) (Figure 2C). We then compared the association signals between rs4420550 and the genes within the proxy 16p11.2 CNV region in GTEx (38), Brain xQTL (36), and the RNA-Seq results of rs4420550-[A/A] and -[G/G] HEK293T cells. In all these datasets, the mRNA expression levels of *MAPK3* and *INO80E* were significantly affected by different alleles of rs4420550 with the same direction of allelic effects (Figure 2D).

We further confirmed that the MAPK3 protein levels were significantly higher in rs4420550-[A/A] cells compared with rs4420550-[G/G] cells (p = .0015) (Figure 2E). The endogenous proteins of INO80E and KCTD13 were undetectable in HEK293T cells (data not shown). The protein levels encoded by other 16p11.2 genes, such as *HIRIP3*, *TAOK2* and *ALDOA*, were also upregulated in rs4420550-[A/A] cells compared with rs4420550-[G/G] cells, despite the fact that TAOK2 did not reach nominal significance level (p = .086) (Figure 2E).

rs4420550 Affects Chromatin Accessibility at 16p11.2 Region

We then tested whether rs4420550-[A/A] and -[G/G] cells had different chromatin accessibility of the 16p11.2 region using the ATAC-Seq method. After quality control, we obtained an average of 11.7 million nonduplicated paired fragments only mapped to human reference genome GRCh37 autosomes and X-chromosome. After exclusion of blacklisted regions, an average of 34,490 narrow peaks called using MACS2 (https://pypi.org/project/MACS2/) remained. The results showed that the insert size between peaks was about ~180 bp or multiple 180 bp (Figure S10 in Supplement 1), so that Tn5 can only access the linker DNA rather than nucleosome-occupied regions (~147 bp). We also found that sequence tag intensity was significantly enriched around TSS in all samples (Figure S11 in Supplement 1). To analyze the chromatin accessibility differences between rs4420550-[A/A] and -[G/G] HEK293T cells, the read counts of 50,644 merged peaks were used as inputs and analyzed using the DESeq2 program. After comparison, we noticed substantial quantitative differences in peak signals at 16p11.2 between rs4420550-[A/A] and -[G/G]

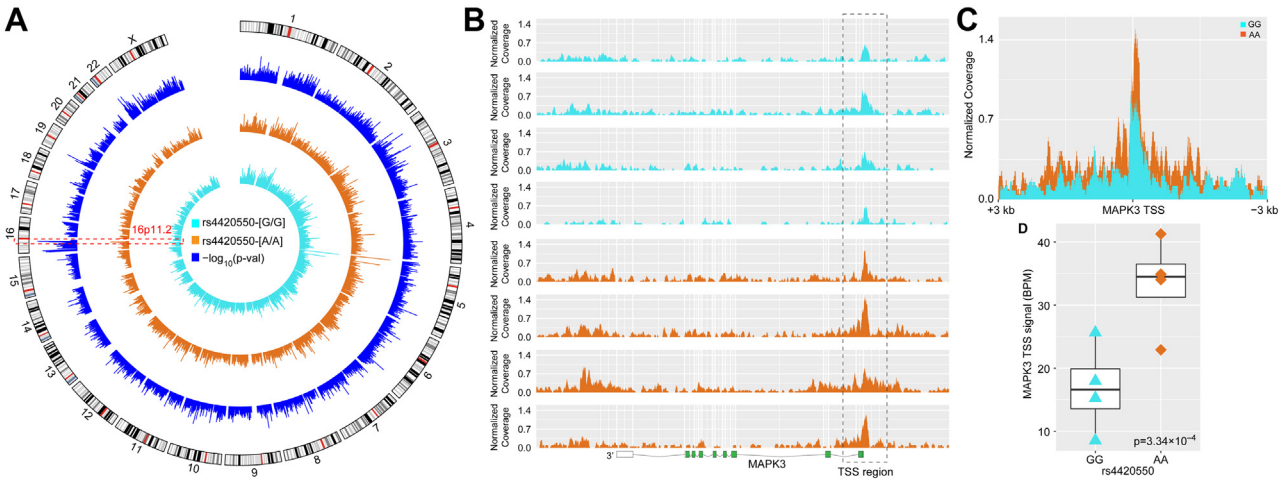


Figure 3. ATAC-Seq analysis of rs4420550-[A/A] and -[G/G] HEK293T cells. **(A)** Genome-wide chromatin accessibility in rs4420550-[A/A] and -[G/G] HEK293T cells. **(B)** The open chromatin of the *MAPK3* gene in the 8 rs4420550-[A/A] and -[G/G] HEK293T samples. **(C)** The chromatin accessibility in the TSS regulatory region of the *MAPK3* gene. **(D)** Boxplot of ATAC-Seq signal at the TSS regulatory region of *MAPK3* for the 8 rs4420550-[A/A] and -[G/G] HEK293T samples. ATAC-Seq, assay for transposase-accessible chromatin using sequencing; BPM, bins per million mapped reads; TSS, transcription start site.

cells (Figure 3A). One of the most significant differences across the genome was *MAZ* in the 16p11.2 region ($p = 7.28 \times 10^{-7}$) (Figure S12 in Supplement 1 and Table S14 in Supplement 2). Remarkably, in the TSS regulatory regions of *MAPK3*, the chromatin accessibility was significantly higher in the rs4420550-[A/A] cells compared with the rs4420550-[G/G] cells ($p = 3.34 \times 10^{-4}$) (Figure 3B–D). The chromatin accessibility at the *INO80E* TSS regulatory region was also higher in rs4420550-[A/A] cells ($p = 3.90 \times 10^{-4}$) (Figure S13 in Supplement 1).

To reveal the potential mechanisms explaining the impact of rs4420550 on gene expression, we screened potential transcription factor(s) that could bind the rs4420550 locus. By functional predictions using RegulomeDB (46) and HaploReg v4.1 (47), we found that the DNA sequences spanning rs4420550 likely bound some transcription factors, such as CREB3L1, PAX6, and HDAC2 (Figure S14 in Supplement 1). We then manipulated the expression of these three transcription factors respectively and found that overexpression of PAX6 or HDAC2 significantly increased *MAPK3* mRNA expression, whereas knockdown of them led to decreased *MAPK3* mRNA levels (Figure S15 in Supplement 1). We conducted electrophoretic mobility shift assay assays using nuclear extracts from wild-type HEK293T cells, HEK293T cells overexpressing PAX6, and HEK293T cells overexpressing HDAC2, but no specific shift bands were observed for either allele of rs4420550 (Figure S16 in Supplement 1). We then examined whether different alleles of rs4420550 affected gene expression by modulating the effects of HDAC2 or PAX6. Briefly, we overexpressed either HDAC2 or PAX6 in the HEK293T cells with rs4420550-[A/A] or -[G/G], but overexpression of either transcription factor did not result in any *MAPK3* mRNA expression differences between rs4420550-[A/A] and -[G/G] cells (Figure S17 in Supplement 1). We also examined whether the rs4420550 locus was highlighted in the ChIP-Seq analyses of diverse transcription factors in multiple cell lines or tissues using ENCODE dataset (44). However, rs4420550 was not implicated in the ChIP-Seq assays of any transcription factors (data not shown). Thus, the molecular mechanisms underlying the regulatory effects of rs4420550 remain unclear. We speculate that its regulatory function might involve mechanisms other than transcription factor binding (e.g., noncoding RNAs), and this hypothesis awaits further explorations.

rs4420550 Affects Cell Proliferation

Previous studies suggested that several genes in the 16p11.2 region, including *MAPK3*, could affect cell proliferation (8,56); we therefore investigated whether cells carrying different genotypes of rs4420550 showed different proliferation rates. By performing the CCK-8 assay, we observed that the HEK293T cell clones carrying rs4420550-[A/A] showed significantly higher cell proliferation rates compared with the clones carrying rs4420550-[G/G] ($p = .015$ at 48 hours and $p = .017$ at 72 hours) (Figure 4), further confirming the putative physiological impact of rs4420550.

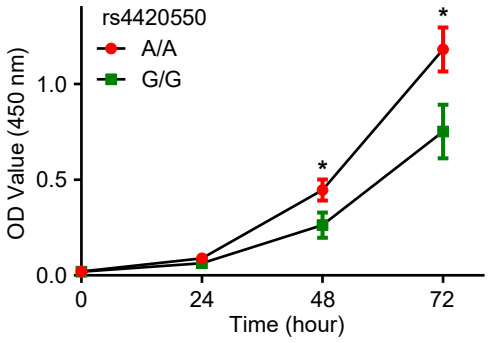


Figure 4. rs4420550 affects HEK293T cell proliferation abilities. The blank absorbance values from wells without cells were subtracted from all absorbance values. Error bars represent SEM of four biological replicates. Data are representative of three independent experiments with consistent results. *$p < .05$, rs4420550-[A/A] versus rs4420550-[G/G]. OD, optical density.

DISCUSSION

Through an integrative functional analysis combining eQTL associations, epigenome annotations, and experimental validations, we identified a functional risk variation rs4420550 in the schizophrenia-associated 16p11.2 locus. Although not showing the most significant association with schizophrenia in GWAS (1), rs4420550 likely locates in the DNA enhancer region and exerts regulatory effects on transcription of multiple genes in the 16p11.2 region. While our results suggest that a number of genes are affected by rs4420550, *MAPK3* remains those most significantly affected. In addition, regulation of the *INO80E* mRNA by rs4420550 is also consistently observed in multiple analyses, and this gene previously showed significant eQTL associations with schizophrenia genetic risk in a transcriptome study of human fetal brains (57).

Previous studies have shown that *MAPK3* mRNA expression is significantly higher in the DLPFC of patients with schizophrenia compared with controls (52), and elevated synaptic ERK-1 (encoded by *MAPK3*) activation leads to long-term memory deficits in mice (53). Conversely, lacking *Mapk3* can result in enhanced synaptic plasticity in the striatum and facilitate striatal-mediated learning and memory (58). Additionally, inhibition of ERK-1 signaling rescues the pathophysiological alterations observed in mice carrying the 16p11.2 microduplication (9). These studies suggest that higher *MAPK3* expression (or ERK-1 level/activity) is a risk factor for schizophrenia and other neurodevelopmental disorders. *INO80E* is a component of chromatin remodeling complex (59), which modulates nucleosome spacing and sliding in an ATP (adenosine triphosphate)-dependent manner (60), and regulates transcription during cortical neurogenesis (61–63).

Despite the fact that consistent impacts of rs4420550 on *MAPK3* and *INO80E* expression have been reported in a series of analyses, inconsistencies exist regarding its effects on other 16p11.2 genes. For example, while many genes at 16p11.2 showed different expression levels between the rs4420550-[A/A] and -[G/G] HEK293T cells, not all of them showed evidence of eQTL associations with rs4420550 in human brains (e.g., *HIRIP3* and *ALDOA* in Figure 2D). Given that expression of a gene is usually regulated by multiple independent or correlated genetic variations, the eQTL associations of a gene revealed in currently available datasets likely reflect the "summed" effects of multiple genetic variations. Therefore, additional functional variations, which either interact with rs4420550 or directly affect transcription, may explain the nonsignificant eQTL associations between some 16p11.2 genes and rs4420550.

In line with hypothesis, we acknowledge the potentially incomplete assessment of other plausible functional SNPs in the 16p11.2 locus as a limitation of the current study. Actually, existence of multiple functional variations (either independent or correlated with each other) in one GWAS risk locus has been commonly seen (33,64–68). In fact, functional predictions of 16p11.2 variations using HaploReg v4.1 (47) suggested that several SNPs are potentially functional. For example, rs12716973, in high LD with rs4420550 and in the active TSS region of *KCTD13*, has 22 bound proteins in the ChIP-Seq assays of human tissues as implemented in HaploReg dataset (Figure S4 in Supplement 1). Further validation of the physiological impact of these potentially functional variations is needed.

In summary, we have performed a detailed molecular characterization of a potential functional variation within the schizophrenia-associated 16p11.2 region. Future studies of rs4420550 using human induced pluripotent stem cells, and reprogrammed cells via genome editing, may provide essential insights into its relevant schizophrenia pathophysiology (69,70).

ACKNOWLEDGMENTS AND DISCLOSURES

This work was supported by National Natural Science Foundation of China Grant Nos. 81722019 (to ML), 31701133 (to XX), and 81871067 (to HC); Strategic Priority Research Program of the Chinese Academy of Sciences Grant Nos. XDB13000000 (to ML) and XDB02020003 (to Y-GY); Innovative Research Team of Science and Technology Department of Yunnan Province Grant No. 2019HC004 (to X-JL); the Chinese Academy of Sciences Western Light Program (to XX); the Youth Innovation Promotion Association, Chinese Academy of Sciences (to XX); the CAS Pioneer Hundred Talents Program (to ML); and the 1000 Young Talents Program (to ML). Data were generated as part of the CommonMind Consortium supported by funding from Takeda Pharmaceuticals Company Limited, F. Hoffman-La Roche and National Institutes of Health Grants Nos. 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 National Institutes of Health 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 National Institute of Mental Health Human Brain Collection Core. CommonMind Consortium 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), Enrico Domenici, Laurent Essioux (F. Hoffman-La Roche), Lara Mangravite, Mette Peters (Sage Bionetworks), Thomas Lehner, Barbara Lipska (National Institute of Mental Health). Data were generated as part of the PsychENCODE Consortium, supported by Grant Nos. U01MH103392, U01MH103365, U01MH103346, U01MH103340, U01MH103339, R21MH109956, R21MH105881, R21MH105853, R21MH103877, R21MH102791, R01MH111721, R01MH110928, R01MH110927, R01MH110926, R01MH110921, R01MH110920, R01MH110905, R01MH109715, R01MH109677, R01MH105898, R01MH105898, R01MH094714, P50MH106934, U01MH116488, U01MH116487, U01MH116492, U01MH116489, U01MH116438, U01MH116441, U01MH116442, R01MH114911, R01MH114899, R01MH114901, R01MH117293, R01MH117291, and R01MH117292 awarded to Schahram Akbarian (Icahn School of Medicine at Mount Sinai), Gregory Crawford (Duke University), Stella Dracheva (Icahn School of Medicine at Mount Sinai), Peggy Farnham (University of Southern California), Mark Gerstein (Yale University), Daniel Geschwind (University of California, Los Angeles), Fernando Goes (Johns Hopkins University), Thomas M. Hyde (Lieber Institute for Brain Development), Andrew Jaffe (Lieber Institute for Brain Development), James A. Knowles (University of Southern California), Chunyu Liu (State University of New York Upstate Medical University), Dalila Pinto (Icahn School of Medicine at Mount Sinai), Panos Roussos (Icahn School of Medicine at Mount Sinai), Stephan Sanders (University of California, San Francisco), Nenad Sestan (Yale University), Pamela Sklar (Icahn School of Medicine at Mount Sinai), Matthew State (University of California, San Francisco), Patrick Sullivan (University of North Carolina), Flora Vaccarino (Yale University), Daniel Weinberger (Lieber Institute for Brain Development), Sherman Weissman (Yale University), Kevin White (University of Chicago), Jeremy Willsey (University of California, San Francisco), and Peter Zandi (Johns Hopkins University).

We thank Yongbin Chen and Cuiping Yang (Kunming Institute of Zoology) for providing assistance during the experiments.

HC, XX, and ML designed the study and interpreted the results. HC and XC conducted the primary experiments. H-JL conducted the RNA-Seq and ATAC-Seq analysis, as well as other bioinformatics analyses. W-PL, L-JZ, C-YZ, LW, J-YW, J-WL, and X-LM provided assistance during the experiments. Y-GY and X-JL provided advice and discussion during the design and practice of the current study. XX and ML drafted the manuscript, and all authors contributed to the final version of the article.

The authors report no biomedical financial interests or potential conflicts of interest.

ARTICLE INFORMATION

From the Key Laboratory of Animal Models and Human Disease Mechanisms of the Chinese Academy of Sciences and Yunnan Province (HC, XC, H-JL, W-PL, L-JZ, C-YZ, J-YW, J-WL, X-LM, LW, Y-GY, X-JL, ML, XX) and KIZ-CUHK Joint Laboratory of Bioresources and Molecular Research in Common Diseases (Y-GY, X-JL, ML), Kunming Institute of Zoology, Chinese Academy of Sciences; Kunming College of Life Science (XC, H-JL, W-PL, L-JZ, C-YZ, J-YW), University of Chinese Academy of Sciences; and Center for Excellence in Animal Evolution and Genetics (X-JL), Chinese Academy of Sciences, Kunming; and the CAS Center for Excellence in Brain Science and Intelligence Technology (Y-GY, ML), Chinese Academy of Sciences, Shanghai, China.

HC, XC, and H-JL contributed equally to this work.

Address correspondence to Xiao Xiao, Ph.D., at xiaoxiao2@mail.kiz.ac.cn, or Ming Li, Ph.D., at limingkiz@mail.kiz.ac.cn.

Received Mar 5, 2020; revised and accepted Sep 17, 2020.

Supplementary material cited in this article is available online at https://doi.org/10.1016/j.biopsych.2020.09.016.

REFERENCES

1. Schizophrenia Working Group of the Psychiatric Genomics Consortium (2014): Biological insights from 108 schizophrenia-associated genetic loci. Nature 511:421–427.
2. Pardinas AF, Holmans P, Pocklington AJ, Escott-Price V, Ripke S, Carrera N, *et al.* (2018): Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. Nat Genet 50:381–389.
3. Hu Z, Xiao X, Zhang Z, Li M (2019): Genetic insights and neurobiological implications from NRXN1 in neuropsychiatric disorders. Mol Psychiatry 24:1400–1414.
4. Marshall CR, Howrigan DP, Merico D, Thiruvahindrapuram B, Wu W, Greer DS, *et al.* (2017): Contribution of copy number variants to schizophrenia from a genome-wide study of 41,321 subjects. Nat Genet 49:27–35.
5. McCarthy SE, Makarov V, Kirov G, Addington AM, McClellan J, Yoon S, *et al.* (2009): Microduplications of 16p11.2 are associated with schizophrenia. Nat Genet 41:1223–1227.
6. Wu Y, Li X, Liu J, Luo XJ, Yao YG (2020): SZDB2.0: An updated comprehensive resource for schizophrenia research. Hum Genet 139:1285–1297.
7. Chang H, Li L, Li M, Xiao X (2017): Rare and common variants at 16p11.2 are associated with schizophrenia. Schizophr Res 184:105–108.
8. Golzio C, Willer J, Talkowski ME, Oh EC, Taniguchi Y, Jacquemont S, *et al.* (2012): KCTD13 is a major driver of mirrored neuroanatomical phenotypes of the 16p11.2 copy number variant. Nature 485:363–367.
9. Blizinsky KD, Diaz-Castro B, Forrest MP, Schurmann B, Bach AP, Martin-de-Saavedra MD, *et al.* (2016): Reversal of dendritic phenotypes in 16p11.2 microduplication mouse model neurons by pharmacological targeting of a network hub. Proc Natl Acad Sci U S A 113:8520–8525.
10. Richter M, Murtaza N, Scharrenberg R, White SH, Johanns O, Walker S, *et al.* (2019): Altered TAOK2 activity causes autism-related neurodevelopmental and cognitive abnormalities through RhoA signaling. Mol Psychiatry 24:1329–1350.
11. Escamilla CO, Filonova I, Walker AK, Xuan ZX, Holehonnur R, Espinosa F, *et al.* (2017): Kctd13 deletion reduces synaptic transmission via increased RhoA. Nature 551:227–231.
12. Yadav S, Oses-Prieto JA, Peters CJ, Zhou J, Pleasure SJ, Burlingame AL, *et al.* (2017): TAOK2 kinase mediates PSD95 stability and dendritic spine maturation through Septin7 phosphorylation. Neuron 93:379–393.
13. Penzes P, Cahill ME, Jones KA, VanLeeuwen JE, Woolfrey KM (2011): Dendritic spine pathology in neuropsychiatric disorders. Nat Neurosci 14:285–293.
14. Forrest MP, Parnell E, Penzes P (2018): Dendritic structural plasticity and neuropsychiatric disease. Nat Rev Neurosci 19:215–234.
15. MacDonald ML, Alhassan J, Newman JT, Richard M, Gu H, Kelly RM, *et al.* (2017): Selective loss of smaller spines in schizophrenia. Am J Psychiatry 174:586–594.
16. Focking M, Lopez LM, English JA, Dicker P, Wolff A, Brindley E, *et al.* (2015): Proteomic and genomic evidence implicates the postsynaptic density in schizophrenia. Mol Psychiatry 20:424–432.
17. Osimo EF, Beck K, Reis Marques T, Howes OD (2019): Synaptic loss in schizophrenia: A meta-analysis and systematic review of synaptic protein and mRNA measures. Mol Psychiatry 24:549–561.
18. Berdenis van Berlekom A, Muflihah CH, Snijders G, MacGillavry HD, Middeldorp J, Hol EM, *et al.* (2020): Synapse pathology in schizophrenia: A meta-analysis of postsynaptic elements in postmortem brain studies. Schizophr Bull 46:374–386.
19. Steinberg S, de Jong S, Mattheisen M, Costas J, Demontis D, Jamain S, *et al.* (2014): Common variant at 16p11.2 conferring risk of psychosis. Mol Psychiatry 19:108–114.
20. Yang Z, Cai X, Qu N, Zhao L, Zhong BL, Zhang SF, *et al.* (2020): Identification of a functional 339-bp Alu polymorphism in the schizophrenia-associated locus at 10q24.32. Zool Res 41:84–89.
21. Yang Z, Zhou D, Li H, Cai X, Liu W, Wang L, *et al.* (2020): The genome-wide risk alleles for psychiatric disorders at 3p21.1 show convergent effects on mRNA expression, cognitive function and mushroom dendritic spine. Mol Psychiatry 25:48–66.
22. Edwards SL, Beesley J, French JD, Dunning AM (2013): Beyond GWASs: Illuminating the dark road from association to function. Am J Hum Genet 93:779–797.
23. Li M, Jaffe AE, Straub RE, Tao R, Shin JH, Wang Y, *et al.* (2016): A human-specific AS3MT isoform and BORCS7 are molecular risk factors in the 10q24.32 schizophrenia-associated locus. Nat Med 22:649–656.
24. Liu W, Li W, Cai X, Yang Z, Li H, Su X, *et al.* (2020): Identification of a functional human-unique 351-bp Alu insertion polymorphism associated with major depressive disorder in the 1p31.1 GWAS risk loci. Neuropsychopharmacology 45:1196–1206.
25. Cai X, Yang Z-H, Li H-J, Xiao X, Li M, Chang H (2020): A human-specific schizophrenia risk tandem repeat affects alternative splicing of a human-unique isoform AS3MTd2d3 and mushroom dendritic spine density [published online ahead of print Jul 14]. Schizophr Bull.
26. Zhang S, Zhang H, Zhou Y, Qiao M, Zhao S, Kozlova A, *et al.* (2020): Allele-specific open chromatin in human iPSC neurons elucidates functional disease variants. Science 369:561–565.
27. French JD, Edwards SL (2020): The role of noncoding variants in heritable disease. Trends Genet 36:880–891.
28. Song M, Yang X, Ren X, Maliskova L, Li B, Jones IR, *et al.* (2019): Mapping cis-regulatory chromatin contacts in neural cells links neuropsychiatric disorder risk variants to target genes. Nat Genet 51:1252–1262.
29. Girdhar K, Hoffman GE, Jiang Y, Brown L, Kundakovic M, Hauberg ME, *et al.* (2018): Cell-specific histone modification maps in the human frontal lobe link schizophrenia risk to the neuronal epigenome. Nat Neurosci 21:1126–1136.
30. Won H, de la Torre-Ubieta L, Stein JL, Parikshak NN, Huang J, Opland CK, *et al.* (2016): Chromosome conformation elucidates regulatory relationships in developing human brain. Nature 538:523–527.
31. Duan J, Shi J, Fiorentino A, Leites C, Chen X, Moy W, *et al.* (2014): A rare functional noncoding variant at the GWAS-implicated MIR137/MIR2682 locus might confer risk to schizophrenia and bipolar disorder. Am J Hum Genet 95:744–753.
32. Roussos P, Mitchell AC, Voloudakis G, Fullard JF, Pothula VM, Tsang J, *et al.* (2014): A role for noncoding variation in schizophrenia. Cell Rep 9:1417–1429.

33. Forrest MP, Zhang H, Moy W, McGowan H, Leites C, Dionisio LE, *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.
34. Song JHT, Lowe CB, Kingsley DM (2018): Characterization of a human-specific tandem repeat associated with bipolar disorder and schizophrenia. Am J Hum Genet 103:421–430.
35. Jaffe AE, Straub RE, Shin JH, Tao R, Gao Y, Collado-Torres L, *et al.* (2018): Developmental and genetic regulation of the human cortex transcriptome illuminate schizophrenia pathogenesis. Nat Neurosci 21:1117–1125.
36. Ng B, White CC, Klein HU, Sieberts SK, McCabe C, Patrick E, *et al.* (2017): An xQTL map integrates the genetic architecture of the human brain's transcriptome and epigenome. Nat Neurosci 20:1418–1426.
37. Fromer M, Roussos P, Sieberts SK, Johnson JS, Kavanagh DH, Perumal TM, *et al.* (2016): Gene expression elucidates functional impact of polygenic risk for schizophrenia. Nat Neurosci 19:1442–1453.
38. GTEx Consortium; Laboratory Data Analysis, and Coordinating Center—Analysis Working Group; Statistical Methods groups—Analysis Working Group; Enhancing GTEx groups; NIH Common Fund; NIH/NCI, *et al.* (2017): Genetic effects on gene expression across human tissues. Nature 550:204–213.
39. Wang D, Liu S, Warrell J, Won H, Shi X, Navarro FCP, *et al.* (2018): Comprehensive functional genomic resource and integrative model for the human brain. Science 362:eaat8464.
40. Zhu Z, Zhang F, Hu H, Bakshi A, Robinson MR, Powell JE, *et al.* (2016): Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. Nat Genet 48:481–487.
41. Wu Y, Zeng J, Zhang F, Zhu Z, Qi T, Zheng Z, *et al.* (2018): Integrative analysis of omics summary data reveals putative mechanisms underlying complex traits. Nat Commun 9:918.
42. Barrett JC, Fry B, Maller J, Daly MJ (2005): Haploview: Analysis and visualization of LD and haplotype maps. Bioinformatics 21:263–265.
43. Genomes Project Consortium, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, *et al.* (2015): A global reference for human genetic variation. Nature 526:68–74.
44. ENCODE Project Consortium (2012): An integrated encyclopedia of DNA elements in the human genome. Nature 489:57–74.
45. Roadmap Epigenomics Consortium, Kundaje A, Meuleman W, Ernst J, Bilenky M, Yen A, *et al.* (2015): Integrative analysis of 111 reference human epigenomes. Nature 518:317–330.
46. Boyle AP, Hong EL, Hariharan M, Cheng Y, Schaub MA, Kasowski M, *et al.* (2012): Annotation of functional variation in personal genomes using RegulomeDB. Genome Res 22:1790–1797.
47. Ward LD, Kellis M (2012): HaploReg: A resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. Nucleic Acids Res 40:D930–D934.
48. Naumova N, Smith EM, Zhan Y, Dekker J (2012): Analysis of long-range chromatin interactions using Chromosome Conformation Capture. Methods 58:192–203.
49. Hagege H, Klous P, Braem C, Splinter E, Dekker J, Cathala G, *et al.* (2007): Quantitative analysis of chromosome conformation capture assays (3C-qPCR). Nat Protoc 2:1722–1733.
50. Gupta RM, Hadaya J, Trehan A, Zekavat SM, Roselli C, Klarin D, *et al.* (2017): A genetic variant associated with five vascular diseases is a distal regulator of Endothelin-1 gene expression. Cell 170:522–533.e15.
51. Love MI, Huber W, Anders S (2014): Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15:550.
52. Gandal MJ, Zhang P, Hadjimichael E, Walker RL, Chen C, Liu S, *et al.* (2018): Transcriptome-wide isoform-level dysregulation in ASD, schizophrenia, and bipolar disorder. Science 362:eaat8127.
53. Seese RR, Maske AR, Lynch G, Gall CM (2014): Long-term memory deficits are associated with elevated synaptic ERK1/2 activation and reversed by mGluR5 antagonism in an animal model of autism. Neuropsychopharmacology 39:1664–1673.
54. Schmitt AD, Hu M, Jung I, Xu Z, Qiu Y, Tan CL, *et al.* (2016): A compendium of chromatin contact maps reveals spatially active regions in the human genome. Cell Rep 17:2042–2059.
55. Wang Y, Song F, Zhang B, Zhang L, Xu J, Kuang D, *et al.* (2018): The 3D Genome Browser: A web-based browser for visualizing 3D genome organization and long-range chromatin interactions. Genome Biol 19:151.
56. Gusev A, Mancuso N, Won H, Kousi M, Finucane HK, Reshef Y, *et al.* (2018): Transcriptome-wide association study of schizophrenia and chromatin activity yields mechanistic disease insights. Nat Genet 50:538–548.
57. Walker RL, Ramaswami G, Hartl C, Mancuso N, Gandal MJ, de la Torre-Ubieta L, *et al.* (2019): Genetic control of expression and splicing in developing human brain informs disease mechanisms. Cell 179:750–771.e22.
58. Mazzucchelli C, Vantaggiato C, Ciamei A, Fasano S, Pakhotin P, Krezel W, *et al.* (2002): Knockout of ERK1 MAP kinase enhances synaptic plasticity in the striatum and facilitates striatal-mediated learning and memory. Neuron 34:807–820.
59. Ayala R, Willhoft O, Aramayo RJ, Wilkinson M, McCormack EA, Ocloo L, *et al.* (2018): Structure and regulation of the human INO80-nucleosome complex. Nature 556:391–395.
60. Clapier CR, Iwasa J, Cairns BR, Peterson CL (2017): Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes. Nat Rev Mol Cell Biol 18:407–422.
61. Morrison AJ, Shen X (2009): Chromatin remodelling beyond transcription: The INO80 and SWR1 complexes. Nat Rev Mol Cell Biol 10:373–384.
62. Ueda T, Watanabe-Fukunaga R, Ogawa H, Fukuyama H, Higashi Y, Nagata S, *et al.* (2007): Critical role of the p400/mDomino chromatin-remodeling ATPase in embryonic hematopoiesis. Genes Cells 12:581–592.
63. Sokpor G, Castro-Hernandez R, Rosenbusch J, Staiger JF, Tuoc T (2018): ATP-dependent chromatin remodeling during cortical neurogenesis. Front Neurosci 12:226.
64. Chatterjee S, Kapoor A, Akiyama JA, Auer DR, Lee D, Gabriel S, *et al.* (2016): Enhancer variants synergistically drive dysfunction of a gene regulatory network in hirschsprung disease. Cell 167:355–368.e10.
65. Huo Y, Li S, Liu J, Li X, Luo XJ (2019): Functional genomics reveal gene regulatory mechanisms underlying schizophrenia risk. Nat Commun 10:670.
66. He H, Li W, Liyanarachchi S, Srinivas M, Wang Y, Akagi K, *et al.* (2015): Multiple functional variants in long-range enhancer elements contribute to the risk of SNP rs965513 in thyroid cancer. Proc Natl Acad Sci U S A 112:6128–6133.
67. Roman TS, Marvelle AF, Fogarty MP, Vadlamudi S, Gonzalez AJ, Buchkovich ML, *et al.* (2015): Multiple hepatic regulatory variants at the GALNT2 GWAS locus associated with high-density lipoprotein cholesterol. Am J Hum Genet 97:801–815.
68. Wu Y, Bi R, Zeng C, Ma C, Sun C, Li J, *et al.* (2019): Identification of the primate-specific gene BTN3A2 as an additional schizophrenia risk gene in the MHC loci. EBioMedicine 44:530–541.
69. Falk A, Heine VM, Harwood AJ, Sullivan PF, Peitz M, Brustle O, *et al.* (2016): Modeling psychiatric disorders: From genomic findings to cellular phenotypes. Mol Psychiatry 21:1167–1179.
70. Hoffman GE, Schrode N, Flaherty E, Brennand KJ (2019): New considerations for hiPSC-based models of neuropsychiatric disorders. Mol Psychiatry 24:49–66.