Contents lists available at ScienceDirect



EBioMedicine

**EBioMedicine** 

Published by THE LANCET

journal homepage: www.ebiomedicine.com

# Identification of the primate-specific gene *BTN3A2* as an additional schizophrenia risk gene in the MHC loci



Yong Wu <sup>a,b</sup>, Rui Bi <sup>a</sup>, Chunhua Zeng <sup>a,b</sup>, Changguo Ma <sup>a,b</sup>, Chunli Sun <sup>b,c</sup>, Jingzheng Li <sup>b,c</sup>, Xiao Xiao <sup>a</sup>, Ming Li <sup>a,b,e</sup>, Deng-Feng Zhang <sup>a</sup>, Ping Zheng <sup>b,c,d,f</sup>, Nengyin Sheng <sup>b,c,d,\*</sup>, Xiong-Jian Luo <sup>a,b,d,\*</sup>, Yong-Gang Yao <sup>a,b,e,f,\*</sup>

<sup>a</sup> Key Laboratory of Animal Models and Human Disease Mechanisms of the Chinese Academy of Sciences & Yunnan Province, Kunming Institute of Zoology, Kunming, Yunnan 650223, China

<sup>b</sup> Kunming College of Life Science, University of Chinese Academy of Sciences, Kunming, Yunnan 650204, China

<sup>c</sup> State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, Yunnan 650223, China

<sup>d</sup> Center for Excellence in Animal Evolution and Genetics, Chinese Academy of Sciences, Kunming, Yunnan 650223, China

<sup>e</sup> CAS Center for Excellence in Brain Science and Intelligence Technology, Chinese Academy of Sciences, Shanghai 200031, China

<sup>f</sup> KIZ/CUHK Joint Laboratory of Bioresources and Molecular Research in Common Diseases, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, Yunnan 650223, China

#### ARTICLE INFO

Article history: Received 9 March 2019 Received in revised form 26 April 2019 Accepted 3 May 2019 Available online 24 May 2019

Keywords: Schizophrenia Integrative analysis BTN3A2 Brain development Glutamatergic synapse

#### ABSTRACT

*Background:* Schizophrenia is a complex mental disorder resulting in poor life quality and high social and economic burden. Despite the fact that genome-wide association studies (GWASs) have successfully identified a number of risk loci for schizophrenia, identifying the causal genes at the risk loci and elucidating their roles in disease pathogenesis remain major challenges.

*Methods:* The summary data-based Mendelian randomization analysis (SMR) was used to integrate a large-scale GWAS of schizophrenia with brain expression quantitative trait loci (eQTL) data and brain methylation expression quantitative trait loci (meQTL) data, to identify novel risk gene(s) for schizophrenia. We then analyzed the mRNA expression and methylation statuses of the gene hit *BTN3A2* during the early brain development. Electrophysiological analyses of CA1 pyramidal neurons were performed to evaluate the excitatory and inhibitory synaptic activity after overexpression of BTN3A2 in rat hippocampal slices. Cell surface binding assay was used to test the interaction of BTN3A2 and neurexins.

*Findings*: We identified *BTN3A2* as a potential risk gene for schizophrenia. The mRNA expression and methylation data showed that *BTN3A2* expression in human brain is highest post-natally. Further electrophysiological analyses of rat hippocampal slices showed that BTN3A2 overexpression specifically suppressed the excitatory synaptic activity onto CA1 pyramidal neurons, most likely through its interaction with the presynaptic adhesion molecule neurexins.

*Interpretation:* Increased expression of BTN3A2 might confer risk for schizophrenia by altering excitatory synaptic function. Our result constitutes a paradigm for distilling risk gene using an integrative analysis and functional characterization in the post-GWAS era.

*Fund:* This study was supported by the Strategic Priority Research Program (B) of the Chinese Academy of Sciences (XDB02020003 to Y-GY), the National Natural Science Foundation of China (31730037 to Y-GY), and the Bureau of Frontier Sciences and Education, Chinese Academy of Sciences (QYZDJ-SSW-SMC005 to Y-GY).

© 2019 Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http:// creativecommons.org/licenses/by-nc-nd/4.0/).

#### 1. Introduction

Schizophrenia is a complex, highly heritable and heterogeneous disease characterized by episodic psychosis and altered cognitive function [1]. It affects approximately 0.5–1% of the world population and causes significant global burdens. Its morbidity and mortality have been both

\* Corresponding authors at: Kunming College of Life Science, University of Chinese Academy of Sciences, Kunming, Yunnan 650204, China.

E-mail addresses: shengnengyin@mail.kiz.ac.cn (N. Sheng),

very high that this disease is widely referred to as the "cancer of mental illness" [2]. So far, the pathophysiology of schizophrenia has not been well elucidated, probably because of the complexity and heterogeneity of the psychopathology and the associated cognitive impairments [1]. An increasing evidence from epidemiology has linked this disorder to prenatal and early postnatal life, and evidence from schizophrenia genetic studies has also suggested that many risk genes of this illness enriched in early neurodevelopmental process such as neuronal differentiation and migration [2]. The 'neurodevelopmental model' of schizophrenia posited that perturbations in 'normal' brain development, which were associated with molecular changes in the developmental

2352-3964/© 2019 Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

luoxiongjian@mail.kiz.ac.cn (X.-J. Luo), yaoyg@mail.kiz.ac.cn (Y.-G. Yao).

#### Research in context

Evidence before this study

Schizophrenia is a complex, highly heritable and heterogeneous disease with estimated heritability ranging up to 80%. During the past decade, the genome-wide association studies (GWASs) have identified numerous schizophrenia-associated variants and loci in world-wide populations, but most of the risk single-nucleotide polymorphisms (SNPs) or variants identified by GWASs were located in intergenic, intronic or other non-coding regions. To identify the causal genes at the reported risk loci and to elucidate how these risk genes influence the pathogenesis of schizophrenia remains a daunting task.

#### Added value of this study

By integrating a large-scale schizophrenia GWAS data and brain prefrontal cortex expression quantitative trait loci (eQTL) data, we identified *BTN3A2* as a potential schizophrenia risk gene (independent of the *C4* gene) in the MHC region. We found that *BTN3A2* expression in human brain is highest post-natally. Overexpression of this gene specifically suppressed the excitatory synaptic activity onto CA1 pyramidal neurons and decreased the presynaptic release probability, most likely through its interaction with the presynaptic adhesion molecule neurexins. These results suggested that increased expression of BTN3A2 might confer risk for schizophrenia by altering excitatory synaptic function.

#### Implications of all the available evidence

We identified *BTN3A2* as a risk gene for schizophrenia via integrative analysis of GWAS, eQTL data and brain methylation expression quantitative trait loci (meQTL) data. The role of this primatespecific gene was further characterized by electrophysiological experiments and cell surface binding assays. Our result constituted a paradigm for distilling risk gene using an integrative analysis in the post-GWAS era and provided helpful information for understanding the complex pathogenesis of schizophrenia.

stage and affected by environmental factors, might lead to an altered brain developmental trajectory and consequently cause schizophrenia in early adulthood [2]. However, the underlying neurodevelopmental pathogenesis of schizophrenia has not been well determined.

Due to a high heritability of schizophrenia, many researchers aimed at parsing its pathogenesis through genetic analyses, in particular the genome-wide association studies (GWASs). During the past decade, these GWASs have identified numerous schizophrenia-associated variants and loci in world-wide populations [3-11]. In 2014, the Schizophrenia Working Group of the Psychiatric Genomics Consortium (PGC) carried out a large-scale schizophrenia GWAS (PGC2 GWAS) [10] and reported 108 independent risk loci based on a multi-stage GWAS of >150,000 samples. Through integrating the genetic association signals from the schizophrenia GWAS and expression quantitative trait loci (eQTL), new susceptibility genes have been identified [12-16]. Most recently, Pardiñas et al. [7] carried out another large-scale schizophrenia GWAS (CLOZUK + PGC GWAS) and reported 145 independent risk loci based on a multi-stage GWAS of >100,000 samples. Although GWASs have identified many loci associated with schizophrenia, there has no one-to-one Mendelian mapping between these schizophrenia risk alleles and diagnosis [17]. Instead, schizophrenia is truly complex and appears to result from a myriad of genetic variants each exerting small effects on the overall disease risk, conforming closely to a classical polygenic model [4,18]. Indeed, our recent whole-genome sequencing of monozygotic twins discordant for schizophrenia revealed multiple genetic risk factors for this disease [19]. Approximately 90% of the single-nucleotide polymorphisms (SNPs) or variants identified by GWASs were located in intergenic, intron or other non-coding regions [10]. Elucidating how these risk loci influence the pathogenesis of schizophrenia remains a daunting task. Previous studies showed that most complex disease-associated variants confer risk for the illnesses by acting as eQTL to influence gene expression in *cis* or in *trans* [20–23]. Moreover, pathway and integrative analyses showed that the schizophrenia risk genes identified by GWASs were enriched in neurodevelopment-associated pathways like glutamatergic neurotransmission and neuronal calcium signaling pathways, supporting the notation that disturbation of the neurodevelopment processes may play a pivotal role in the pathogenesis of schizophrenia [10,24].

In this study, we aimed to identify genes that are *cis*-regulated by nearby schizophrenia GWAS risk variants based on a two-stage integrative analysis, followed by functional characterization of the target gene. We identified *BTN3A2* as a potential novel schizophrenia risk gene, and found that increased expression of BTN3A2 might confer risk for schizophrenia by altering excitatory synaptic function.

#### 2. Materials and methods

#### 2.1. GWAS associations of schizophrenia

Many GWAS studies of schizophrenia have been conducted in the past decade [3,4,7,10,11]. Because the samples in these GWAS studies were (partially) overlapped [8,11,25] or the dataset were not publically released [3–6], we focused on the result of the recent GWAS study (CLOZUK + PGC GWAS) conducted by Pardiñas et al. [7] in the current integrative study. Briefly, Pardiñas et al. [7] collected genome-wide genotype information for 11,260 schizophrenia cases from the United Kingdom and 24,542 controls from public repositories or through collaborations. They combined these samples with 29,415 cases and 40,101 controls from the PGC2 GWAS [10] and performed meta-analysis, which resulted in a total of 105,318 individuals (including 40,675 schizophrenia cases and 64,643 controls) [7]. More details about sample description, genotyping and statistical analyses could be found in the original paper [7]. We downloaded the summary statistics of this meta-analysis [7] from http://walters.psycm.cf.ac.uk/.

#### 2.2. Brain eQTL and meQTL data

The BrainCloud dataset, which was generated using Illumina Human 49 K Oligo array (30,176 probes) and contained 268 human dorsolateral prefrontal cortex (DLPFC) samples collected from fetal (negative ages) through gerontal (80 years old) donors, was downloaded from the GEO database (https://www.ncbi.nlm.nih.gov/geo/; GSE30272) [26]. The corresponding SNP data for these subjects (268 individuals), generated using Illumina Infinium II 650 K or Illumina Infinium HD Gemini 1 M BeadChips, were downloaded at the dbGaP (https://www.ncbi.nlm.nih.gov/gap; phs000417.v1.p1) [26]. A total of 654,333 SNPs were successfully genotyped. More detailed information about the BrainCloud dataset can be found in the original paper [26].

For the eQTL analysis, we firstly excluded SNPs with a minor allele frequency (MAF) <0.01 (MAF was estimated on the basis of the 1000 Genomes data [27]) from further analyses. In addition, only probes for which the *P*-value of the top associated *cis*-eQTL (<1 Mb away from the probe) was  $<5 \times 10^{-8}$  were included in subsequent analyses, as SNPs showing a significant *P*-value would be strongly associated with the gene expression based on the basic assumption for a Mendelian randomization analysis that the instrumental variable has a strong effect on exposure [28]. In brief, a total of 542,091 SNPs and 4876 probes for eQTL analysis were used. Next, we used Plink (version v1.07) [29] to assess the association between each SNP and each probe. We excluded transcripts with a missing value of >5% of the total samples from the

analysis. The following minimum SNP cut-off values were used: per sample call rate should be at least 90%, per SNP call rate should be at least 90%, per SNP MAF should be at least 1%, and lack of significance (*P*-value >.05) should be observed in the Hardy-Weinberg equilibrium tests. SNP position and gene position were based on the hg19 genome assembly (http://www.ensembl.org/info/data/ftp/index.html).

We used the data of the Psychiatric Encyclopedia of DNA Elements (PsychENCODE) project, which collected prenatal and postnatal postmortem brain specimens across the developmental continuum to refine and characterize schizophrenia risk loci [30], as an independent brain eQTL dataset for validation. We downloaded the eQTL results with a false discovery rate < 0.05 and a filter requiring genes to have an expression >0.1 FPKM (fragments per kilobase per million mapped fragments) in at least 10 samples from the PsychENCODE Integrative Analysis website (http://resource.psychencode.org/) [31]. The PsychENCODE Integrative Analysis [31] contained adult brain prefrontal cortex data of 1387 individuals from the PsychENCODE [30] and the Genotype-Tissue Expression (GTEx, https://www.gtexportal.org) data [32]. Detailed information about the data collection and analysis process can be found in the original study [31].

We also retrieved the release v.7 summary statistics from the eQTL data of 14 brain-related tissues of the GTEx project [32]. The GTEx project is a resource database and associated tissue bank for the scientific community to study the relationship between genetic variation and gene expression in human tissues [32]. Currently, the release of GTEx (release v.7) contains a total of 11,688 samples from 53 tissues of 714 donors within an age range from 20 years to 79 years [32]. In brief, the RNA expression data was collected using Illumina TrueSeq RNA sequencing (non-stranded, ployA+ selection) and Affymetrix Human Gene 1.1 ST Expression Array. Whole genome/exome sequencing and Illumina SNP Array were then used to get the genotype data. More detailed summary and data processing method of the dataset can be found in the GTEx website (https://www.gtexportal.org) [32].

The binary format methylation quantitative trait loci (meQTLs) datasets were downloaded from the summary data-based Mendelian randomization analysis (SMR) website (http://cnsgenomics.com/software/smr) [28]. Briefly, Qi et al. [33] conducted a meta-analysis of meQTL data from ROSMAP [34], Hannon et al. [35] and Jaffe et al. [36]. These meQTL data in binary format were stored at the SMR website [28]. The detailed information about the sample size and method can be found in the original publication of meQTL meta-analysis [33] and the SMR website [28].

#### 2.3. SMR integrative analysis

The basic assumption of SMR [28] is if the expression level of a gene is influenced by an eQTL, then there will be differences in gene expression levels among individuals carrying different genotypes of the genetic variant. Then, if the expression level of the gene has an effect on a trait, we will observe differences in phenotype among the different genotype groups: that is, the genetic variant will also show an effect on the trait [28]. The SMR employs the Mendelian randomization to test for pleiotropic associations between gene expression and complex traits using eQTL and trait GWAS summary data [28]. This method can be used to integrate data from multiple-omics studies, such as GWAS data and meQTL data [28].

We integrated the eQTL data [26,32,33] and the meQTL data [33] with the GWAS summary data [7] to perform the SMR integrative analysis [28]. To prepare these data, a sparse BESD file was made by SMR (version 0.631) [28] using the –make-besd argument, which provides an efficient way to store the eQTL summary data in binary format. The sparse BESD file stored the data for eQTLs within 2 Mb of a probe in either direction, as well as SNPs within 1 Mb of any eQTLs with *P*-value <1  $\times 10^{-5}$  in either direction in the rest of the genome. For GWAS data [7],

we removed the SNP with a MAF <0.01 (MAF was estimated based on the 1000 Genomes data [27]). SMR (version 0.631) [28] was then used to integrate the eQTL (or meQTL) data (BESD format) and GWAS data [7] using the following settings: -bfile (individual-level SNP genotype data in PLINK binary format), --gwas-summary (GWAS summary data) and -beqtl-sumary (BESD format eQTL or meQTL data).

#### 2.4. Sherlock integrative analysis

Sherlock (http://sherlock.ucsf.edu/submit.html) is a powerful integrative method and computes the Bayes factor for each gene [37]. The basic assumption of Sherlock is that the expression level of a specific gene(s) may influence the risk of a disease. For a given gene, there may be many variants in the genome affecting its expression (expression SNPs, eSNPs). A change of genotype at any of these eSNPs would lead to mRNA expression change of the gene, which could in turn affect the disease risk. Sherlock computes the LBF score for each eSNP of the gene being tested to represent how strongly the SNP supports a functional role of the gene [37]. The total LBF score of a gene evaluates evidence supporting that the gene is associated versus not associated with the disease. The detailed modeling assumption and calculation theory are described in the original study [37].

We firstly used the standalone version of Sherlock [37] to integrate the BrainCloud eQTL data [26] with GWAS data [7] under the following settings: N\_expr = 268 (sample size of the BrainCloud eQTL data), is\_pheno\_binary = 1 (whether the phenotypic trait is binary (1) or quantitative (0)), N\_pheno = 105,318 (40,675 + 64,643 = 105,318), K = 0.01 (the disease prevalence). Other settings were set as the default arguments. The integration of GWAS data [7] with the GTEx eQTL data [32] were conducted using Sherlock webserver (http://sherlock.ucsf. edu/submit.html). We uploaded the GWAS data [7] and selected GTEx eQTL (release v.7) [32] data related to brain tissues as the eQTL data sets.

#### 2.5. Gene expression and methylation data

We explored the gene expression pattern from Human Brain Transcriptome (HBT, http://hbatlas.org/) [38] and the PsychENCODE Human Brain Development website (http://development. psychencode.org) [39]. HBT is a public database containing transcriptome data and associated metadata obtained from human brain tissues at various development stages [38]. A total of 1340 tissues collected from 57 postmortem human brains from 16 brain regions were sampled, and the exon-level transcriptome data was generated using Affymetrix GeneChip Human Exon 1.0 ST Arrays. The detailed information about the sample origin, sample size and data processing can be found in the original publication [38]. Recently, the PsychENCODE generated transcriptomic profiling data (mRNA-seq) of 607 histologically verified, high-quality tissue samples from 16 anatomical brain regions that were dissected from 41 brains with age range from 8 postconceptional weeks to 40 postnatal years [39]. These data were systematically analyzed and made queryable in the PsychENCODE website (http://development.psychencode.org) [39]. The detailed information about sample origin, sample size and data processing can be found in the original publication [39]. Jaffe et al. [36] generated DNA methylation data of the DLPFC brain tissues from 335 non-psychiatric controls across the lifespan and 191 patients with schizophrenia by using the Illumina HumanMethylation450 microarray, and they identified widespread changes in the transition from prenatal to postnatal life [36]. The sample size was expanded to 450 psychiatric disorder-free samples and 225 schizophrenia samples based on the downloaded data (GEO accession number GSE74193). Detailed information about the sample source, sample size and data processing can be found in the original publication [36] and the GEO dataset GSE74193.

## 2.6. CommonMind Consortium data and human induced pluripotent stem cells (hiPSC) expression data

The normalized DLPFC gene expression profile of schizophrenia and control subjects were downloaded from Synapse (https://www. synapse.org, syn5607581) from the CommonMind Consortium (CMC) project [40] with permissions. This dataset contained mRNA expression data of the post-mortem human brain specimens from three brain banks (the Icahn School of Medicine at Mount Sinai, the University of Pennsylvania and the University of Pittsburg). Briefly, the total RNA was isolated from 258 schizophrenia cases and 279 control subjects, and was sequenced by Illumina HiSeq 2500 after quality control by RNA integrity number. The cleaned sequencing reads were mapped to human reference genome hg19 (http://www.ensembl.org/info/data/ftp/index.html) and gene expression levels were quantified using log (CPM) (read counts per million total reads) [40].

hiPSC gene expression profiles [41] were downloaded at GEO database (GSE25673). This dataset contains the neurons differentiated from hiPSC that were derived from the primary human fibroblasts from 4 schizophrenia patients and 4 control subjects. Most of these hiPSC neurons were presumably glutamatergic and the others were GABAergic and dopaminergic [41]. Gene expression analyses were performed on 6-week-old hiPSC neurons using Affymetrix Human Gene 1.0ST arrays. Three independent neural differentiations from fibroblasts for each donor were compared. More details about the expression of hiPSC neurons can be found in the original paper [41].

#### 2.7. Antibodies and cell culture

The following antibodies and cells were used in this study: mouse monoclonal anti-Flag (Eno-Gene, E12-001), peroxidase-conjugated anti-mouse antibody (KPL, 474–1086), Alexa Fluor 488 conjugated don-key anti-mouse secondary antibody (Invitrogen, A21202). HEK293T cells were obtained from the Kunming Cell Bank of Kunming Institute of Zoology. Cells were maintained in DMEM medium (Gibco-BRL, 11965–092) supplemented with 10% fetal bovine serum (Gibco-BRL, 10099–141) at 37 °C in a humidified atmosphere incubator with 5% CO<sub>2</sub>.

#### 2.8. Plasmid construction and transfection

The full length human *BTN3A2* (BTN3A2-L, GenBank accession number NM\_001197247.2) and a truncated *BTN3A2* lacking transmembrane domain (BTN3A2-S, XM\_005248831.4) were cloned into Flag-tagged pCMV-3Tag-8 with *Hind*III and *Xho*I for constructing overexpression vectors pCMV-BTN3A2-L-Flag and pCMV-BTN3A2-S-Flag, respectively. Mouse *Nlgn1* (NM\_138666.4), *Nlgn3* (NM\_172932.4), *Nrxn1β* (NM\_001346959.1), *Nrxn1βA4(-)* (NM\_001346960.1) and *Nrxn3β* (NM\_001252074.2) were also cloned into pCMV-DsRed2, respectively. The cDNA of full length *BTN3A2* was also sub-cloned into pCAGGS-IRES-GFP vector for biolistic transfection. All the primers used for constructing plasmids were listed in Table S1. Each construct was verified by direct sequencing.

HEK239T cells were transfected using Lipofectamine<sup>™</sup> 3000 (Invitrogen L3000008) following the manufacture's protocol. In brief, cells were cultured in 6-well plates and allowed to grow to 70% confluence. Then cells in each well were transfected with a total volume of around 270 µL mixture containing 2.5 µg DNA, 250 µL Opti-MEM medium (Gibco-BRL, 31985–070) and 12.5 µL Lipofectamine<sup>™</sup> 3000. The cells were incubated with the transfection mixture for 6 h, and then the medium was changed into growth medium for 42 h before the harvest.

#### 2.9. Electrophysiology in rat hippocampal slice cultures

Electrophysiology in rat slice cultures was conducted as described in our previous studies [42,43]. Briefly, organotypic hippocampal slice cultures were made from postnatal 6 days – postnatal 8 days rats. The vectors were transfected on days in vitro (DIV) 2 during culturing by a Helios Gene Gun (Bio-Rad, America) with 1 µm DNA-coated gold particles. Slices were maintained at 34 °C with media changes every other day. On DIV 8, the responses of pyramidal neurons, which were identified by morphology and location, in area CA1 were simultaneously recorded from a fluorescent transfected neuron and a neighboring untransfected control one. Dual whole-cell recordings measuring evoked excitatory or inhibitory postsynaptic responses used artificial cerebrospinal fluid (ACSF) bubbled with 95% O<sub>2</sub>/5% CO<sub>2</sub> consisting of 119 µM NaCl, 2.5 µM KCl, 4 µM CaCl<sub>2</sub>, 4 µM MgSO<sub>4</sub>, 1 µM NaH<sub>2</sub>PO<sub>4</sub>, 26.2 µM NaHCO<sub>3</sub>, and 11 µM glucose. A total of 100 µM picrotoxin was added to block inhibitory currents when measuring AMPA and NMDA excitatory postsynaptic responses (EPSCs) or 10 µM NBQX/50 µM APV when measuring inhibitory postsynaptic responses. In addition, 2chloroadenosine (4 µM) was used to control epileptiform activity. A bipolar stimulation electrode was placed in stratum radiatum, and responses were evoked at 0.2 Hz. Peak  $\alpha$ -amino-3-hydroxy-5-methyl-4isoxazolepropionic acid receptor (AMPAR) currents were recorded at -70 mV, and N-methyl-D-aspartate receptor (NMDAR) current amplitudes 100 ms following the stimuli were recorded at +40 mV. Pairedpulse ratio was determined by delivering two stimuli 40 ms apart and dividing the peak response to stimulus 2 by the peak response to stimulus 1. Peak GABA receptor-mediated currents were recorded at 0 mV. All these data were analyzed by custom software (IGOR Pro).

#### 2.10. Cell surface binding assay

Cell supernatant containing the truncated isoform of BTN3A2 tagged by Flag (BTN3A-S-Flag) was produced from the HEK293T cells after transfection with expression vector pCMV-BTN3A2-S-Flag (Table S1) for 48 h. The supernatant was added to the transfected HEK293T cells overexpressing DsRed2, Nlgn1-DsRed2, Nlgn3-DsRed2, Nrxn1<sub>β</sub>-DsRed2, or Nrxn3<sub>B</sub>-DsRed2 (Table S1), respectively. After an incubation for 4 h, the medium was removed and cells were fixed with 4% paraformaldehyde for 15 min. Cells were then washed for three times with phosphate-buffered saline (PBS) and incubated at room temperature for 30 min in a blocking solution containing 5% bovine serum albumin in PBS. Mouse anti-Flag antibody was then added (Eno-Gene, E12-001, 1:500 ratio) and incubated for 2 h. The slides were washed three times with PBS and incubated with anti-mouse Alexa Fluor 488 fluorescent secondary antibody (Invitrogen, A21202, 1:500 ratio) for 1 h at room temperature to label BTN3A2. The slides were visualized under an Olympus FluoViewTM 1000 confocal microscope (Olympus, America).

#### 2.11. Statistical analysis

The mRNA expression levels of different genes in CMC DLPFC [17] and hiPSC neurons [41] were compared using the *Student's t*-test using R software (https://www.r-project.org/). For the electrophysiology experiment, all the statistical analysis was compared with the respective control neurons using two-tailed Wilcoxon signed-rank sum test. A *P*-value <.05 was considered to be statistically significant.

#### 2.12. Data availability

The datasets used and/or re-analyzed in this study are available at the original sources [7,17,26,31–33,38,39,41] and could be found in our webserver SZDB (www.szdb.org/integrative1.php) [12].

#### 2.13. Ethics statement

All animal experiments were performed in accordance with established protocols approved by the Institutional Review Board of Kunming Institute of Zoology, Chinese Academy of Sciences.

#### 3. Results

3.1. Identification of BTN3A2 as a schizophrenia risk gene by SMR integrative analysis

The overall study design and relevant rationale were shown in Fig. 1. The genotype and brain DLPFC expression data of 268 individuals were retrieved from BrainCloud [26], which contained 542,091 SNPs (each SNP with a MAF  $\geq$  0.01) and 4876 probes (each probe with a  $P_{eQTL} < 5 \times 10^{-8}$ ), and were used for the following analyses. As shown in Table 1, *BTN3A2*, a primate-specific gene [44,45], which was located in the extended major histocompatibility complex (MHC) region, had the most significant association with schizophrenia ( $P_{SMR} = 2.59 \times 10^{-11}$ , Bonferroni corrected *P*-value <.05). SNP rs1979, which was located in the 3'UTR region of *BTN3A2*, significantly affected the mRNA expression of *BTN3A2* ( $P_{eOLT} = 1.44 \times 10^{-17}$ , Fig. S1) [26] and was significantly associated with schizophrenia (*P*-value =  $1.15 \times 10^{-26}$ ) [7]. Another previously uncharacterized non-coding RNA (ncRNA) *LOC100131289* also survived the SMR association test (*P*<sub>SMR</sub> = 4.11  $\times 10^{-6}$ , Bonferroni corrected *P*-value <.05; Table 1). So far, there are no reports for potential function of this ncRNA according to the PubMed search with keyword "LOC100131289". This ncRNA has a relatively higher expression in digestive system, spleen, ovary / testis and skin, than that of brain (https://www.ncbi.nlm.nih.gov/gene/?term=LOC100131289). Further study should be carried out to discern its potential role in schizophrenia.

#### 3.2. Validation of BTN3A2 as a schizophrenia risk gene

We conducted a different integrative analysis (Sherlcok [37]) so as to use GWAS data (CLOZUK + PGC GWAS) [7] and eQTL data (BrainCloud eQTL) [26] to validate the results of the SMR integrative analysis. The



Fig. 1. Workflow of integrative analyses and functional assay in this study. The SMR [28] was used to integrate BrainCloud eQTL data [26] and CLOZUK + PGC GWAS data [7]. The Sherlock [37] was used to integrate BrainCloud [26] and GTEx brain tissue eQTL data [32] with CLOZUK + PGC GWAS data [7] to validate the result of SMR. Brain meQTL data [36] was also used to identify schizophrenia risk genes. Next, brain expression data [26,38,39,76] and methylation data [36] were explored to check the expression and methylation pattern during the prenatal and postnatal neurodevelopment. Data from GEO (GSE25673) [41] and CMC (syn5607581) [17] were used to plot the *BTN3A2* mRNA expression pattern between normal controls and schizophrenia cases. The electrophysiology assays in rat brain slice and cell surface binding assays were conducted to characterize the potential function of this gene.

#### Table 1

Integrative analysis (SMR) of schizophrenia GWAS (CLOZUK + PGC GWAS) and BrainCloud eQTL reveals BTN3A2 as a schizophrenia risk gene.

Probe	Gene	Supporting SNP (trans or cis)	P <sub>GWAS</sub> <sup>a</sup>	$P_{eQTL}^{b}$	P <sub>SMR</sub> <sup>c</sup>	$P_{\rm HEIDI}^{\rm d}$	$P_{adj}^{e}$
hHR012536	BTN3A2	rs1979 ( <i>cis</i> )	$1.15\times10^{-26}$	$1.44\times10^{-17}$	$2.59\times10^{-11}$	$2.64\times10^{-4}$	< 0.05
hHR010906	LOC100131289	rs175597 (cis)	$1.19 \times 10^{-30}$	$5.00 \times 10^{-7}$	$4.11 \times 10^{-6}$	$1.54 \times 10^{-2}$	< 0.05
hHC024109	GPR75-ASB3	rs3770407 (cis)	$6.10 \times 10^{-5}$	$4.53 \times 10^{-25}$	$1.84 \times 10^{-4}$	$6.50 \times 10^{-2}$	NS
hHC030929	THAP5	rs40919 (cis)	$1.02  imes 10^{-4}$	$2.65  imes 10^{-37}$	$1.98  imes 10^{-4}$	$3.73  imes 10^{-3}$	NS
hHC007598	NDUFAF7	rs1056021 (cis)	$1.59  imes 10^{-4}$	$5.58  imes 10^{-46}$	$2.58  imes 10^{-4}$	$3.72  imes 10^{-2}$	NS
hHC028708	THAP5	rs40926 (cis)	$9.95  imes 10^{-5}$	$2.14\times10^{-23}$	$2.89  imes 10^{-4}$	$1.98  imes 10^{-3}$	NS
hHR014285	LINC00339	rs2143103 (cis)	$7.04  imes 10^{-6}$	$3.69 \times 10^{-9}$	$3.51 \times 10^{-4}$	NA	NS
hHR021334	BOP1	rs3802232 (cis)	$8.51 \times 10^{-7}$	$2.74 \times 10^{-7}$	$3.77 \times 10^{-4}$	$7.05 \times 10^{-3}$	NS
hHA037802	SRA1	rs10042299 ( <i>cis</i> )	$\textbf{8.03}\times 10^{-5}$	$1.63\times10^{-15}$	$4.04\times10^{-4}$	$\textbf{2.26}\times 10^{-2}$	NS

<sup>a</sup> *P*-values for eQTL SNPs from the CLOZUK + PGC GWAS data [7].

<sup>b</sup> *P*-values for eQTL SNPs from the BrainCloud dataset [26].

<sup>c</sup> P-values from summary data-based Mendelian randomization (SMR) analysis.

<sup>d</sup> P-values from heterogeneity in dependent instruments (HEIDI) test. HEIDI uses multiple SNPs in a *cis*-eQTL region to distinguish pleiotropy from linkage. NA, not available.

<sup>e</sup> We used the Bonferroni correction to account for multiple testing, which resulted in a genome-wide significance level of  $P = 1.03 \times 10^{-5}$  (= 0.05/4876, 4876 probes were retained for SMR analysis). NS, not significant,

Sherlock method [37], which computes the Bayes factor for each gene, reasoned that for a disease-associated gene, any genetic variation that perturbs its expression is also likely to influence the disease risk. We could well validate the result that *BTN3A2* was a schizophrenia risk gene. By integrating CLOZUK + PGC GWAS data [7] with BrainCloud eQTL data [26] using Sherlock [37], we confirmed *BTN3A2* as a schizophrenia risk gene with *P*-value =  $1.68 \times 10^{-7}$  and the logarithm of Bayes factors (LBF) value of 9.38 (Table S2), which means that the posterior probability of the association between *BTN3A2* and schizophrenia was exp. (9.38) = 11,849 times more likely than the opposite hypothesis (no association). By integrating CLOZUK + PGC GWAS data [7] with eQTL data of the PsychENCODE brain PFC [31] and the GTEx 14 brain tissues [32], respectively, we also found that *BTN3A2* was significantly associated with schizophrenia (Tables S3 and S4).

The Sherlock [37] analysis of CLOZUK + PGC GWAS data [7] and GTEx eQTL data [32] also confirmed that *BTN3A2* was a schizophrenia risk gene (all *P*-value <1.00 × 10<sup>-5</sup>), with all LBF larger than 7 in all studied brain regions (Tables S5). All the integrative analysis data could be found in SZDB (www.szdb.org/integrative1.php) [12]. In addition, we investigated the eQTL of rs1979 in Braineac [46] and found that the risk allele G of rs1979 could significantly elevate the mRNA expression of *BTN3A2* (Fig. S3), which was consistent with our observations in BrainCloud [26] (Fig. S1) and GTEx [32] (Fig. S2) datasets.

Taken together, we validated that *BTN3A2* was a schizophrenia risk gene using independent integrative approaches (SMR [28] and Sherlock [37]), GWAS (CLOZUK + PGC GWAS) [7] and brain eQTLs (BrainCloud [26], PsychENCODE [31] and GTEx [32]) (Fig. 1). More importantly, we consistently detected the elevated mRNA expression levels of *BTN3A2* in brain tissues from carriers of the risk allele G of rs1979, suggesting that increased expression of *BTN3A2* was a potential risk factor for schizophrenia.

#### 3.3. BTN3A2 was independent of the C4 signal in the MHC region

Of note, integrating GWAS data [7] with GTEx eQTL data [32] also identified another MHC region gene *C4A* as a schizophrenia risk gene (Tables S6), which has been previously reported to mediate synapse elimination during postnatal development and therefore affects the risk of schizophrenia [47]. The MHC region is the most complex region in human genome because of its unintelligible linkage disequilibrium (LD), and this region exhibited the strongest genetic association with schizophrenia [7,10]. We therefore tested whether *BTN3A2* was an independent signal or was just a hitchhiking effect to the *C4A* gene by analyzing the LD relationship between the respective SNPs supporting the involvement of *BTN3A2* and *C4A* in the risk of schizophrenia. Interestingly, these SNPs were distributed in two evident blocks, with *BTN3A2* supporting SNPs in one block and *C4A* supporting SNPs in the other

block (Fig. S4). This result suggested that *BTN3A2* might be an independent schizophrenia risk gene in the MHC region.

## 3.4. Integrative analysis of methylation and GWAS data indicated BTN3A2 as a schizophrenia risk gene

Through integrative analysis, we found multiple GWAS SNPs conferring the risk of schizophrenia by affecting the mRNA expression of *BTN3A2* (Table S4). As elaborated previously, noncoding risk variants might leverage the expression of nearby genes by affecting the binding of *trans*-acting factors or changing the epigenetic conditions of a *cis* region, like histone and/or chromatin modification, or DNA methylation [2]. Based on these hypotheses and observations, we used SMR [28] to integrate schizophrenia CLOZUK + PGC GWAS data [7] with brainmMeta meQTL data [33], and observed significant association between *BTN3A2* and schizophrenia (Table 2). Taken together, our integrative analyses combining GWAS data [7] with gene expression [26,31,32] and methylation profiles [33] indicated that GWAS SNPs likely affected the mRNA expression levels of *BTN3A2* through modulating the methylation statuses of its adjacent sites.

### 3.5. Inverse pattern of methylation level and BTN3A2 mRNA expression level during fetal brain development

Transcriptomic and epigenomic profilings associated with schizophrenia might affect the early neurodevelopmental trajectory [39]. A recent report from the PsychENCODE Consortium revealed that schizophrenia risk genes were enriched in a co-expression module which was significantly associated with neurodevelopment, and expression pattern was significantly changed during the prenatal and postnatal neurodevelopment [39]. Therefore, we explored the dynamic pattern of mRNA expression and methylation conditions during the fetal neurodevelopment. We found that *BTN3A2* was upregulated during the fetal brain development (Fig. 2A), but the methylation levels of the above described schizophrenia-associated loci (Table 2) showed an opposite trend (except for cg10795676, Fig. 2B). These results further suggested that GWAS risk SNPs might affect the expression of *BTN3A2* through altering the methylation profiles.

### 3.6. BTN3A2 was differentially expressed between schizophrenia cases and controls

Post-mortem gene-expression studies of schizophrenia patients and controls suggested subtle abnormalities in multiple brain regions (including the prefrontal and temporal cortex, hippocampus, etc.) and several specific cell types [17,48,49]. Given the fact that integrative analyses identified *BTN3A2* as a risk gene for schizophrenia and the risk allele G of rs1979 was significantly associated with a higher mRNA expression of

536 Table 2

Integrative analysis (SMR) of schizophrenia GWAS (CLOZUK + PGC GWAS) and ROSMAP Hannon Jaffe brain meta meQTL reveals BTN3A2 as a schizophrenia risk gene.

Gene	Probe	Supporting SNP (cis or trans)	P <sub>GWAS</sub> <sup>a</sup>	$P_{eQTL}^{b}$	P <sub>SMR</sub>	P <sub>HEIDI</sub>	$P_{adj}$
BTN3A2 BTN3A2 BTN3A2 BTN3A2 BTN3A2	cg23465465 cg14345882 cg02045355 cg10795676	rs1977 (cis) rs9393714 (cis) rs3734536 (cis) rs34565965 (cis)	$\begin{array}{c} 6.63\times 10^{-27} \\ 6.52\times 10^{-27} \\ 9.20\times 10^{-15} \\ 2.21\times 10^{-13} \end{array}$	$\begin{array}{c} 1.76 \times 10^{-21} \\ < 1 \times 10^{-20} \\ < 1 \times 10^{-20} \\ 2.83 \times 10^{-10} \end{array}$	$\begin{array}{c} 1.06 \times 10^{-12} \\ 4.49 \times 10^{-25} \\ 3.03 \times 10^{-14} \\ 1.73 \times 10^{-6} \end{array}$	$\begin{array}{c} 1.16 \times 10^{-2} \\ 3.06 \times 10^{-5} \\ 1.04 \times 10^{-6} \\ 2.70 \times 10^{-6} \end{array}$	<0.05 <0.05 <0.05 NS

<sup>a</sup> *P*-values for eQTL SNPs from the CLOZUK + PGC GWAS data [7].

<sup>b</sup> *P*-values for eQTL SNPs from the ROSMAP Hannon Jaffe brain meta meQTL methylation data [33].

*BTN3A2*, we speculated that this gene might be up-regulated in schizophrenia cases. To test this hypothesis, we downloaded the gene expression data from CMC [17] which sequenced RNA from DLPFC of both schizophrenia patients (N = 258) and control subjects (N = 279) [17]. We found that *BTN3A2* was significantly up-regulated (*P*-value = .018) in schizophrenia cases compared with control subjects (Fig. 3A).

To further validate this result, we downloaded the gene expression data obtained from neurons differentiated from the human induced pluripotent stem cells [41]. This dataset contained microarray data of neurons that were differentiated from reprogrammed fibroblasts from schizophrenia patients and control individuals. The mRNA expression levels of *BTN3A2* in neurons were compared between cells derived from schizophrenia cases and control subjects, and we again observed significant increase of the mRNA levels of this gene in schizophrenia cases (Fig. 3B). Overall, these results were consistent with our findings of the integrative analysis, and suggested that a higher expression level of *BTN3A2* might increase the risk of schizophrenia.

#### 3.7. BTN3A2 specifically regulated excitatory synaptic transmission

Our analyses strongly suggested that *BTN3A2* plays essential roles in schizophrenia pathogenesis. We therefore further explored the function of *BTN3A2* to provide insights into its role in this illness. BTN3A2 is an immunoglobulin (Ig) superfamily receptor protein of the butyrophilin and butyrophilin-like families, and is known to play critical immunomodulatory roles [45]. Moreover, Ig superfamily adhesion molecules are among the most abundant trans-synaptic signaling molecules and

exert diverse functions during synaptic physiological and pathological processes [50–52]. To study the function of BTN3A2 during synaptic transmission, we used cultured rat hippocampal slice as the study system and expressed BTN3A2 exogenously in CA1 neurons through biolistic transfection. Then dual whole-cell recordings were applied to measure the evoked excitatory or inhibitory postsynaptic responses. We found that overexpression of BTN3A2 decreased both AMPAR and NMDAR-mediated excitatory synaptic transmission compared with the control neurons (Fig. 4A and B), while no effect was seen on the ratio of AMPA and NMDA EPSCs (Fig. 4C). In addition, the paired-pulse ratio (Fig. 4D), a parameter for presynaptic release probability, was increased in BTN3A2-overexpressed neurons. These results suggested that BTN3A2 regulated excitatory synaptic transmission through decreasing the presynaptic glutamate release. In contrast, BTN3A2 had no effect on GABA receptor-mediated inhibitory postsynaptic transmission (Fig. 4E). All these results suggested that BTN3A2 was specifically involved in excitatory synaptic maintenance and its dysregulated expression might lead to malfunction of excitatory/inhibitory balance.

#### 3.8. BTN3A2 interacted with pre-synaptic neurexins

Neurexins (Nrxns) are a family of presynaptic transmembrane proteins that function as the organizer of synapses [53]. They play important roles in presynaptic release regulation and have been implicated in schizophrenia by human genetic studies [54–57]. Therefore, we hypothesized that BTN3A2 might interact with presynaptic neurexins and thereby affect the excitatory synaptic maintenance and



Fig. 2. BTN3A2 expression in human brain is highest post-natally. (A) BTN3A2 expression pattern in human brain according to PsychENCODE [39], Human Brain Transcriptome [38], BrianCloud [26] and Brainspan [76] datasets. (B) Level of methylation in four sites identified by meQTL integrative analysis based on the data reported by Jaffe et al. [36].



Fig. 3. *BTN3A2* is differentially expressed between schizophrenia cases and controls. The mRNA level of *BTN3A2* is significantly up-regulated in (A) post-mortem dorsolateral prefrontal cortex tissues and (B) iPSC neurons from schizophrenia patients compared to that of healthy controls. Data were taken from CMC [17] and GSE25673 [41], respectively. The *P*-values were calculated by using unpaired *Student's t*-test.

excitatory/inhibitory balance [53]. To test this hypothesis, we used cell surface binding assay to test whether BTN3A2 could interact with neurexins. The Flag-tagged and truncated BTN3A2 (BTN3A2-S-Flag), which lacked the transmembrane domain (Fig. 5A), was successfully expressed in HEK293T cells and its secretion to extracellular supernatant was confirmed by Western blot (Fig. 5B). We then overexpressed DsRed2-tagged Nrxn1 $\beta$  or Nrxn3 $\beta$  in HEK293T cells for 48 h, and replaced the extracellular medium with culture supernatant containing secreted BTN3A2-S-Flag for another 4 h. Immunostaining was performed to examine the surface expression of Nrxns and BTN3A2. We found that BTN3A2 was co-localized with either Nrxn1 $\beta$  or Nrxn3 $\beta$  at the cell surface (Fig. 5C), suggesting that they were directly interacted through the Nrxns extracellular domain. However, the alternative

splicing at the 4th canonical sites of Nrxn1 $\beta$  (Nrxn1 $\beta$ A4 (-)) that is critical for its binding to its postsynaptic ligand, was not required for this interaction as Nrxn1 $\beta$ A4 (-) still significantly bound with BTN3A2 (Fig. 5C). The same strategy was used to determine the interaction between BTN3A2 and postsynaptic neuroligins (Nlgns). We found that neither Nlgn1 nor Nlgn3 interacted with BTN3A2 (Fig. 5C). These results suggest that BTN3A2 specifically interacts with presynaptic Nrxns. Based on these observations, we propose a model that BTN3A2 functions as a postsynaptic ligand for neurexins to regulate presynaptic glutamate release, thereby maintaining the excitatory/inhibitory balance of our brain. Abnormality of *BTN3A2* expression might cause an imbalance of synaptic transmission and lead to schizophrenia (Fig. 5D).



**Fig. 4.** BTN3A2-mediated regulation is specific to excitatory synapses. Overexpression of BTN3A2 decreases both AMPAR (A) and NMDAR (B)-mediated synaptic transmissions compared with the control neurons. Overexpression of BTN3A2 increases the paired-pulse ratio, a parameter for presynaptic release probability (C), but does not change the ratio of AMPAR and NMADR-mediated EPSCs related to neighboring wild-type neurons (D), and has no effect on GABA receptor-mediated inhibitory postsynaptic transmission (E). Differences between neurons overexpressing *BTN3A2* and respective control neurons are compared using two-tailed Wilcoxon signed-rank sum test. \**P*-value <.05, \*\**P*-value <.01, \*\*\**P*-value <.001. Bars represent mean ± SEM.



D



**Fig. 5.** Interaction of BTN3A2 with presynaptic neurexins and a working model of the role of BTN3A2 in synapses. (A) Schematic profile of the two isoforms of BTN3A2. The short isoform BTN3A2-S has a deletion of the transmembrane domain of the full length BTN3A2 (BTN3A2-L). (B) The overexpressed BTN3A2-S in HEK293T cells can be secreted into culture supernatant. The Western blot was detected by using the Flag antibody. (C) Cell surface binding assay showing potential interaction of BTN3A2-S with neurexins. HEK293T cells were transfected with overexpression vector pCMV-NIgn1-DsRed2, pCMV-NIgn3-DsRed2, pCMV-Nrxn1β-DsRed2, pCMV-Nrxn1β-DsRed2, pCMV-Nrxn1β-DsRed2, pCMV-Nrxn3β-DsRed2 or pCMV-DsRed2 (empty vector) alone for 48 h, then were incubated with the culture supernatant of HEK293T overexpression pCMV-3Tag-8 (green) for 4 h before fixing for visualization. Scale bar, 5 µm. (D) A working model for the unbalanced expression of *BTN3A2* in schizophrenia. I: inhibitory synaptic transmission; E: excitatory synaptic transmission.

#### 4. Discussion

Hitherto, GWAS studies have identified hundreds of loci associated with schizophrenia [7,10]. We now know that the genetic basis and heritability of schizophrenia were largely facilitated by common variations [7,10]. However, as many of these risk variants were located in intergenic regions or other non-coding regions, identifying 'targetable' genes in sizable studies has been proven to be difficult, with a few exceptions [40,47,58,59]. Moreover, due to the complexity of schizophrenia and different genetic backgrounds among various ethnic groups, it is even harder to find out the causal genes conferring the risk of schizophrenia. Considering the fact that most of the identified genetic risk loci are located in non-coding regions, these factors may confer the risk of schizophrenia by affecting the expression of nearby gene (s) and/or changing the chromatin state [13,14,60]. Integrative studies with brain gene expression data or other multiple-omics data like DNA methylation, proteins or metabolites can help us parse the result of GWAS studies, as we and others had done recently [12-16]. However, the result of integrative analysis only provided a list of prioritized genes [12–16]. Further follow-up functional experiments were essential to verify the results [13], which constitutes the main work in the post-GWAS era.

In this study, we used a large-scale genotype and expression data from 268 human DLPFC samples from the BrainCloud [26] to conduct eQTL analyses. Then we integrated these eQTL data [26] with CLOZUK + PGC GWAS data [7] results to identify risk genes whose expression changes might confer risk of schizophrenia. We found that BTN3A2 was a potential schizophrenia risk gene, and increased expression of BTN3A2 may contribute to schizophrenia pathogenesis. Further validation using GTEx brain eQTL datasets [32], PsychENCODE brain PFC eQTL datasets [31] and a different integrative analysis method [37] confirmed our findings. Moreover, integrating GWAS data [7] with meQTL data [33] also identified the same gene. These integrative studies from gene expression and methylation suggested that BTN3A2 might affect the risk of schizophrenia, and altered methylation statuses of its nearby sites might play a role. While integrating GWAS and eQTL data is widely accepted as a promising strategy to reveal mechanisms underlying the genetic risk of schizophrenia, few of these studies conducted in recent years have successfully explained how the GWAS risk SNPs affected the expression of nearby genes [13,14,16]. In this study, we combined many newly published datasets like CLOZUK + PGC GWAS data [7] and PsychENCODE brain PFC eQTL data [31] that had not been used in previous schizophrenia integrative studies [13,16,24,61]. Reports from Ma et al. [24], Luo et al. [16] and Yang et al. [13] used the eOTL dataset generated by Myers et al. [62], which was published many years ago. However, in this study, we combined the newly published CLOZUK + PGC GWAS data [7] with RNA-seq generated eQTL dataset [32] using newly developed SMR method [28] and deposited all the integrative analysis results in SZDB (www.szdb.org/integrative1.php) [12] for data sharing. Moreover, we found that the expression and methylation levels of BTN3A2 were significantly changed during the prenatal and postnatal neurodevelopment, with an absolute opposite trend. We further showed that the mRNA levels of BTN3A2 were significantly upregulated in brain tissues of schizophrenia cases compared with controls. Consistent with the finding in brains, we found that the BTN3A2 gene was also significantly up-regulated in induced pluripotent stem cells derived neurons from schizophrenia patients [41]. Furthermore, overexpression of BTN3A2 in rat hippocampal CA1 pyramidal neurons could decrease both AMPAR and NMDAR-mediated synaptic transmissions, but had no effect on GABA receptor-mediated inhibitory postsynaptic transmission. Finally, cell surface binding assays indicated that this effect might be mediated by the binding of BTN3A2 to neurexins. This study provided a good example of identifying risk gene for schizophrenia via integrative analyses and functional characterization.

Schizophrenia has long been considered as a neurodevelopmental disorder resulted from aberrant developmental processes in fetal, childhood, or adolescent stage [63]. The central nervous system, especially the brain, is vulnerable to the impact of genetic and environmental risk factors for schizophrenia in the prenatal and postnatal developmental periods [63]. Indeed, previous studies have found that many genes that play key roles in neurodevelopment, including brain circuits and synaptic organization, are potential schizophrenia risk genes. For example, ZNF804A is the first putative schizophrenia risk gene identified by GWAS [3,64]. The risk allele of schizophrenia risk SNP rs1344706 was found to decrease ZNF804A cis-expression in the fetal brain during the second trimester, but had no regulatory effects in the first trimester or the adult brain [65]. In utero mouse overexpression experiment verified that ZNF804A was required for normal progenitor proliferation and neuronal migration [66]. In addition, NRGN, another important synapse organizer, could ameliorate ZFP804Amediated migration defect [66]. In this study, we found that the mRNA expression of BTN3A2 was significantly changed during the prenatal and postnatal neurodevelopment, which might be disturbed by different alleles of GWAS risk SNPs. Importantly, BTN3A2 could interact with presynaptic neurexins and this interaction might lead to a decrease of the excitatory presynaptic release. These lines of evidence further suggested that BTN3A2 might be a schizophrenia susceptibility gene by altering excitatory synaptic function. It should be mentioned that we recently identified independent major depressive disorder (MDD) risk SNPs in the MHC region, which would also tag BTN3A2 as a MDD risk gene [67].

Based on the data from Gentree database (http://gentree.ioz.ac.cn/) [44] and evolutionary analysis, we found that *BTN3A2* is specific to some Old World monkeys and human species [45]. Owing to the complexity of cerebral cortex, the human brain developed advanced cognitive, emotional and social capacity. However, the more complex the cerebral cortex is, the more likely that aberrant neural connectivity would appear and result in mental disorders [68,69]. A good case is human-specific *NOTCH2NL* gene, which contributes to the rapid evolution of a larger human neocortex with the trade-off effect of loss of genomic stability at the 1q21.1 locus, finally confers recurrent neurodevelopmental disorders [70]. Whether this newly evolved *BTN3A2* in primates is associated with increased the complexity of the brain, with the cost of higher schizophrenia risk, remains to be answered.

The current study has some limitations. First, the  $P_{\text{HEIDI}}$  values of SMR [28] integrating CLOZUK + PGC GWAS data [7] with BrainCloud eQTL [26] or CLOZUK + PGC GWAS data [7] with PsychENCODE eQTL data [31] were < 0.01 (Table 1 and Table S3), which suggested that the SMR association was most likely driven by linkage [28,71]. However, the region containing *BTN3A2* has a very complex LD pattern (Fig. S4), and it would be hard to identify the causal SNP(s). Second, the risk gene list distilled by integrative analyses was dependent on gene expression profiles of eQTL datasets. Different eQTL datasets might offer different results. For instance, a recent integrative analysis using SMR [28] to integrate CLOZUK + PGC GWAS data [7] with fetal brain eQTL [72] did not identify the association of BTN3A2 with schizophrenia. Evidently, refined expression data would promote the authenticity of the identified risk gene. A recent analysis of brain single cells by single cell RNA-sequencing highlighted a key role of four types of cells (hippocampal CA1 pyramidal cells, striatal medium spiny neurons, neocortical somatosensory pyramidal cells and cortical interneurons) in schizophrenia [49]. A comprehensive integration of expression data at the single-cell level would cast new insights into the etiology of schizophrenia. Third, the electrophysiological analyses of rat brain with overexpression of BTN3A2 might not be a perfect system to show the neurobiological function of this gene as this gene was primate-specific [45]. Tests using non-human primates (NHPs) lacking BTN3A2 and/or overexpression of this gene should be a promising way to approach the final answer, given the recent advancements in genetic modification of NHPs [73-75].

In summary, our integrative analyses and functional assays showed that *BTN3A2* is a schizophrenia susceptibility gene. Altered expression of *BTN3A2* could affect excitatory and inhibitory synaptic balance through regulating excitatory presynaptic release. Further studies using the schizophrenia related single cell RNA-seq eQTL data and NHP models may help with elucidating the etiology of schizophrenia.

#### Acknowledgements

We thank Futao Zhang, the author of SMR, for the help of data analyzing. The BrainCloud data was downloaded from http://braincloud. jhmi.edu/ and dbGap. We thank the families who donated tissue to make this study possible. The gene expression data of DLPFC was taken from CommonMind Consortium. These data 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, R01-MH-075916, P50M096891, P50MH084053S1, and R37MH057881S1, HHSN271201300031C, R37MH057881 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.

#### **Funding sources**

This study was supported by the Strategic Priority Research Program (B) of the Chinese Academy of Sciences (XDB02020003 to Y-GY), the National Natural Science Foundation of China (31730037 to Y-GY), and the Bureau of Frontier Sciences and Education, Chinese Academy of Sciences (QYZDJ-SSW-SMC005 to Y-GY). The funders had no role in study design, data collection, data analysis, interpretation, writing of the report.

#### **Declaration of interests**

The authors declare no competing interests.

#### Author contributions

Y-G. Yao, X-J. Luo, N. Sheng, and Y. Wu designed research; Y. Wu, N. Sheng, and C. Zeng performed research; R. Bi, C. Ma, C. Sun, J. Li, and P. Zheng, contributed new reagents/analytic tools; Y. Wu, D-F. Zhang, M. Li, X. Xiao, and Y-G. Yao analyzed data; Y. Wu and Y-G. Yao wrote the paper. All authors approved the submission.

#### Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi. org/10.1016/j.ebiom.2019.05.006.

#### References

- [1] Owen MJ, Sawa A, Mortensen PB. Schizophrenia. Lancet 2016;388:86-97.
- [2] Birnbaum R, Weinberger DR. Genetic insights into the neurodevelopmental origins of schizophrenia. Nat Rev Neurosci 2017;18:727–40.

- [3] O'Donovan MC, Craddock N, Norton N, et al. Identification of loci associated with schizophrenia by genome-wide association and follow-up. Nat Genet 2008;40: 1053–5.
- [4] International Schizophrenia Consortium, Purcell SM, Wray NR, et al. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. Nature 2009;460:748–52.
- [5] Shi Y, Li Z, Xu Q, et al. Common variants on 8p12 and 1q24.2 confer risk of schizophrenia. Nat Genet 2011;43:1224–7.
- [6] 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: 1228–31.
- [7] Pardiñas AF, Holmans P, Pocklington AJ, et al. Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. Nat Genet 2018;50:381–9.
- [8] Li Z, Chen J, Yu H, et al. Genome-wide association analysis identifies 30 new susceptibility loci for schizophrenia. Nat Genet 2017;49:1576–83.
- [9] Yu H, Yan H, Wang L, et al. Five novel loci associated with antipsychotic treatment response in patients with schizophrenia: a genome-wide association study. Lancet Psychiatry 2018;5:327–38.
- [10] Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. Nature 2014;511:421–7.
- [11] Ripke S, O'Dushlaine C, Chambert K, et al. Genome-wide association analysis identifies 13 new risk loci for schizophrenia. Nat Genet 2013;45:1150–9.
- [12] Wu Y, Yao YG, Luo XJ. SZDB: a database for schizophrenia genetic research. Schizophr Bull 2017;43:459–71.
- [13] Yang CP, Li X, Wu Y, et al. Comprehensive integrative analyses identify GLT8D1 and CSNK2B as schizophrenia risk genes. Nat Commun 2018;9:838.
   [14] Dobbyn A, Huckins LM, Boocock L et al. Landscape of conditional eOTL in dorsolat-
- [14] Dobbyn A, Huckins LM, Boocock J, et al. Landscape of conditional eQTL in dorsolateral prefrontal cortex and co-localization with schizophrenia GWAS. Am J Hum Genet 2018;102:1169–84.
- [15] Hauberg ME, Zhang W, Giambartolomei C, et al. Large-scale identification of common trait and disease variants affecting gene expression. Am J Hum Genet 2017; 100:885–94.
- [16] Luo XJ, Mattheisen M, Li M, et al. Systematic integration of brain eQTL and GWAS identifies ZNF323 as a novel schizophrenia risk gene and suggests recent positive selection based on compensatory advantage on pulmonary function. Schizophr Bull 2015;41:1294–308.
- [17] Fromer M, Roussos P, Sieberts SK, et al. Gene expression elucidates functional impact of polygenic risk for schizophrenia. Nat Neurosci 2016;19:1442–53.
- [18] Purcell SM, Moran JL, Fromer M, et al. A polygenic burden of rare disruptive mutations in schizophrenia. Nature 2014;506:185–90.
- [19] Tang J, Fan Y, Li H, et al. Whole-genome sequencing of monozygotic twins discordant for schizophrenia indicates multiple genetic risk factors for schizophrenia. J Genet Genomics 2017;44:295–306.
- [20] Nicolae DL, Gamazon E, Zhang W, et al. Trait-associated SNPs are more likely to be eQTLs: annotation to enhance discovery from GWAS. PLoS Genet 2010;6:e1000888.
- [21] Guo X, Lin W, Bao J, et al. A comprehensive cis-eQTL analysis revealed target genes in breast cancer susceptibility loci identified in genome-wide association studies. Am J Hum Genet 2018;102:890–903.
- [22] Ward LD, Kellis M. Interpreting noncoding genetic variation in complex traits and human disease. Nat Biotechnol 2012;30:1095–106.
- [23] Maurano MT, Humbert R, Rynes E, et al. Systematic localization of common diseaseassociated variation in regulatory DNA. Science 2012;337:1190–5.
- [24] Ma C, Gu C, Huo Y, Li X, Luo XJ. The integrated landscape of causal genes and pathways in schizophrenia. Transl Psychiatry 2018;8:67.
- [25] Schizophrenia Psychiatric Genome-Wide Association Study (GWAS) Consortium. Genome-wide association study identifies five new schizophrenia loci. Nat Genet 2011;43:969–76.
- [26] Colantuoni C, Lipska BK, Ye T, et al. Temporal dynamics and genetic control of transcription in the human prefrontal cortex. Nature 2011;478:519–23.
- [27] 1000 Genomes Project Consortium, Abecasis GR, Altshuler D, et al. A map of human genome variation from population-scale sequencing. Nature 2010;467:1061–73.
- [28] Zhu Z, Zhang F, Hu H, et al. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. Nat Genet 2016;48:481–7.
- [29] Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007;81:559–75.
- [30] PsychEncode Consortium, Akbarian S, Liu C, et al. The PsychENCODE project. Nat Neurosci 2015;18:1707–12.
- [31] Wang D, Liu S, Warrell J, et al. Comprehensive functional genomic resource and integrative model for the human brain. Science 2018;362:eaat8464.
- [32] GTEx Consortium. The genotype-tissue expression (GTEx) project. Nat Genet 2013; 45:580–5.
- [33] Qi T, Wu Y, Zeng J, et al. Identifying gene targets for brain-related traits using transcriptomic and methylomic data from blood. Nat Commun 2018;9:2282.
- [34] Ng B, White CC, Klein HU, et al. An xQTL map integrates the genetic architecture of the human brain's transcriptome and epigenome. Nat Neurosci 2017;20:1418–26.
- [35] Hannon E, Spiers H, Viana J, et al. Methylation QTLs in the developing brain and their enrichment in schizophrenia risk loci. Nat Neurosci 2016;19:48–54.
- [36] Jaffe AE, Gao Y, Deep-Soboslay A, et al. Mapping DNA methylation across development, genotype and schizophrenia in the human frontal cortex. Nat Neurosci 2016;19:40–7.
- [37] He X, Fuller CK, Song Y, et al. Sherlock: detecting gene-disease associations by matching patterns of expression QTL and GWAS. Am J Hum Genet 2013;92:667–80.
- [38] Kang HJ, Kawasawa YI, Cheng F, et al. Spatio-temporal transcriptome of the human brain. Nature 2011;478:483–9.

- [39] Li M, Santpere G, Imamura Kawasawa Y, et al. Integrative functional genomic analysis of human brain development and neuropsychiatric risks. Science 2018;362 eaat7615.
- [40] Fromer M, Pocklington AJ, Kavanagh DH, et al. De novo mutations in schizophrenia implicate synaptic networks. Nature 2014;506:179–84.
- [41] Brennand KJ, Simone A, Jou J, et al. Modelling schizophrenia using human induced pluripotent stem cells. Nature 2011;473:221–5.
- [42] Sheng N, Shi YS, Nicoll RA. Amino-terminal domains of kainate receptors determine the differential dependence on Neto auxiliary subunits for trafficking. Proc Natl Acad Sci U S A 2017;114:1159–64.
- [43] Zhang D-F, Fan Y, Xu M, et al. Complement C7 is a novel risk gene for Alzheimer's disease in Han Chinese. Natl Sci Rev 2019;6:257–74.
- [44] Zhang YE, Vibranovski MD, Landback P, Marais GA, Long M. Chromosomal redistribution of male-biased genes in mammalian evolution with two bursts of gene gain on the X chromosome, PLoS Biol 2010;8:e1000494.
- [45] Afrache H, Pontarotti P, Abi-Rached L, Olive D. Evolutionary and polymorphism analyses reveal the central role of BTN3A2 in the concerted evolution of the BTN3 gene family. Immunogenetics 2017;69:379–90.
- [46] Ramasamy A, Trabzuni D, Guelfi S, et al. Genetic variability in the regulation of gene expression in ten regions of the human brain. Nat Neurosci 2014;17:1418–28.
- [47] Sekar A, Bialas AR, de Rivera H, et al. Schizophrenia risk from complex variation of complement component 4. Nature 2016;530:177–83.
- [48] Horváth S, Janka Z, Mirnics K. Analyzing schizophrenia by DNA microarrays. Biol Psychiatry 2011;69:157–62.
- [49] Skene NG, Bryois J, Bakken TE, et al. Genetic identification of brain cell types underlying schizophrenia. Nat Genet 2018;50:825–33.
- [50] Liu H, Focia PJ, He X. Homophilic adhesion mechanism of neurofascin, a member of the L1 family of neural cell adhesion molecules. J Biol Chem 2011;286:797–805.
- [51] Fogel AI, Stagi M, Perez de Arce K, Biederer T. Lateral assembly of the immunoglobulin protein SynCAM 1 controls its adhesive function and instructs synapse formation. EMBO | 2011;30:4728–38.
- [52] Gangwar SP, Zhong X, Seshadrinathan S, et al. Molecular mechanism of MDGA1: regulation of neuroligin 2:neurexin trans-synaptic bridges. Neuron 2017;94:1132–41 e4.
- [53] Südhof TC. Synaptic neurexin complexes: a molecular code for the logic of neural circuits. Cell 2017;171:745–69.
- [54] Kirov G, Gumus D, Chen W, et al. Comparative genome hybridization suggests a role for NRXN1 and APBA2 in schizophrenia. Hum Mol Genet 2008;17:458–65.
- [55] Walsh T, McClellan JM, McCarthy SE, et al. Rare structural variants disrupt multiple genes in neurodevelopmental pathways in schizophrenia. Science 2008;320: 539–43.
- [56] Vrijenhoek T, Buizer-Voskamp JE, van der Stelt I, et al. Recurrent CNVs disrupt three candidate genes in schizophrenia patients. Am J Hum Genet 2008;83:504–10.
- [57] Kasem E, Kurihara T, Tabuchi K. Neurexins and neuropsychiatric disorders. Neurosci Res 2018;127:53–60.

- [58] Genovese G, Fromer M, Stahl EA, et al. Increased burden of ultra-rare proteinaltering variants among 4,877 individuals with schizophrenia. Nat Neurosci 2016; 19:1433–41.
- [59] Singh T, Kurki MI, Curtis D, et al. Rare loss-of-function variants in SETD1A are associated with schizophrenia and developmental disorders. Nat Neurosci 2016;19: 571–7.
- [60] González-Peñas J, Amigo J, Santomé L, et al. Targeted resequencing of regulatory regions at schizophrenia risk loci: role of rare functional variants at chromatin repressive states. Schizophr Res 2016;174:10–6.
- [61] Zhao Y, He A, Zhu F, et al. Integrating genome-wide association study and expression quantitative trait locus study identifies multiple genes and gene sets associated with schizophrenia. Prog Neuropsychopharmacol Biol Psychiatry 2018;81:50–4.
- [62] Myers AJ, Gibbs JR, Webster JA, et al. A survey of genetic human cortical gene expression. Nat Genet 2007;39:1494–9.
- [63] Catts VS, Fung SJ, Long LE, et al. Rethinking schizophrenia in the context of normal neurodevelopment. Front Cell Neurosci 2013;7:60.
- [64] Chang H, Xiao X, Li M. The schizophrenia risk gene ZNF804A: clinical associations, biological mechanisms and neuronal functions. Mol Psychiatry 2017;22:944–53.
- [65] Hill MJ, Bray NJ. Evidence that schizophrenia risk variation in the ZNF804A gene exerts its effects during fetal brain development. Am J Psychiatry 2012;169:1301–8.
- [66] Zhou Y, Dong F, Lanz TA, et al. Interactome analysis reveals ZNF804A, a schizophrenia risk gene, as a novel component of protein translational machinery critical for embryonic neurodevelopment. Mol Psychiatry 2018;23:952–62.
- [67] Li H, Chang H, Song X, et al. Integrative analyses of major histocompatibility complex loci in the genome-wide association studies of major depressive disorder. Neuropsychopharmacology 2019. https://doi.org/10.1038/s41386-019-0346-3.
- [68] Takahashi T. Complexity of spontaneous brain activity in mental disorders. Prog Neuropsychopharmacol Biol Psychiatry 2013;45:258–66.
- [69] Fernández A, Gómez C, Hornero R, López-Ibor JJ. Complexity and schizophrenia. Prog Neuropsychopharmacol Biol Psychiatry 2013;45:267–76.
- [70] Fiddes IT, Lodewijk GA, Mooring M, et al. Human-specific NOTCH2NL genes affect NOTCH signaling and cortical neurogenesis. Cell 2018;173:1356–69.
- [71] Wu Y, Zeng J, Zhang F, et al. Integrative analysis of omics summary data reveals putative mechanisms underlying complex traits. Nat Commun 2018;9:918.
- [72] O'Brien HE, Hannon E, Hill MJ, et al. Expression quantitative trait loci in the developing human brain and their enrichment in neuropsychiatric disorders. Genome Biol 2018;19:194.
- [73] Chen Y, Niu Y, Ji W. Genome editing in nonhuman primates: approach to generating human disease models. J Intern Med 2016;280:246–51.
- [74] Luo X, Li M, Su B. Application of the genome editing tool CRISPR/Cas9 in non-human primates. Zool Res 2016;37:214–9.
- [75] Liu Z, Cai Y, Liao Z, et al. Cloning of a gene-edited macaque monkey by somatic cell nuclear transfer. Natl Sci Rev 2019;6:101–8.
- [76] Brianspan. BrainSpan: Atlas of the developing human brain. Internet http://www. brainspan.org; 2013, Accessed date: 28 April 2015.

### **Supplemental Information**



**Fig. S1.** The eQTL effect of rs1979 on *BTN3A2* mRNA expression in BrainCloud DLPFC data [1]. None of the BrainCloud samples had genotype TT. Statistical differences were calculated by the *Student's t*-test. Bars represent mean  $\pm$  SD.



**Fig. S2.** Relative mRNA expression levels of *BTN3A2* in 14 brain tissues affected by the eQTL of rs1979. The original data were taken from the GTEx data [2].



#### Affymetrix ID t2899333 stratified by rs1979 (chr6:26377591)

**Fig. S3.** Relative mRNA expression levels of *BTN3A2* in 10 brain tissues affected by the eQTL of rs1979. The original data were taken from the Braineac [3]. MEDU: the inferior olivary nucleus, SNIG: substantianigra, WHMT: intralobular white matter, THAL: thalamus, HIPP: hippocampus, PUTM: putamen, CRBL: cerebellar cortex, FCTX: frontal cortex, TCTX: temporal cortex, OCTX: occipital cortex.



Fig. S4. BTN3A2 supporting SNPs and C4A supporting SNPs were distributed in independent blocks (Data from the 1000 Genomes dataset [4]).

Table S1. Primers used for constructing the plasmids

Vector	Restriction enzyme	Primer pair <sup>a</sup>
pCMV-BTN3A2-L-Flag	HindIII & XhoI	5'-CCCAAGCTTATGAAAATGGCAAGTTCCCTGGC-3' / 5'-CCGCTCGAGGGCTGACTTATTGGTATCGGACG-3'
pCMV-BTN3A2-S-Flag	HindIII & XhoI	5'-CCCAAGCTTATGAAAATGGCAAGTTCCCTGGC-3' / 5'-CCGCTCGAGGGCTGACTTATTGGTATCGGACG-3'
pCMV-Nlgn1-DsRed2	SacI & HindIII	5'-GCGAGCTCATGGCACTTCCCAGATGCATG-3' / 5'-CCCAAGCTTTACCCTGGTTGTTGAATGTG-3'
pCMV-Nlgn3-DsRed2	SacIF & HindIII	5'-CGAGCTCATGTGGCTGCAGCCCTCGCTG-3' / 5'-CCCAAGCTTTACACGGGTAGTGGAGTGTG-3'
pCMV-Nrxn1β-DsRed2	BamHI & XhoI	5'-CGGGATCCATGTACCAGAGGATGCTTCG-3' / 5'-CCGCTCGAGGACGTAATACTCCTTATCC-3'
pCMV-Nrxn1βA4(-)-DsRed2	BamHI & XhoI	5'-CGGGATCCATGTACCAGAGGATGCTTCG-3' / 5'-CCGCTCGAGGACGTAATACTCCTTATCC-3'
pCMV-Nrxn3β-DsRed2	BamHI & XhoI	5'-CGGGATCCATGCACCTGAGAATCCACCC-3' / 5'-CCGCTCGAGCACATAATACTCCTTGTCC-3'

<sup>a</sup> The region marked in red refers to restriction enzyme digestion site in each primer.

Gene	LBF	<i>P</i> -value <sup>b</sup>	Supporting SNP <sup>c</sup> ( <i>trans</i> or <i>cis</i> )	PeQTL <sup>d</sup>	<b>P</b> GWAS <sup>e</sup>	<b>P</b> adj <sup>f</sup>
HOGA1	12.60	1.68×10 <sup>-7</sup>	rs7794036 (trans)	4.54×10 <sup>-7</sup>	1.22×10 <sup>-7</sup>	< 0.05
ZKSCAN4	10.15	1.68×10 <sup>-7</sup>	rs916403 (cis)	6.76×10 <sup>-5</sup>	3.97×10 <sup>-15</sup>	< 0.05
ZNF322	9.75	1.68×10 <sup>-7</sup>	rs9379895 (trans)	4.08×10 <sup>-4</sup>	3.98×10 <sup>-22</sup>	< 0.05
BTN3A2	9.38	1.68×10 <sup>-7</sup>	rs9379858 (cis)	1.99×10 <sup>-15</sup>	3.25×10 <sup>-25</sup>	< 0.05
USH1C	8.54	1.68×10 <sup>-7</sup>	rs4712981 (trans)	2.00×10 <sup>-7</sup>	5.60×10 <sup>-17</sup>	< 0.05
SCGB2B2	7.09	1.68×10 <sup>-7</sup>	rs1004046 (cis)	1.12×10 <sup>-4</sup>	1.82×10 <sup>-8</sup>	< 0.05
LOC100131289	6.97	1.68×10 <sup>-7</sup>	rs6904596 (cis)	8.63×10 <sup>-8</sup>	2.74×10 <sup>-39</sup>	< 0.05
ASB3	6.32	1.68×10 <sup>-7</sup>	rs2357896 (cis)	6.66×10 <sup>-8</sup>	3.42×10 <sup>-6</sup>	< 0.05
NRIP2	5.84	1.68×10 <sup>-7</sup>	rs10491964 (cis)	9.07×10 <sup>-5</sup>	1.42×10 <sup>-8</sup>	< 0.05
SFXN2	5.72	1.68×10 <sup>-7</sup>	rs11191425 (trans)	3.16×10 <sup>-5</sup>	6.06×10 <sup>-14</sup>	< 0.05
ELAC2	5.47	1.68×10 <sup>-7</sup>	rs2523 (trans)	4.75×10 <sup>-13</sup>	2.19×10 <sup>-6</sup>	< 0.05
FAM114A2	5.17	1.68×10 <sup>-7</sup>	rs566604 (trans)	2.79×10 <sup>-7</sup>	2.35×10 <sup>-5</sup>	< 0.05
VWA5A	5.15	1.68×10 <sup>-7</sup>	rs3759929 (trans)	4.77×10 <sup>-7</sup>	1.68×10 <sup>-7</sup>	< 0.05
OYUD7A	5.00	1.68×10 <sup>-7</sup>	rs11975170 (trans)	1.79×10 <sup>-6</sup>	2.17×10 <sup>-8</sup>	< 0.05

Table S2. Integrative analysis (Sherlock) of schizophrenia GWAS (CLOZUK + PGC GWAS) and BrainCloud eQTL revealed that *BTN3A2* was associated with schizophrenia

<sup>a</sup>LBF (Log Bayes Factors for each gene) evaluates whether the combined evidence from GWAS and expression studies support a gene being associated with schizophrenia. For example, a LBF of 4.6 means that the alternative hypothesis (associated with Schizophrenia) is 100 times (exp[4.6] = 100) more likely than the null hypothesis (no association). High Bayes factors tend to correlate with low *P*-values. For each of the eSNP, if it is also associated with schizophrenia in GWAS, a positive score would be given; otherwise, a negative score. Association only in GWAS (ie, non-eSNPs) does not alter the score. The total score of a gene increases with the number of SNPs with combined evidence (SNPs that are associated with schizophrenia and expression). For each gene, Sherlock computes individual LBFs for each SNP pair in the alignment, and the sum of these constitutes the final LBF score for the gene. <sup>b</sup>*P*-values from Sherlock integrative analysis.

<sup>c</sup>SNP with the total LBF score.

<sup>d</sup> *P*-values for eQTL SNPs from the GTEx data [2].

<sup>e</sup>*P*-values for eQTL SNPs from the CLOZUK + PGC GWAS data [5].

<sup>f</sup>The analysis was performed at the Sherlock web server (<u>http://sherlock.ucsf.edu</u>) [6]. The

GTEx brain tissues gene expression data (release v.7) was selected for eQTL data [2].

Gene	Supporting SNP (trans or cis)	PGWAS	Peqtl	P <sub>SMR</sub>	PHEIDI	Padj
BTN3A2	rs9393704 (cis)	$3.70 \times 10^{-26}$	$3.14\times10^{\text{-}78}$	$3.28\times10^{\text{-}20}$	$5.51  imes 10^{-11}$	< 0.05
ZSCAN23	rs13203055 (cis)	$4.08\times10^{\text{-}20}$	$4.75 \times 10^{-93}$	$5.17  imes 10^{-17}$	$5.86 imes10^{-8}$	< 0.05
ZNF192P1	rs4711164 (cis)	$7.09  imes 10^{-14}$	$2.47  imes 10^{-83}$	$2.96 \times 10^{-12}$	$6.95 \times 10^{-12}$	< 0.05
ZKSCAN3	rs1361387 (cis)	$3.49\times10^{15}$	$2.37  imes 10^{-50}$	$3.23  imes 10^{-12}$	$1.38  imes 10^{-10}$	< 0.05
ZNF192P2	rs2859368 (cis)	$4.79\times10^{\text{-}15}$	$6.58 imes10^{-51}$	$3.89  imes 10^{-12}$	$3.80  imes 10^{-6}$	< 0.05
ZNF602P	rs9368552 (cis)	$1.56  imes 10^{-13}$	$8.69 imes10^{-49}$	$4.27  imes 10^{-11}$	$1.46  imes 10^{-10}$	< 0.05
ZNF603P	rs203892 (cis)	$5.99  imes 10^{-12}$	$2.50  imes 10^{-71}$	$1.38  imes 10^{-10}$	$6.08  imes 10^{-9}$	< 0.05
ZSCAN31	rs203892 (cis)	$5.99\times10^{\text{-}12}$	$2.08  imes 10^{-58}$	$2.52  imes 10^{-10}$	$5.81  imes 10^{-12}$	< 0.05
OR2B8P	rs6934769 (cis)	$7.94 imes10^{-14}$	$1.59  imes 10^{-31}$	$3.13  imes 10^{-10}$	$5.83  imes 10^{-7}$	< 0.05
FTCDNL1	rs2949006 (cis)	$3.69  imes 10^{-17}$	$1.11  imes 10^{-20}$	$4.13 \times 10^{-10}$	$3.67 \times 10^{-3}$	< 0.05
ZKSCAN8	rs4711164 (cis)	$7.09  imes 10^{-14}$	$1.91  imes 10^{-20}$	$5.80  imes 10^{-9}$	$4.44  imes 10^{-8}$	< 0.05
AL121944.1	rs34466819 (cis)	$1.19 \times 10^{-13}$	$7.83  imes 10^{-21}$	$6.05  imes 10^{-9}$	$2.25  imes 10^{-5}$	< 0.05
AC104129.1	rs4449693 (cis)	$1.04 \times 10^{-9}$	$2.59\times10^{\text{-}54}$	$1.34  imes 10^{-8}$	$5.81  imes 10^{-6}$	< 0.05
THOC7	rs832190 (cis)	$4.21  imes 10^{-10}$	$1.80 \times 10^{-37}$	$1.97  imes 10^{-8}$	$7.16  imes 10^{-2}$	< 0.05
OR2B7P	rs9393892 (cis)	$1.11 \times 10^{-13}$	$1.46  imes 10^{-17}$	$2.12  imes 10^{-8}$	$4.14  imes 10^{-5}$	< 0.05

 Table S3. Confirmation of BTN3A2 as a schizophrenia risk gene by integrative analysis (SMR) of schizophrenia GWAS (CLOZUK + PGC GWAS) and PsychENCODE brain PFC eQTL

<sup>b</sup>*P*-values for eQTL SNPs from PsychENCODE brain PFC [7].

Gene	Supporting SNP ( <i>trans</i> or <i>cis</i> )	PGWAS	PeQTL	P <sub>SMR</sub>	Pheidi	Tissue type
BTN3A2	rs9393703 (cis)	4.10×10 <sup>-26</sup>	1.24×10 <sup>-13</sup>	1.28×10 <sup>-9</sup>	0.07	Brain Amygdala
BTN3A2	rs3047288 (cis)	4.51×10 <sup>-27</sup>	3.17×10 <sup>-16</sup>	7.57×10 <sup>-11</sup>	< 0.05	Brain Anterior Cingulate Cortex
BTN3A2	rs3047288 (cis)	4.51×10 <sup>-27</sup>	5.80×10 <sup>-17</sup>	3.84×10 <sup>-11</sup>	< 0.05	Brain Caudate Basal Ganglia
BTN3A2	rs9393703 (cis)	4.10×10 <sup>-26</sup>	8.62×10 <sup>-16</sup>	1.53×10 <sup>-10</sup>	< 0.05	Brain Cerebellar Hemisphere
BTN3A2	rs9379853 (cis)	4.10×10 <sup>-26</sup>	9.85×10 <sup>-20</sup>	5.48×10 <sup>-12</sup>	< 0.05	Brain Cerebellum
BTN3A2	rs3047288 (cis)	4.51×10 <sup>-27</sup>	5.44×10 <sup>-13</sup>	2.04×10 <sup>-9</sup>	0.05	Brain Cortex
BTN3A2	rs3047288 (cis)	4.51×10 <sup>-27</sup>	$1.84 \times 10^{-14}$	4.26×10 <sup>-10</sup>	< 0.05	Brain Frontal Cortex
BTN3A2	rs9393703 (cis)	4.10×10 <sup>-26</sup>	$1.20 \times 10^{-12}$	3.69×10 <sup>-9</sup>	0.10	Brain Hippocampus
BTN3A2	rs72841536 (cis)	3.94×10 <sup>-27</sup>	6.78×10 <sup>-13</sup>	2.24×10 <sup>-9</sup>	< 0.05	Brain Hypothalamus
BTN3A2	rs28551159 (cis)	4.79×10 <sup>-27</sup>	2.82×10 <sup>-17</sup>	2.94×10 <sup>-11</sup>	< 0.05	Brain Nucleus Accumbens Basal Ganglia
BTN3A2	rs3047288 (cis)	4.51×10 <sup>-27</sup>	9.16×10 <sup>-14</sup>	8.81×10 <sup>-10</sup>	< 0.05	Brain Putamen Basal Ganglia
BTN3A2	rs71557332 (cis)	2.63×10 <sup>-25</sup>	8.93×10 <sup>-15</sup>	5.12×10 <sup>-10</sup>	0.32	Brain Spinal Cord Cervical
BTN3A2	rs28551159 (cis)	4.79×10 <sup>-27</sup>	9.92×10 <sup>-10</sup>	1.07×10 <sup>-7</sup>	< 0.05	Brain Substantia Nigra
BTN3A2	rs58825580 (cis)	3.02×10 <sup>-27</sup>	1.36×10 <sup>-23</sup>	2.04×10 <sup>-13</sup>	0.08	Brain Pituitary

Table S4. SMR integrative analysis of CLOZUK + PGC GWAS data and GTEx 14 brain tissues eQTL data verified *BTN3A2* as a schizophrenia risk gene

<sup>b</sup>*P*-values for eQTL SNPs from the GTEx data [2].

Gene	LBF	<i>P</i> -value	Supporting SNP (trans or cis)	Peqtl	PGWAS	Tissue type
BTN3A2	14.99	6.55×10 <sup>-6</sup>	rs9379858 (cis)	2.38×10 <sup>-13</sup>	3.25×10 <sup>-25</sup>	Brain Amygdala
BTN3A2	7.44	4.20×10 <sup>-6</sup>	rs9379858 (cis)	5.68×10 <sup>-14</sup>	3.25×10 <sup>-25</sup>	Brain Anterior Cingulate Cortex
BTN3A2	14.30	3.03×10 <sup>-6</sup>	rs13207082 (cis)	1.10×10 <sup>-6</sup>	8.16×10 <sup>-39</sup>	Brain Caudate Basal Ganglia
BTN3A2	15.14	2.49×10 <sup>-6</sup>	rs9379858 (cis)	1.77×10 <sup>-15</sup>	3.25×10 <sup>-25</sup>	Brain Cerebellar Hemisphere
BTN3A2	7.54	1.86×10 <sup>-6</sup>	rs9379858 (cis)	7.12×10 <sup>-19</sup>	3.25×10 <sup>-25</sup>	Brain Cerebellum
BTN3A2	14.52	2.78×10 <sup>-6</sup>	rs13207082 (cis)	9.62×10 <sup>-7</sup>	8.16×10 <sup>-39</sup>	Brain Cortex
BTN3A2	7.18	3.54×10 <sup>-6</sup>	rs9379858 (cis)	3.34×10 <sup>-14</sup>	3.25×10 <sup>-25</sup>	Brain Frontal Cortex
BTN3A2	15.02	4.88×10 <sup>-6</sup>	rs9379858 (cis)	2.03×10 <sup>-12</sup>	3.25×10 <sup>-25</sup>	Brain Hippocampus
BTN3A2	15.01	5.02×10 <sup>-6</sup>	rs9379858 (cis)	6.78×10 <sup>-13</sup>	3.25×10 <sup>-25</sup>	Brain Hypothalamus
BTN3A2	14.52	3.49×10 <sup>-6</sup>	rs13207082 (cis)	2.66×10 <sup>-7</sup>	8.16×10 <sup>-39</sup>	Brain Nucleus Accumbens Basal Ganglia
BTN3A2	7.39	4.01×10 <sup>-6</sup>	rs9379858 (cis)	5.65×10 <sup>-13</sup>	3.25×10 <sup>-25</sup>	Brain Putamen Basal Ganglia
BTN3A2	7.53	5.68×10 <sup>-6</sup>	rs9379858 (cis)	5.28×10 <sup>-14</sup>	3.25×10 <sup>-25</sup>	Brain Spinal Cord Cervical
BTN3A2	7.59	7.31×10 <sup>-6</sup>	rs9379858 (cis)	5.25×10 <sup>-9</sup>	3.25×10 <sup>-25</sup>	Brain Substantia Nigra
BTN3A2	7.57	2.52×10 <sup>-6</sup>	rs9379858 (cis)	2.67×10 <sup>-23</sup>	3.25×10 <sup>-25</sup>	Brain Pituitary

 Table S5. Confirmation of *BTN3A2* as a schizophrenia risk gene by integrative analysis (Sherlock) of schizophrenia GWAS (CLOZUK + PGC GWAS) and GTEx brain tissues eQTL

<sup>b</sup>*P*-values for eQTL SNPs from the GTEx data [2].

 Table S6. SMR integrative analysis of CLOZUK + PGC GWAS data and GTEx 14 brain tissues eQTL data identified C4A as a schizophrenia risk gene

Gene	Supporting SNP (trans or cis)	<b>P</b> GWAS <sup>a</sup>	<b>P</b> eQTL <sup>b</sup>	<b>P</b> <sub>SMR</sub>	<b>P</b> HEIDI	Tissue type
C4A	rs3129956 (cis)	2.95×10 <sup>-29</sup>	3.46×10 <sup>-11</sup>	1.15×10 <sup>-8</sup>	0.23	Brain Amygdala
C4A	rs1794282 (cis)	$1.58 \times 10^{-28}$	4.34×10 <sup>-9</sup>	2.13×10 <sup>-7</sup>	0.42	Brain Anterior Cingulate Cortex
C4A	rs9268158 (cis)	2.64×10 <sup>-29</sup>	$1.52 \times 10^{-13}$	6.74×10 <sup>-10</sup>	0.09	Brain Caudate Basal Ganglia
C4A	rs9268158 (cis)	2.64×10 <sup>-29</sup>	8.70×10 <sup>-13</sup>	1.61×10 <sup>-9</sup>	< 0.05	Brain Cerebellar Hemisphere
C4A	rs1265905 (cis)	3.18×10 <sup>-30</sup>	7.64×10 <sup>-17</sup>	1.65×10 <sup>-11</sup>	< 0.05	Brain Cerebellum
C4A	rs1265905 (cis)	3.18×10 <sup>-30</sup>	5.60×10 <sup>-9</sup>	$2.08 \times 10^{-7}$	< 0.05	Brain Frontal Cortex
C4A	rs625421 (cis)	8.10×10 <sup>-31</sup>	6.49×10 <sup>-9</sup>	2.16×10 <sup>-7</sup>	< 0.05	Brain Hippocampus
C4A	rs9268158 (cis)	2.64×10 <sup>-29</sup>	$1.42 \times 10^{-8}$	4.13×10 <sup>-7</sup>	0.69	Brain Hypothalamus
C4A	rs3129855 (cis)	8.91×10 <sup>-29</sup>	2.97×10 <sup>-13</sup>	1.05×10 <sup>-9</sup>	0.36	Brain Pituitary

<sup>b</sup>*P*-values for eQTL SNPs from the GTEx data [2].

### References

- [1] Colantuoni C, Lipska BK, Ye T, et al. Temporal dynamics and genetic control of transcription in the human prefrontal cortex. Nature 2011;478:519-23.
- [2] GTEx Consortium. The Genotype-Tissue Expression (GTEx) project. Nat Genet 2013;45:580-5.
- [3] Ramasamy A, Trabzuni D, Guelfi S, et al. Genetic variability in the regulation of gene expression in ten regions of the human brain. Nat Neurosci 2014;17:1418-28.
- [4] 1000 Genomes Project Consortium, Abecasis GR, Altshuler D, et al. A map of human genome variation from population-scale sequencing. Nature 2010;467:1061-73.
- [5] Pardiñas AF, Holmans P, Pocklington AJ, et al. Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. Nat Genet 2018;50:381-89.
- [6] He X, Fuller CK, Song Y, et al. Sherlock: detecting gene-disease associations by matching patterns of expression QTL and GWAS. Am J Hum Genet 2013;92:667-80.
- [7] Wang D, Liu S, Warrell J, et al. Comprehensive functional genomic resource and integrative model for the human brain. Science 2018;362.